CHAPTER IX
IN-SILICO SCREENING OF LIGANDS FOR THEIR ANTICANCER ACTIVITY USING MOLECULAR DOCKING STUDIES

1. Introduction

Cancer therapies target cell cycle-based mechanism which imitates the body’s natural process in order to prevent the growth of cancer cells. This approach can limit damages to normal cells and the associated side effects caused by usual chemotherapeutic agents. It has been unambiguously proved that the genome contains information in two forms: genetic and epigenetic. The genetic information provides the scheme of assembling of all the proteins necessary to create a living organism, while the epigenetic information directs when, where, and how the genetic information must be used. Epigenetics has just moved to the forefront of studies describing various processes such as transcriptional regulation, chromatin structure, genome integrity and tumorigenesis.1

1.1 Role of histone deacetylases in human cancer

Inhibition of histone deacetylases (HDAC) enzymes is emerging as an innovative and important approach for the treatment of cancer. HDAC inhibitor interacts with chromosomes in the cancer cell and stops their growth. Acetylation levels are the outcome of the activities of histone acetyltransferase (HAT) and HDAC. They have a critical role in the regulation of gene expression. HATs induce histone acetylation. It is coupled with gene transcription. HDAC induces histone hypo-acetylation. It is associated with gene silencing. Altered appearance and mutations of genes that encode HDACs have been correlated to tumour growth. These changes induce the abnormal transcription of genes which regulates the important cellular functions such as cell proliferation, cell-cycle regulation and apoptosis.2 HDACs are expressed in colon, breast,3 lung cancers, hepato-cellular carcinoma, gastric cancer,4 pancreatic tumour tissue, metastatic melanoma, acute myeloid leukemia (AML), acute lymphocytic leukemia (ALL) as well as in childhood tumors of the nervous system, such as neuroblastoma.5–8 As a result, HDACs have become one of the most potential therapeutic targets for cancer treatment.

Four families of human HDACs have been described and categorized into four classes. Class I includes HDAC 1–3 and 8; class II includes HDAC 4–7, 9 and 10; class III includes sirtuins 1–7; and class IV includes HDAC 11. The HDACs belonging to classes I, II and IV have Zn2+ as a cofactor and class III enzymes are characterised for being NADP-dependent.9,10
1.2. Role of histone deacetylase 8 in oncogenesis

Histone deacetylase 8 (HDAC8) is a class I HDAC. It is considered as a therapeutic target in various diseases, including cancer, X-linked intellectual disability and parasitic infections. This enzyme also deacetylates non-histone proteins. This enzyme is a major ‘epigenetic player’ and it is linked to uncontrolled expression or interaction with transcription factors important for tumourigenesis. In the parasite *Schistosoma mansoni* and in viral infections, HDAC8 is a novel target to suppress infection. The current challenge is to develop potent selective inhibitors that would solely target HDAC8 with less adverse effects.

1.3. Role of DNA methyltransferases in human cancer

Contemporary works have revealed how DNA methylation and chromatin structure are correlated at the molecular level and how methylation plays a direct role in tumourigenesis and genetic disease. A great deal of information is available concerning the cellular methylation mechanism. DNA methyltransferases are the main components of larger complexes which are actively involved in transcriptional control and chromatin structure modulation. These findings enhance our perspective about the roles of DNA methylation in novel therapies to prevent or repair these defects. Variation in DNA methylation process can cause changes in gene transcription patterns and can also promote mutational events. Abnormal DNA methylation patterns also play an essential role in the development of cancer. Cancer cells show an overall decrease in the level of genomic cytosine methylation, mainly in centromeric satellite and other repeated sequences. This genome hypomethylation is believed to be linked to genome instability resulting in a variety of chromosomal defects and tumorigenesis. In addition to the loss of methylation at satellite sequences, cancer cells acquire methylation in normally un-methylated promoter regions. In many cancer types, one can see tumour suppressor genes epigenetically silenced by hypermethylation.

Azacytidine (Vidaza; Celgene) and Decitabine (5 aza 2′ deoxycytidine) (Dacogen; SuperGen) are the two DNMT inhibitors (azanucleosides) currently approved by the US Food and Drug Administration (FDA). These two types of drugs are the first molecules that have been characterised as the archetypal DNMT inhibitors and the only epidergics that have been approved for the treatment of patients with AML and myelodysplastic syndrome (MDS). Azacytidine has also been approved by the FDA and the European Medicines Agency (EMA) for the use against chronic myelomonocytic leukemia (CMML). Even though several new inhibitors of DNMT have been identified, none of them can substitute azacytidine and decitabine.

1.4. Role of ribonucleotide reductase subunit, R2 in human cancer

Ribonucleotide reductase (RR), a distinct enzyme which is responsible for reducing ribonucleotides to deoxyribonucleotides, the building blocks required for DNA replication and repair. Dysregulated RR activity is allied with genomic instability, malignant transformation and cancer development. The use of RR inhibitors, either as a single agent or combined with other therapies, has proven to be a promising approach for treating solid tumors and hematological malignancies.
RR subunit, R2 regulates catalytic action of the enzyme to provide DNA synthesis via reduction. It has been constantly investigated as anticancer drug target for design and discovery of its inhibitors. RR plays an important role in determining cell fate and increased expression. Activity of human RR has been associated with many types of cancer. Failure to control the levels and balance of dNTPs can lead to mutagenesis, carcinogenesis and even cell death. After the occurrence of transformation, cancer cells require increased RR activity to meet the demand for dNTPs that are needed to support their rapid proliferation. However, accumulating evidences also indicate that RR subunits play biological roles in promoting cancer development that are distinct from their role in forming the active RR holoenzyme to supply dNTPs.

RRM2 has been proposed as an oncoprotein. It not only can cooperate with a variety of oncogenes to promote cellular transformation but also can enhance the invasive potential of cancer cells. Discovering novel RR inhibitors is a challenge for improving cancer therapies in the future. Investigation of RR expression and action in clinical samples will provide a logical platform for modified cancer therapy using RR inhibitors.

1.5. Molecular docking studies

Molecular docking is a well-known computational technique which predicts the interaction energy between two molecules. Docking is frequently used to predict the binding affinity and to identify the bound conformation of ligands with the protein targets. It can provide information regarding the intermolecular interactions between the ligands and the targets. This in turn can be used for the development of novel drugs. Considerable efforts have been made to improve the methods used to predict docking, due to their significance in the biological and pharmaceutical fields.

This chapter portrays the molecular docking studies of the ligands synthesized by us with different enzyme targets.

2. Materials and methods

Molecular docking studies were performed using the Glide program. Glide means Grid-Based Ligand Docking With Energetics. The Maestro (version 10.4) was used as graphical user interface. It was used to set up and execute the docking protocol and also for analysis of the docking results. The three-dimensional structures of all the proteins were retrieved from the Protein Data Bank (PDB). The ligands under study were docked into the binding sites of the proteins. PyMOL was used for the three-dimensional visualization of proteins to superimpose the structures and for the characterisation of ligand–protein interactions. PyMOL supports different macromolecular structural representations such as ribbon structure, wired structure, ball and sick model, surface model, cartoon model and so on. The selection of the best pose was done based on the interaction energy between the ligand and the protein. The results were analyzed by a statistical scoring function which converts interacting energy into numerical values called as the glide score.
Docking methodology consists of four basic steps:

1. Preparation of protein
2. Ligand preparation
3. Receptor grid generation
4. Ligand docking

2.1. Preparation of protein

Protein preparation was done by protein preparation wizard. The PDB structure file consists of different metal ions, hetero atoms, cofactors, water molecules, co-crystallized ligands and so on. Therefore, the PDB structure files as such are not suitable for molecular modelling calculations. Schrodinger contains all the tools to prepare appropriate protein for its calculations. The appropriate protein structure is refined through optimization and minimization\textsuperscript{22} during protein preparation. Minimization process is coordinated by a defined RMSD of 0.30 Å and force fields OPLS 2005.\textsuperscript{23} Protein preparation wizard prepares accurate protein structure for docking. The different procedures used for protein preparation include pre-processing, review, modification, refinement and minimization.

2.2. Ligand preparation

The ligand preparation was done by ligand preparation wizard, which is also known as LigPrep. The ligands were prepared using LigPrep with parameter force field, OPLS 2005 in Schrodinger. This is used for the conversion of 2D molecules to 3D structures. The steps involved in the process are convert to the structure format (sd format), select the structures, add hydrogen atoms, remove unwanted molecules, neutralize charged groups, generate ionization states, generate tautomers, filter the structures, generate alternative chiralities, generate low-energy ring conformations, remove problematic structures, optimize the geometries and finally convert the output file. Many of these steps are optional and are controlled by selecting options in the LigPrep panel or by specifying command-line options. Prepared ligands are saved with maegz extension or in maestro format and all the ligands are available in the project table. Now, the ligands are ready for the docking process. We can export the project table as csv format and keep the details of ligands.

2.3. Receptor grid generation

The prepared protein was loaded into maestro environment and the active site was defined using Glide grid generation wizard. For grid generation, we want to specify the active site residue. Certain proteins may have metal ions or may possess one or more ligands. These already existing ligands or metal ions can be selected as active site residues. If no ligand or metal ion is present, then active site amino acids are used for grid generation. Grid centre was defined for the active site and box sizes were set to 15 Å.

Receptor grid should be generated by using Receptor Grid Generation feature of maestro. The van der Waals radius scaling involves a scaling factor as 1.0 and partial charge cut off as 0.25. The shapes and properties of receptor molecule were represented on a grid.
The prepared protein structure was used for receptor grid generation. It determines the position and size of the active site. Receptor grid generation was selected from the application menu and it includes three fields such as receptor, site and constraints. Receptor section defines the respective target molecule and picks the ligand molecule. Several constraints can be applied in grid generation, which are positional constraint, H-bond constraint, metal constraint and hydrophobic constraint. The generated grid was used for docking.

2.4. Ligand docking

This was carried out using Glide dock. Ligands were kept flexible, while the protein was rigid and docking started with extra precision mode (XP mode). The docking calculation generated a few poses for each ligand. The selection of the best pose was done based on the interaction energy between the ligand and the protein. The least negative XP g-score (negative value) indicates the best binding score for the ligand with protein.

3. Results and discussion

Crystal structures of HDAC8 (PDB ID: 1T69), DNA methyltransferase (PDB ID: 2QRV) and RR M2 subunit (PDB ID: 2UW2) (Fig. 1) were downloaded from PDB.

The following novel ligands synthesized by us were selected for performing molecular docking studies:

1. N-Methyl-2-((3-methyl-5-oxo-1-phenyl-4,5-dihydro-1H-pyrazol-4-yl)(phenyl)methylene)-N-phenylhydrazinecarbothioamide (H2L1)
2. 2-[2-(Diphenylmethylidene)hydrazine]-N-methyl-N-phenylhydrazine-1-carbothioamide (H2L2)
3. 3-Formylchromone N(4)-methyl-N(4)-phenylthiosemicarbazone (HL3)
4. 2-2-[(4-Hydroxy-3-methoxyphenyl)methylidene]hydrazinecarbonothioyl]-N-methyl-N-phenylhydrazine-1-carbothioamide (H2L4)
5. 2-2-[(3,4-Dimethoxyphenyl)methylidene]hydrazinecarbonothioyl]-N-methyl-N-phenylhydrazine-1-carbothioamide (H2L5)

The compounds were characterised by spectroscopic analysis (electronic, FT-IR, FT-Raman and 1H NMR). The compounds 1 and 5 were further characterised by single-crystal XRD and others by powder XRD. The details of identification of selected compounds are presented in the earlier chapters of the thesis (Chapters III–VII). For interpreting the docking results, mainly four parameters were considered: Glide score, Glide energy, H-bonds and non-bonded interactions (van der Waals and electrostatic). Based on these, the binding affinity of the ligand towards the receptor has been discussed. The result demonstrates that all the compounds could effectively be docked into the same binding site of each protein. Details of binding affinity of each ligand towards different receptors are discussed in detail below.

1. N-Methyl-2-((3-methyl-5-oxo-1-phenyl-4,5-dihydro-1H-pyrazol-4-yl)(phenyl)methylene)-N-phenylhydrazinecarbothioamide (H2L1)

The amino acid residues in the active site of RR M2 subunit are Glu 232, Gly 233, Le 234, Phe 236, Ser 237, Val 327, Arg 330, Leu 331, Glu 334, Asp 271, Glu 260, Cys 270, Leu 268, Gly 267, Glu 266, Arg 264, Phe 244, Met 350 and Ser 263.
The compound was docked deeply into the active site region making interactions with the residues Asp 271 and Ser 263. The compound formed three hydrogen bonded contacts through Ser 263 and Asp 271 with the NH functionality and Ser 263 with the SH group. Certain very prominent interactions between the compound and the active site residues Gly 703, Ar 887, Glu 660 and Phe 636 of DNA methyltransferases were observed. The other residues that were found in the locality of this docked ligand were Leu 635, Phe 636, Asp 637, Gly 638, Le 639, Ala 640, Thr 641, Gly 642, Glu 752, Trp 889, Ser 888, Arg 887, Ser 665, Cys 662, Gly 660, Ser 659, Gly 703, Ser 704, Pro 705, Cys 706, Cys 662, Val 661, Glu 660, Asp 682, Val 683, Arg 684 and Leu 726. The amino acid residues in the active site of HDAC8 were Arg 37, Le 34, Lys 33, Tyr 306, Gly 305, Gly 304, Gly 303, Asp 178, He 180, Gly 151, Phe 152, Cys 153, Asp 101, Asp 267, Phe 206, Phe 208, His 143, Gly 206, Met 274 and Trp 141. H$_2$L1 binds in the pockets defined by Asp 101 and Gly 206 of HDAC8. It is found that H$_2$L1 is able to bind all the selected targets (Fig. 1). The
Docking results for the inhibitory compound, H₂L₁ are reported in Table 1. Fig. 2 shows the docked pose of the inhibitor, H₂L₁ inside the binding pockets of the proteins (a) PDB ID: 2UW2, (b) PDB ID: 2QRV and (c) PDB ID: 1T69.

**Table 1.** Glide score and Glide energy of H₂L₁

<table>
<thead>
<tr>
<th>Protein</th>
<th>Glide Score</th>
<th>Glide Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribonucleoside diphosphate reductase M2 subunit (PDB ID: 2UW2)</td>
<td>−7.047803</td>
<td>−47.1708</td>
</tr>
<tr>
<td>DNA methyltransferase (PDB ID: 2QRV)</td>
<td>−4.32107</td>
<td>−18.399939</td>
</tr>
<tr>
<td>Histone deacetylase 8 (PDB ID: 1T69)</td>
<td>−5.034831</td>
<td>−31.793303</td>
</tr>
</tbody>
</table>
2. 2-(2-(Diphenylmethylidene)hydrazine)-N-methyl-N-phenylhydrazine-1-carbothioamide (H₂L₂)

The compound, H₂L₂ binds in the pocket defined by Asp 271 and Ser 263 of RR M2 subunit. It binds in the pocket defined by Val 661 and Glu 663 of DNA methyltransferase. It binds in the pocket defined by Asp 271 of HDAC8. The docking results for H₂L₂ are reported in Table 2. Fig. 3 shows the docked pose of the inhibitor, H₂L₂ inside the binding pockets of the proteins (a) PDB ID: 2QRV and (b) PDB ID: 1T69.

![Docked pose of the inhibitor, H₂L₂ inside the binding pockets of proteins. (a) PDB ID: 2QRV and (b) PDB ID: 1T69.](image)

Table 2. Glide score and Glide energy of H₂L₂

<table>
<thead>
<tr>
<th>Protein</th>
<th>Glide Score</th>
<th>Glide Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribonucleoside diphosphate reductase M2 subunit (PDB ID: 2UW2)</td>
<td>−5.13638</td>
<td>−47.9122</td>
</tr>
<tr>
<td>DNA methyltransferase (PDB ID: 2QRV)</td>
<td>−3.11166</td>
<td>−44.3094</td>
</tr>
<tr>
<td>Histone deacetylase 8 (PDB ID: 1T69)</td>
<td>−3.03165</td>
<td>−35.7943</td>
</tr>
</tbody>
</table>

3. 3-Formylchromone N(4)-methyl-N(4)-phenylthiosemicarbazone (HL3)

The compound docked deeply into the active site region of ribonucleoside diphosphate reductase M2 subunit, making three hydrogen bonding interactions with the residues Arg 330, Glu 334, Asp 271 and Ser 263. The compound docked deeply into the active site region of DNA methyltransferase, making interactions with the residues Glu 660, Phe 636, Gly 703, Trp 889 and Ser 263. It docked deeply into the active site region of HDAC8, making hydrogen bonded interactions with the residues His 143, He 180, Asp 178, Asp 101, Gly 151 and Ser 263. The docking results for HL3 are reported in Table 3. Fig. 4 shows the docked pose of the inhibitor HL3 inside the binding pockets of the proteins (a) PDB ID: 2UW2, (b) PDB ID: 2QRV and (c) PDB ID: 1T69.
Table 3. Glide score and Glide energy of HL3

<table>
<thead>
<tr>
<th>Protein</th>
<th>Glide Score</th>
<th>Glide Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribonucleoside diphosphate reductase M2 subunit (PDB ID: 2UW2)</td>
<td>−6.438</td>
<td>−43.039</td>
</tr>
<tr>
<td>DNA methyltransferase (PDB ID: 2QRV)</td>
<td>−6.222</td>
<td>−45.642</td>
</tr>
<tr>
<td>Histone deacetylase 8 (PDB ID: 1T69)</td>
<td>−8.148</td>
<td>−34.238</td>
</tr>
</tbody>
</table>

Fig. 4. Docked pose of the inhibitor, HL3 inside the binding pockets of proteins. (a) PDB ID: 2UW2, (b) PDB ID: 2QRV and (c) PDB ID: 1T69.
4. 2-{2-[(4-Hydroxy-3-methoxyphenyl)methylidene]hydrazinecarbonothioyl}-N-methyl-N-phenylhydrazine-1-carbothioamide (H$_2$L4)

The compound $H_2L4$ bound in the pocket defined by Asp 271, Ser 237 and Tyr 323 of RR. $H_2L4$ stabilized within the active site of DNA methyltransferase through hydrogen bonding interactions with Arg 887, Val 661 and Glu 663 residues. It bound in the pocket defined by Asp 101, Phe 208, Gly 206 of HDAC8. The docking results for the inhibitory compound $H_2L4$ are reported in Table 4. Fig. 5 shows the docked pose of the inhibitor $H_2L4$ inside the binding pocket of protein (PDB ID: 2UW2).

![Docked pose](image)

**Fig. 5.** Docked pose of the inhibitor, $H_2L4$ inside the binding pocket of protein (PDB ID: 2UW2).

**Table 4.** Glide score and Glide energy of $H_2L4$

<table>
<thead>
<tr>
<th>Protein</th>
<th>Glide Score</th>
<th>Glide Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribonucleoside diphosphate reductase M2 subunit (PDB ID: 2UW2)</td>
<td>−8.13072</td>
<td>−56.5909</td>
</tr>
<tr>
<td>DNA methyltransferase (PDB ID: 2QRV)</td>
<td>−3.99925</td>
<td>−45.9494</td>
</tr>
<tr>
<td>Histone deacetylase 8 (PDB ID: 1T69)</td>
<td>−4.18525</td>
<td>−38.1185</td>
</tr>
</tbody>
</table>
5. 2-{2-[(3,4-Dimethoxyphenyl)methylidene]hydrazinecarbonothioyl}-N-methyl-N-phenylhydrazine-1-carbothioamide (H₂L₅)

The compound stabilized within the active site of RR through four prominent hydrogen bonding interactions: first between the oxygen atom of methoxy group and Arg330 residue, second and third interactions between NH group and Asp 271, while fourth interactions between NH group and Ser 263 residue. The compound H₂L₅ stabilized within the active site of DNA methyltransferase through two prominent hydrogen bonding interactions. The first between the NH group and Asn 707 residue while the second between SH and Asn 707 residue. H₂L₅ bound in the pocket defined by Asp 708, Asn 707 of DNA methyltransferase. It bound in the pocket defined by Asp 101, Phe 208 of HDAC8. The docking results for the inhibitory compound H₂L₅ are reported in Table 5. Fig. 6 shows the docked poses of the inhibitor, H₂L₅ inside the binding pockets of the proteins (a) PDB ID: 2UW2 and (b) PDB ID: 1T69.

(a) PDB ID: 2UW2
(b) PDB ID: 1T69

Table 5. Glide score and glide energy of H₂L₅

<table>
<thead>
<tr>
<th>Protein</th>
<th>Glide Score</th>
<th>Glide Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribonucleoside diphosphate reductase M2 subunit (PDB ID: 2UW2)</td>
<td>−6.46082</td>
<td>−43.6959</td>
</tr>
<tr>
<td>DNA methyltransferase (PDB ID: 2QRV)</td>
<td>−3.99925</td>
<td>−3.11166</td>
</tr>
<tr>
<td>Histone deacetylase 8 (PDB ID: 1T69)</td>
<td>−3.60968</td>
<td>−36.5944</td>
</tr>
</tbody>
</table>
4. Conclusions

The results of this study can be helpful for the design and development of new compounds having enhanced inhibitory activity against several types of cancer. These compounds will be hopeful candidates for anticancer studies. The molecular docking studies revealed that all these compounds have affinity towards the active site of RR M2 subunit with Glide score ranging from $-5.13638$ to $-8.13072$. They can further be explored as selective RR inhibitors. It is also found that the compound 3-formylchromone N(4)-methyl-N(4)-phenylthiosemicarbazone (HL3) has affinity towards all the selected target with good binding score and is a good candidate for anticancer studies. This can further be validated in wet lab studies.
References