

CHAPTER 1

INTRODUCTION

Finding the ultimate HIV cure remain a challenging tasks for decades. Various active compounds have been tested against various components of the virus in effort to halt the virus development in infected host. More than a decade ago, El-Mekkawy and co-workers (el-Mekkawy et al.1998) have tested active compounds from Ganoderma Lucidum against HIV proliferation and HIV protease. They have successfully identified several compounds with reasonable inhibitory activity against HIV protease. Today, it is a common practice to utilize structural bioinformatics tools to complement conventional wet lab research especially in the field of drug discovery. Structural bioinformatics has developed by leaps and bounds in the last decades in terms of technology and computing power. Emergence of this field has opened new possibilities and ways to learn molecular interaction between molecules.

From HIV drug development perspective, a combination structural bioinformatics tools namely molecular dynamic simulation and molecular docking has led to the discovery of the first HIV-integrase inhibitor. Molecular dynamic simulation and molecular docking has successfully pointed out a previously unidentified region in HIV-integrase, this region in later development became a potent target for HIV-integrase inhibitor (Schames, Henchman et al. 2004).

This research report aims to study compounds with HIV inhibitory activity identified by el-Mekkawy and co-workers (el-Mekkawy et al.1998) from structural bioinformatics perspective using molecular docking. Compounds with HIV inhibitory activity were studied using two molecular docking approaches (reverse molecular docking

and molecular docking). Outcomes from molecular docking study were compared and correlated to the previous experimental findings to elucidate new facts and to compare consistency.

This report aims to achieve these objectives

- To gather new information on HIV-1 and Ganoderic acid A, B, C1, H and interactions with HIV-PR by consulting molecular docking tools.
- To find correlation between information obtained from molecular docking and information from previous study by el-Mekkawy and co-workers (el-Mekkawy et al. 1998)

CHAPTER 2

LITERATURE REVIEW

2.1 Ganoderma Lucidum

Ganoderma lucidum has an impressively long medicinal history, some papers claimed that it is a culmination of the knowledge and wisdom of the east and west for 5,000 years, some even went to claim it has totally no side effects (Matsomoto, 1979). A typical G. lucidum is shown in figure 2.1. Boh and co-workers described G. lucidum as a wood-degrading basidiomycete with numerous pharmacological effects Boh et al (2007).

Triterpenoids are common chemical constituent of G.lucidum. However due to their unique properties triterpenoids importance is established not only within the species but also extends to the chemotaxonomy of Ganoderma genus (Cheng, Yue et al. 2010). According to Boh and co-workers triterpenoids possesses wide range of pharmaceutical activities such as hepatoprotective, anti-hypertensive, hypocholesterolemic and anti-histaminic effects, anti-tumor and anti-angiogenic activity, effects on platelet aggregation and complement inhibition(Boh et al. 2007).

Another pharmaceutically important constituent of G. lucidum is polysaccharides. Polysaccharides are structurally diverse macromolecules which play a wide-range of biological functions, it also possesses a wide-range physicochemical properties. Important bioactive polysaccharides of G. lucidum includes β -1-3 and β -1-6-D glucans (Cheng, Yue et al. 2010). Anti-tumour compounds are glycoproteins(combination of polysaccharides and proteins), heteropolysaccharides and ganoderans A, B and C(Lindequist 1995)



Figure 2.1:(Typical G.lucidum)

Aside from its popular pharmaceutically important constituents, general nutritional composition G. lucidum are fat, carbohydrate and fiber with detailed composition as follow, folin-positive material (68.9%), glucose (11.1%), protein (7.3%) and metals (10.2%) (K, Mg and Ca are the major components) (Babu and Subhasree 2008).

2.2 Human Immunodeficiency Virus (HIV)

HIV is a type of retrovirus where DNA is synthesized from RNA using reverse transcriptase enzyme, this copy of DNA is transported and inserted into the host cell genome (Kartikeyan, bharmal et al. 2007). The nature of this infection mode allows HIV virus persists within the host for years. The deadliest fact about the virus is the virus's DNA is actually become part of and treated as the host genome thus making the host defense system useless. Volberding and co-workers classified HIV to two types, HIV-1 and HIV-2 (Volberding, Sande et al. 2008). According to WHO statistics report, in 2007 approximately 3.3 million people were estimated living with HIV infection with 22 million were identified in sub-Saharan Africa, 4 million in South East Asia. The massive number of infection is a strong indicator of how difficult it is to deal with this virus. The majority of the world's infections are indentified as HIV-1 group M virus infection, of which this

group is further classified to 10 different subtypes (Volberding, Sande et al. 2008). HIV invades its host in a very structural manner as depicted in Figure 2.2.

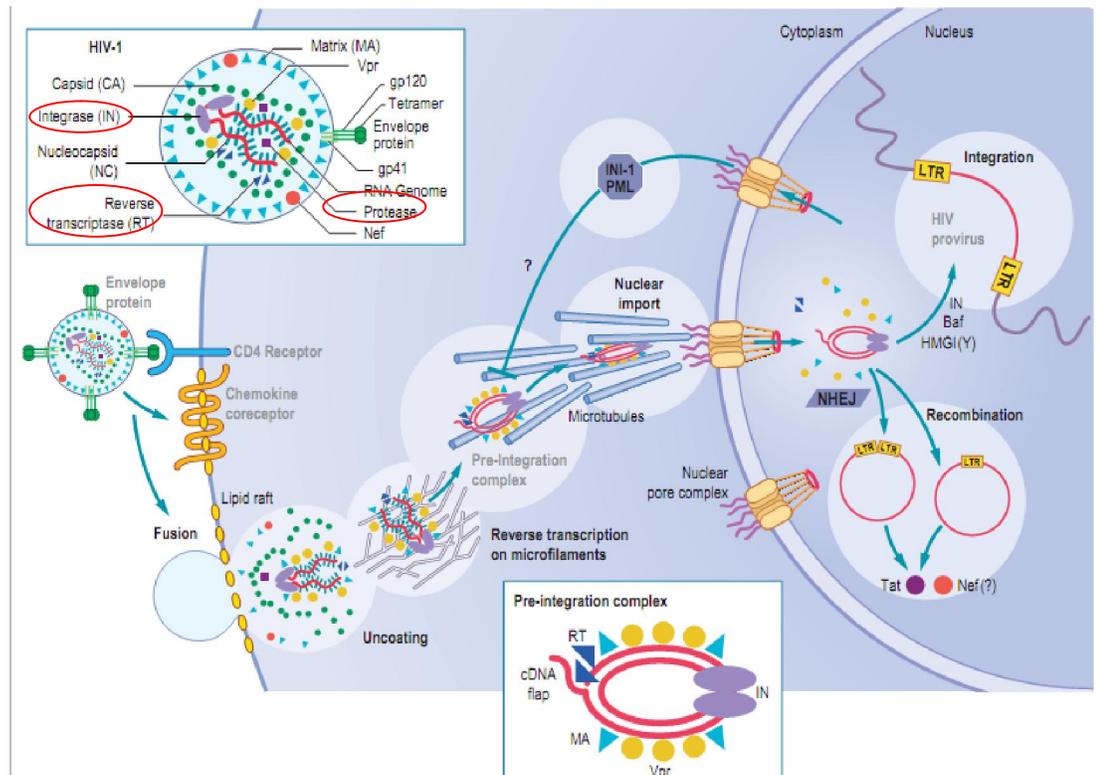


Figure 2.2 : (Schematic description of HIV infection (Volberding, Sande et al. 2008)

The schematic clearly shows there are three major enzymes in the virus, protease, reverse transcriptase, and integrase (highlighted in red circle) and a bunch of other cellular apparatus. The enzymes took the center stage because logically they play crucial role in the virus life cycle. It is not hard to deduce the function of each enzyme because their names naturally describe their function in the virus. Reverse transcriptase deals with the reverse transcription process, integrase deals with integration of the virus genetic materials and protease deals with the processing of the virus polyproteins. All three enzymes are decent candidate for potential drug target because each plays vital role in the virus life cycle.

Therefore knocking off either one of these enzymes might bring chaos towards the life cycle of the virus and hopefully in the long run will help eliminate the virus from human population.

2.3 HIV related Ganoderic acids

Holding the reputation as the most valuable crude drug source, it was not long before researchers started to turn to *G. lucidum* for HIV cure. Researchers started to test various compounds from *G. lucidum* against various components of the virus (enzymes, receptors etc). Triterpenoids are among the most pharmacologically active constituents of *G. lucidum* (Boh, Berovic et al. 2007), arguably compounds from triterpenoids family might yield the best results when tested against the virus components.

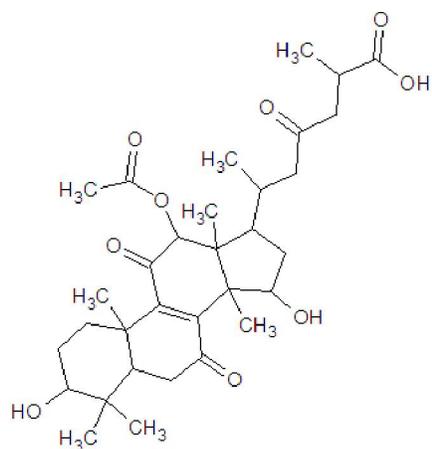
el-Mekkawy and co-workers (el-Mekkawy et al. 1998) tested and identified thirteen compounds with inhibitory activities against HIV protease (HIV-PR) and proliferation of HIV-1 from *G. lucidum*. Five of those compounds belong to Ganoderic acid family namely Ganoderic acid A, B, C1, H and , complete list of the compounds is provided in Table 2.1. Ganoderic acid is a type of triterpenoids, to be exact it is actually a highly oxygenated C₃₀ lanostane-type triterpenoids (Xu, Zhao et al. 2010).

Table 2.1 shows compounds belonging to Ganoderic acid family exhibited inhibitory activity against HIV-PR with IC₅₀ value below 0.20 mM apart from Ganoderic acid A which have surprisingly higher IC₅₀ value (>1.0 mM). Structural representation of Ganoderic acid B, C1, H and is shown in Figure 2.3.

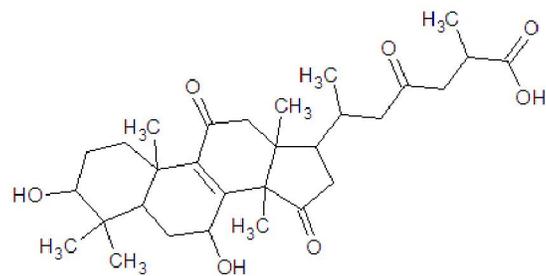
Table 2.1: Inhibitory Activities of Compounds from *Ganoderma lucidum* against Protease and Proliferation of HIV-1 (el-Mekkawy, Meselhy et al. 1998)

Compound	HIV-PR IC ₅₀ (mM)
Ganoderic acid	0.19
Ganoderic acid A	>1.0
Ganoderic acid B	0.17
Ganoderic acid C1	0.18
Ganoderic acid H	0.20
Ganoderiol A	0.23
Ganoderiol B	0.17
Ganoderiol F	0.32
Ganodemanontriol	>1.0
Ergosterol	>1.0
Ergosterol peroxide	>1.0
Crevisterol	>1.0
3,5-Dihydroxy-6-methoxy ergosta-7,22-diene	0.18

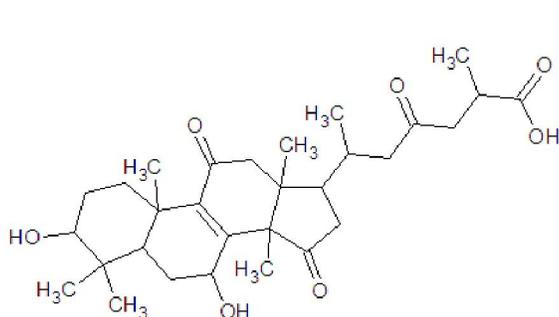
IC₅₀, concentration for half maximum inhibition



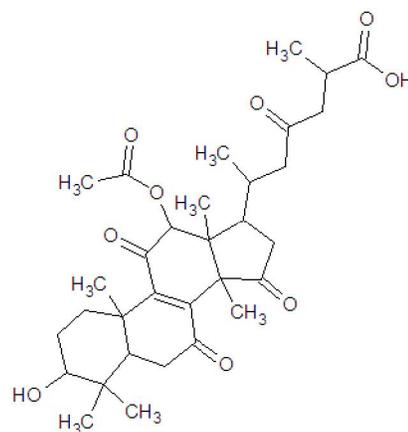
A. Ganoderic acid $C_{32}H_{46}O_9$



B. Ganoderic acid B $C_{30}H_{44}O_7$



C. Ganoderic acid C1 $C_{30}H_{42}O_7$



D. Ganoderic acid H $C_{32}H_{44}O_9$

Figure 2.3A-D: Structural representation of Ganoderic acid B, C1, and H, respectively

(PubChem Compound 2011)

2.4. Molecular docking.

Molecular docking can be separated to two classes. The first is reverse molecular docking and the second is molecular docking. Reverse molecular docking can be described as the process of searching a small molecule-protein target over a large data base of potential protein targets. In contrast to standard molecular docking approach where molecule of interest is screened against various proteins from various and often unrelated databases,

reverse molecular docking screens the molecule against specific database. The reverse molecular docking terms rooted from the fact that this approach starts from the molecule of interest (usually ligands) whilst in conventional molecular docking usually starts with the potential targets (the proteins).

Molecular docking in general can be described in two stages. The first stage is the conformational search and the second stage is conformation's energy calculation. Different algorithms have been implemented to achieve different objectives in conformational stage. For fast conformational search, geometry based algorithm is normally employed. Programs such as DOCK and TarFisDock (Ewing, Makino et al. 2001) which scans large databases will find this type of algorithm relevant. For a more detailed conformational search, genetic algorithm or a combination of genetic algorithm with other conformational search approach usually implemented. Program such as AutoDock which emphasize on a more detailed molecular interaction will benefit from this type of algorithm(Morris, Goodsell et al. 1998).

Energy calculation stage on the other hand is little more conserved. Energy calculations in various molecular docking approaches are generally based on AMBER force field. Different modifications to the standard AMBER force field (shown in equation bellow) were made to achieve different objectives.

$$G = G_{vdw} + G_{hbond} + G_{elec} + G_{conform} + G_{tor} + G_{sol}$$

The first four terms are common molecular mechanics terms for dispersion/repulsion, hydrogen bonding, electrostatic and deviation from covalent geometry. The last two models the torsion angles and desolvation and hydrophobic effects. Faster energy calculation can be achieved by incorporating fewer energy terms, this approach is well demonstrated by DOCK where Energy terms is described by a combination

of Van der Waals terms and electrostatic interaction terms as shown in the following equation.

$$E_{inter} = \sum_{i=1}^{lig} \sum_{j=1}^{rec} \left(\frac{A_{ij}}{r_{ij}^a} - \frac{B_{ij}}{r_{ij}^b} + 332.0 \frac{q_i q_j}{D r_{ij}} \right)$$

each term is a double sum over ligand atoms *i* and receptor atoms *j*; *r_{ij}* is the distance between atom *i* in the ligand and atom *j* in the putative receptor protein; *A_{ij}* and *B_{ij}* are Van der Waals repulsion and attraction parameters, respectively; *a* and *b* are the Van der Waals repulsion and attraction exponents, respectively; *q_i* and *q_j* are point charges on atoms *i* and *j*; *D* is dielectric function; and 332.0 is the factor that converts the electrostatic energy into kcal/mol (Li, Gao et al. 2006)

A more detailed energy calculation is demonstrated by AutoDock. The AMBER energy terms for energy evaluation are described by combination of dispersion/repulsion, hydrogen bonding, torsion, electrostatics, and desolvation as shown in the following equation.

$$\begin{aligned} \Delta G = & \Delta G_{vdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \\ & + \Delta G_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ & + \Delta G_{elec} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{tor} N_{tor} \\ & + \Delta G_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)} \end{aligned}$$

Where the five *G* terms on the right-hand side are coefficients empirically determined using linear regression analysis from a set of protein-ligand complexes with known binding

constants. The summations are performed over all pairs of ligand atoms, i , and protein atoms, j , in addition to all pairs of atoms in the ligand that are separated by three or more bonds (Morris, Goodsell et al. 1998).

CHAPTER 3

METHODOLOGY

3.1 Reverse Docking.

Compound of interest were identified from previous work by el-Mekkawy and co-workers (el-Mekkawy et al. 1998). Four compounds with IC_{50} ranging from 0.17 mM to 0.20 mM were chosen as the compound of interest as shown in Table 3.1. Three dimensional structures of the compounds were obtained from PubChem database (<http://www.ncbi.nlm.nih.gov/pccompound>).

Table 3.1: list of compound of interest

Compound	HIV-PR IC_{50} (mM) (el-Mekkawy et al. 1998)
Ganoderic acid	0.19
Ganoderic acid B	0.17
Ganoderic acid C1	0.18
Ganoderic acid H	0.20

All four structures were submitted to TarFisDock server (<http://www.dddc.ac.cn/tarfisdock>) for target identification by reverse molecular docking. Viral infections were selected as the target criteria and default parameters defined in the server were used. Targets were selected based on crystal structure resolution and distribution in regard to compounds of interest.

3.2 Molecular Docking.

Hydrogen atoms, Gasteiger charges and torsion angles were computed and added to the compounds of interest (Ganoderic acid, B, C1, and H). Numbers of torsion angles assigned to compounds of interest are shown in Table 3.2. Water and ligand were removed from the macromolecules (1HVR and 1DIF). Hydrogen and Gasteiger charges were computed and added to the macromolecules. Active site of the target was highlighted and 50x60x60 Å grid box with 0.375 Å grid spacing was drawn to cover the active site of the target. Grid and molecular docking files were generated based on the target and compound of interest. Each docking trial was initiated with 100 runs. Population size, energy evaluations, mutation and crossover rates and local search probability were kept as default. Further improvements to docking parameters were made accordingly based on control docking results.

Table 3.2: torsion numbers of compounds of interest

Compound	Number of torsion detected by AutoDock
Ganoderic acid	11
Ganoderic acid B	9
Ganoderic acid C1	8
Ganoderic acid H	11

3.3 Analysis.

Control docking results were analyzed from two perspectives, cluster number and free binding energy (G) comparison with experimental G (G_{obs}). Molecular docking results were analyzed by cluster number, and hydrogen bonding comparison with the target's native ligand.

CHAPTER 4

RESULTS

4.1 Reverse Docking

Reverse docking returned ten possible targets from viral infections database for each Ganoderic acid submitted, HIV related targets are shown in Table 4.1

Table 4.1: List of HIV related targets for each Ganoderic acid identified by reverse docking server TarFisDock

Compound Name	PDB ID	Description	Energy score (Kcal mol ⁻¹)	Resolutions (Å ^o)
Ganoderic acid	1RTD	DNA Polymerase/reverse Transcriptase HIV-1 Reverse Transcriptase	-32.54	3.2
	1DIF	HIV Protease	-31.10	1.7
	1HVI	HIV Protease	-28.89	1.8
Ganoderic acid B	1HVR	HIV Protease	-36.55	1.8
	1HVL	HIV Protease	-30.54	1.8
Ganoderic acid C1	1HVR	HIV Protease	-38.40	1.8
	1QS4	HIV-1 Integrase	-30.23	2.1
	1HVL	HIV Protease	-29.95	1.8
Ganoderic acid H	1HVR	HIV Protease	-32.89	1.8

Two types of HIV related proteins for each Ganoderic acid and C1: HIV-1 Reverse Transcriptase, HIV Protease and HIV Protease, HIV-1 Integrase respectively were identified. On the other hand, only HIV Proteases were identified for Ganoderic acid B and H. This is in good agreement with studies carried out by el-Mekkawy and co-workers (el-Mekkawy et al.1998), where they also demonstrated inhibition of HIV Protease by Ganoderic acid , B, C1 and H. Based on these potential targets resolution, their distribution with regard to compounds of interest and correlation to el-Mekkawy and co-

workers (el-Mekkawy et al. 1998) findings, 1DIF and 1HVR (both are HIV-1 Protease) were selected as target for further molecular docking analysis.

4.2 Molecular Docking

4.2.1 Control Docking of 1HVR and 1DIF

Clustering of control docking conformations are shown in Figure 4.1, Summary of conformations clustering is shown in Table 4.2. Control docking of 1HVR returned a total of 14 clusters with cluster 1 being the dominant cluster (cluster with the highest member number). Member of cluster 1 were also observed to have the lowest root mean square deviation (RMSD), 0-2 Å°. On the other hand the first run of control docking of 1DIF returned a total 85 clusters furthermore none of the cluster members are within 0-2 Å° range. Due to the huge number of cluster only the first ten clusters of 1DIF control docking results were reported as shown in Table 4.3.

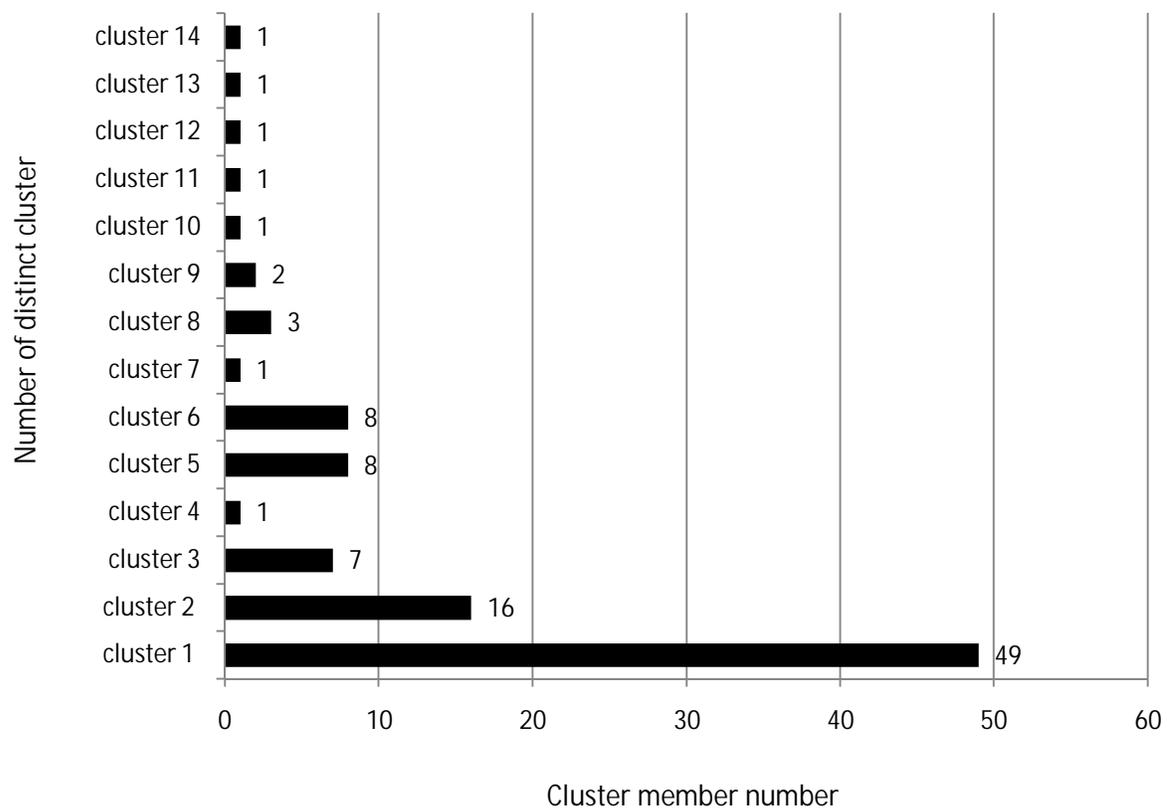


Figure 4.1: (Clustering of control docking conformations of 1HVR)

Table 4.2: Summary of 1HVR control docking conformations

Cluster	No of conformations in cluster	Mean estimated G(Kcal mol ⁻¹)	lowest estimated G(Kcal mol ⁻¹)	RMSD range in A ^o
Cluster 1	49	-13.06	-14.62	0-2
Cluster 2	16	-10.09	-11.21	2-3
Cluster 3	7	-9.20	-10.58	2-3
Cluster 4	1	-9.58	-9.58	3.15
Cluster 5	8	-8.09	-8.50	3-4
Cluster 6	8	-7.91	-8.40	4-6
Cluster 7	1	-7.98	-7.98	2.53
Cluster 8	3	-5.55	-6.93	5-6
Cluster 9	2	-6.56	-6.91	5-5.5
Cluster 10	1	-6.12	-6.12	6.36
Cluster 11	1	-5.99	-5.99	4.68
Cluster 12	1	-5.64	-5.64	5.90
Cluster 13	1	-5.56	-5.56	2.54
Cluster 14	1	-3.26	-3.26	7.04

Table 4.3: Summary of the first ten cluster of 1DIF control docking conformations

Cluster	No of conformations in cluster	Mean estimated G(Kcal mol ⁻¹)	lowest estimated G(Kcal mol ⁻¹)
Cluster 1	6	-9.23	-6.45
Cluster 2	1	-7.12	-7.12
Cluster 3	3	-7.07	-4.78
Cluster 4	1	-6.71	-6.71
Cluster 5	1	-6.49	-6.49
Cluster 6	1	-6.44	-6.44
Cluster 7	2	-6.31	-3.75
Cluster 8	2	-6.78	-4.21
Cluster 9	1	-6.27	-6.27
Cluster 10	3	-6.12	-6.05

4.2.2 Molecular docking of Compounds of Interest to Potential Target

4.2.2.1 Molecular Docking of Compound of interest to 1HVR

Clustering of molecular docking run results for each Ganoderic acids docked to 1HVR is shown in figure 4.2 to 4.5; summary of each docking runs is shown in table 4.5. Molecular docking of Ganoderic acid A and B each produced one dominant cluster, cluster 1(42%) and cluster 2 (78%), respectively. On the other hand molecular docking of Ganoderic acid C1 and H produced three (cluster 1: 23%, cluster 2: 26%, cluster 3: 22%) and two (cluster 1: 24% and cluster 2: 25%) dominant clusters respectively.

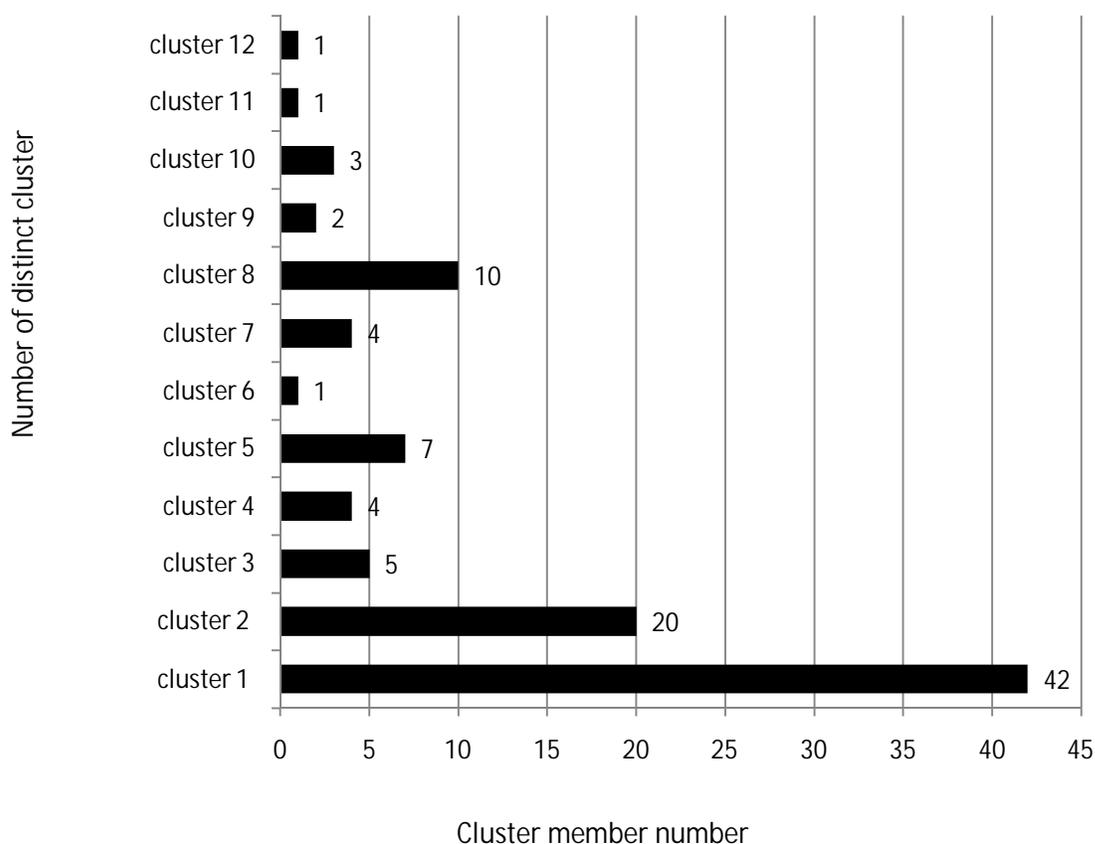


Figure 4.2: (Molecular docking conformations clustering of 1HVR-Ganoderic acid)

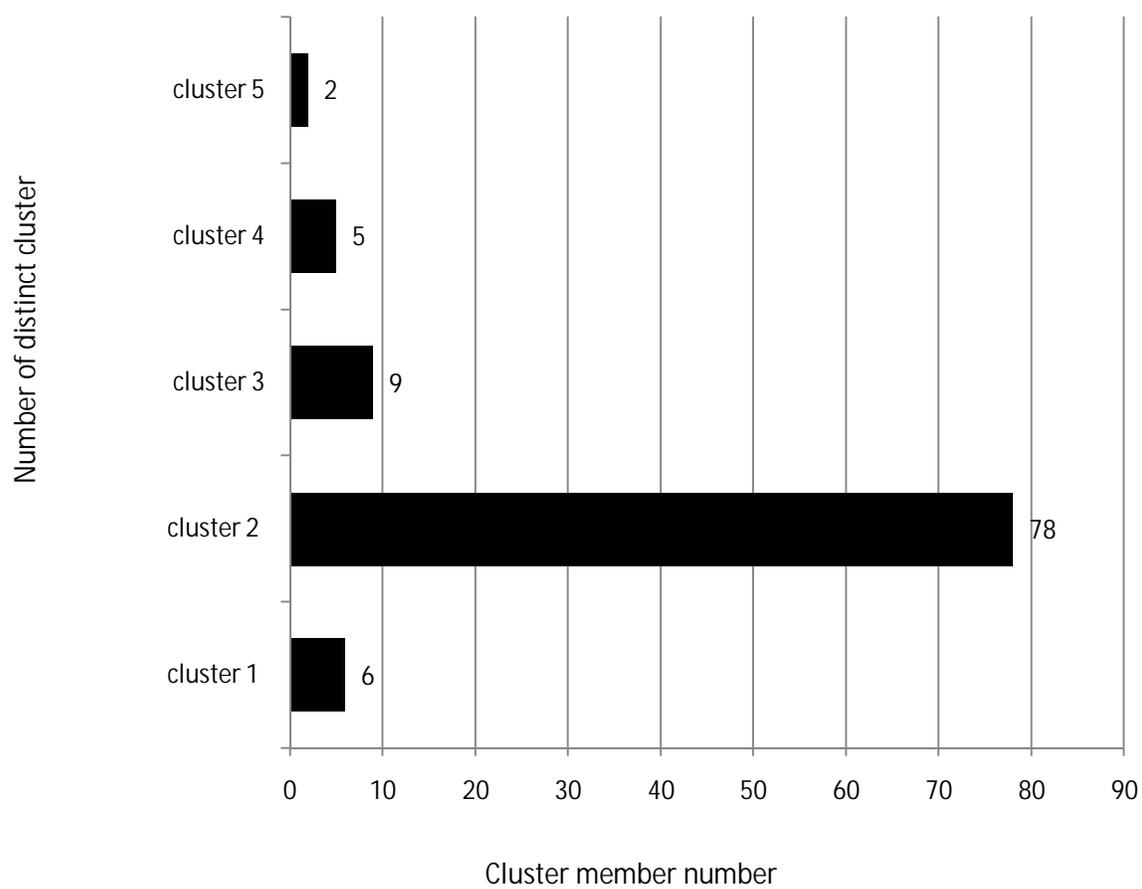


Figure 4.3: (Molecular docking conformations clustering of 1HVR-Ganoderic acid B)

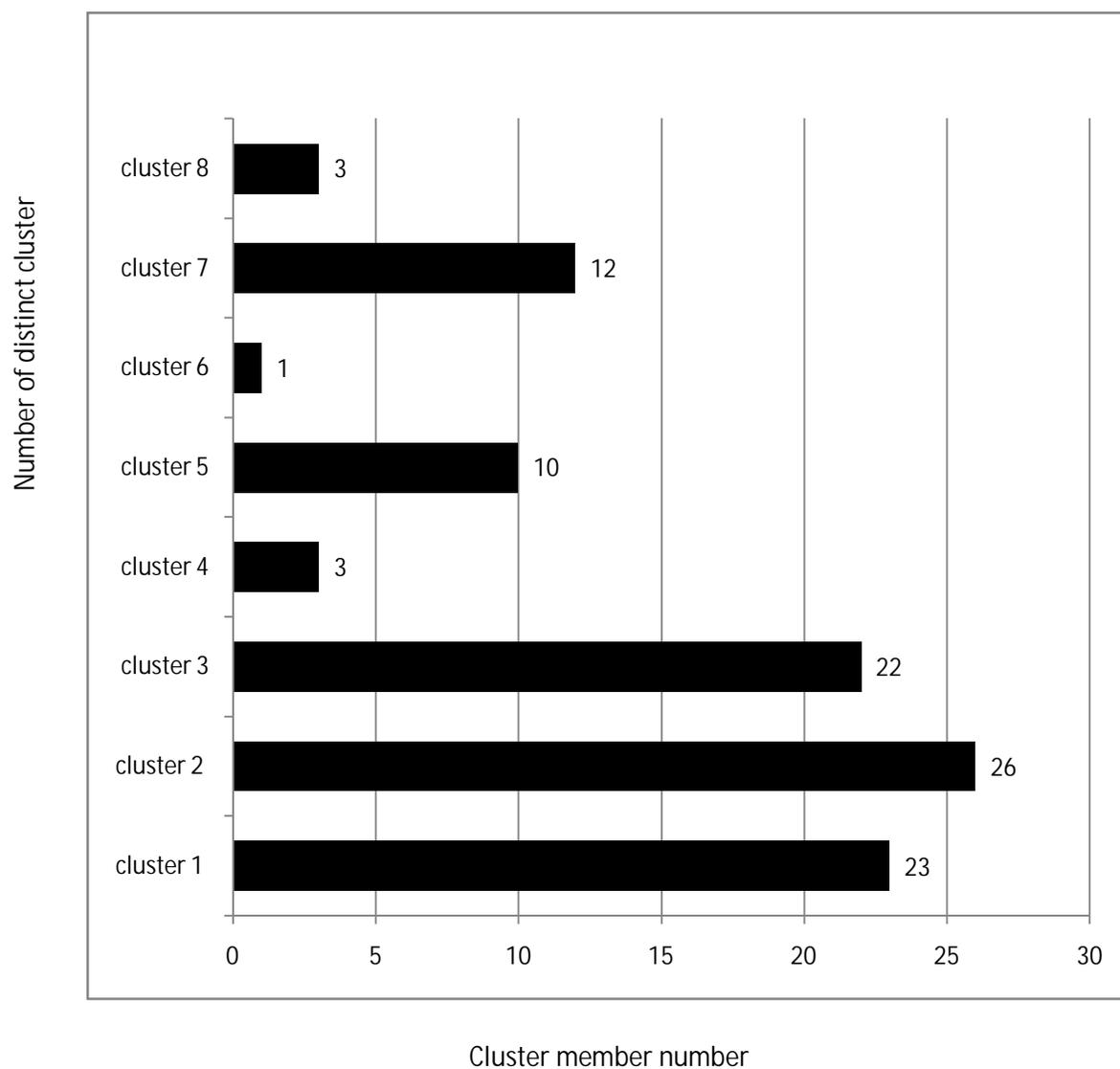


Figure 4.4: 1HVR(Molecular docking conformations clustering of 1HVR-Ganoderic acid

C1)

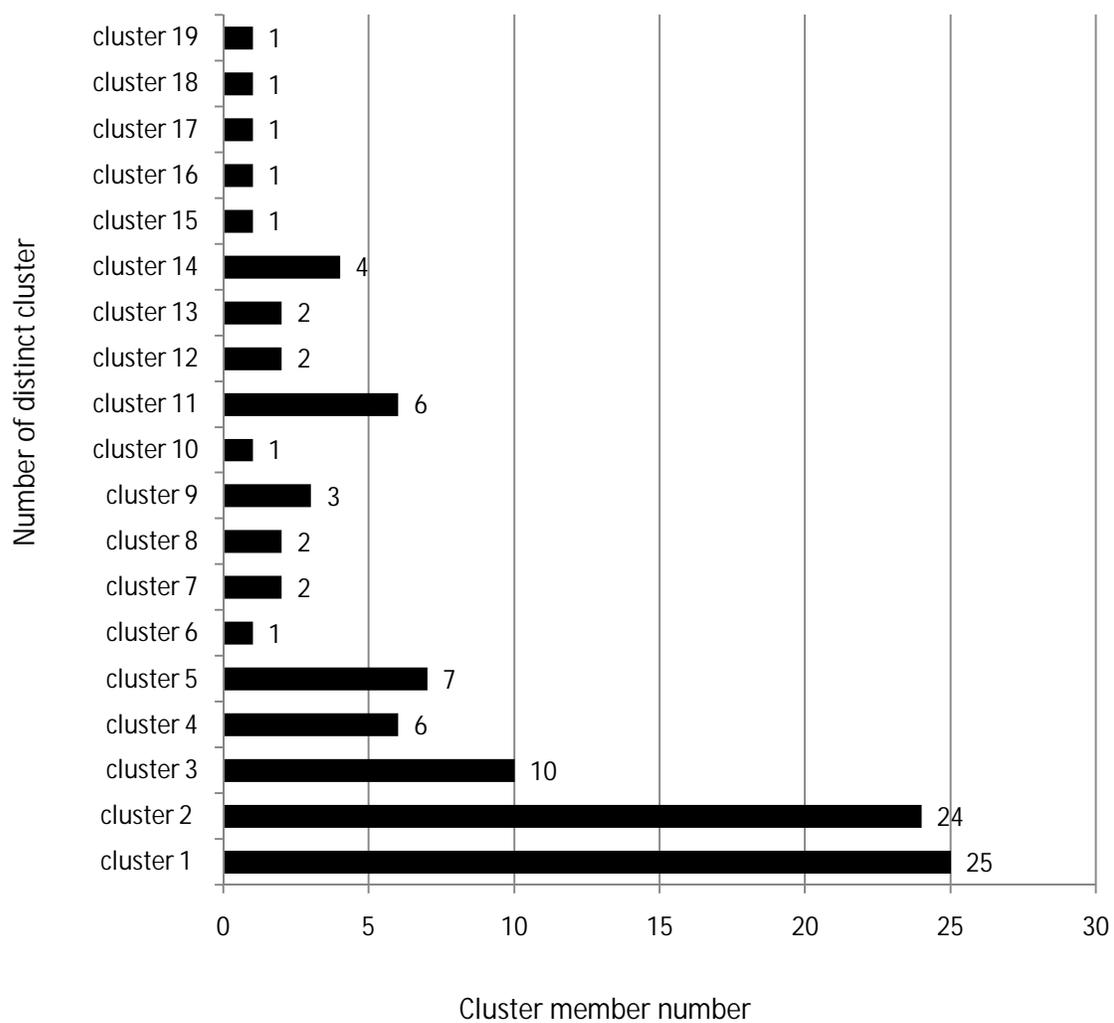


Figure 4.5: (Molecular docking conformations clustering of 1HVR-Ganoderic acid H)

Table 4.5: Summary of dominant cluster result for each Ganoderic acid docked to 1HVR

Compound	Dominant cluster mean estimated G(Kcal mol ⁻¹)	Dominant cluster lowest estimated G(Kcal mol ⁻¹)	Number of cluster	% dominant cluster	1HVR native ligand (XK2) G _{obs} (Kcal mol ⁻¹)
Ganoderic acid	-8.43	-9.71	12	42	
Ganoderic acid B	-8.83	-9.67	5	78	
Ganoderic acid C1	-8.63	-10.10	8	26	-12.97
Ganoderic acid H	-7.02	-8.19	19	25	

4.2.2.2 Molecular Docking of Compound of interest to 1DIF

Clustering of docking run results for each Ganoderic acid docked to 1DIF is shown in Figure 4.6 to 4.9. Similarly to the previous molecular docking, Ganoderic acid A and B produced one dominant cluster 2 (45%) and cluster 2 (40%) respectively. On the other hand Ganoderic acid C1 and H produced three (cluster 1: 20%, cluster 3: 23%, cluster 4: 28%) and two (cluster 1: 32% and cluster 2: 53%) dominant clusters respectively.

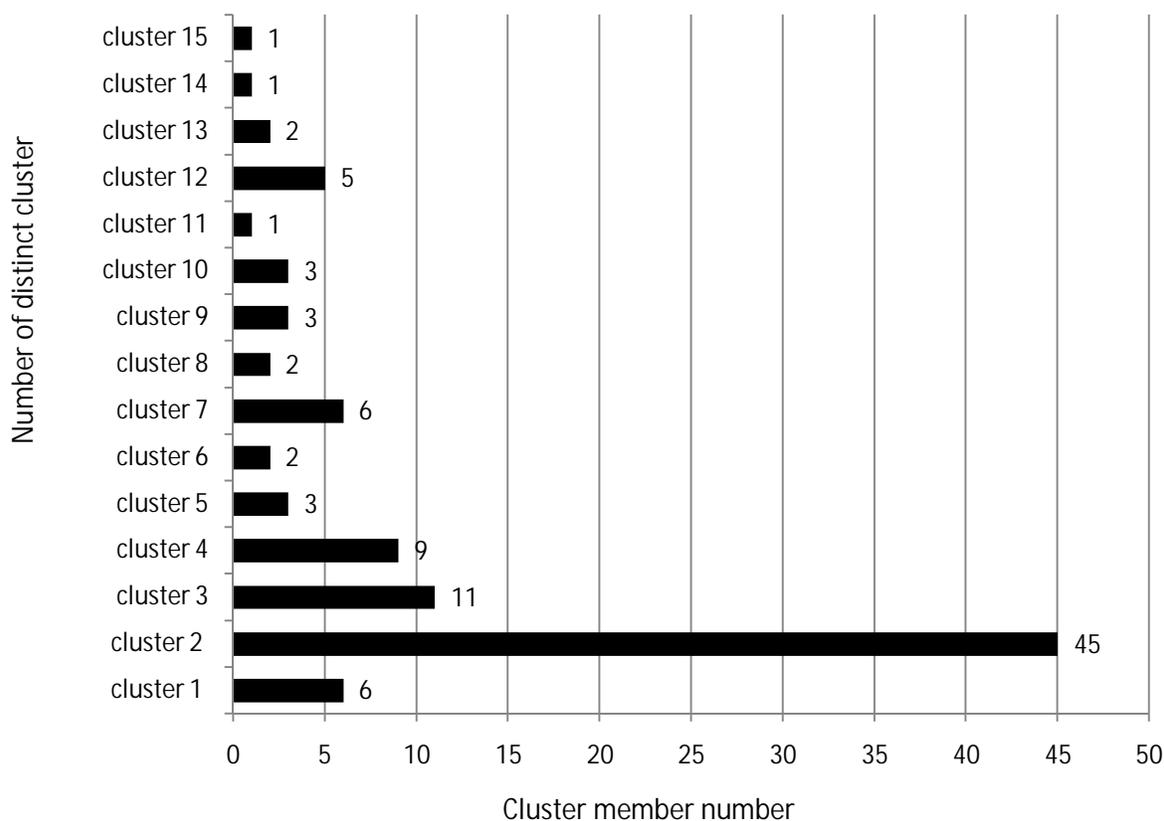


Figure 4.6: (Molecular docking conformations clustering of 1DIF-Ganoderic acid

)

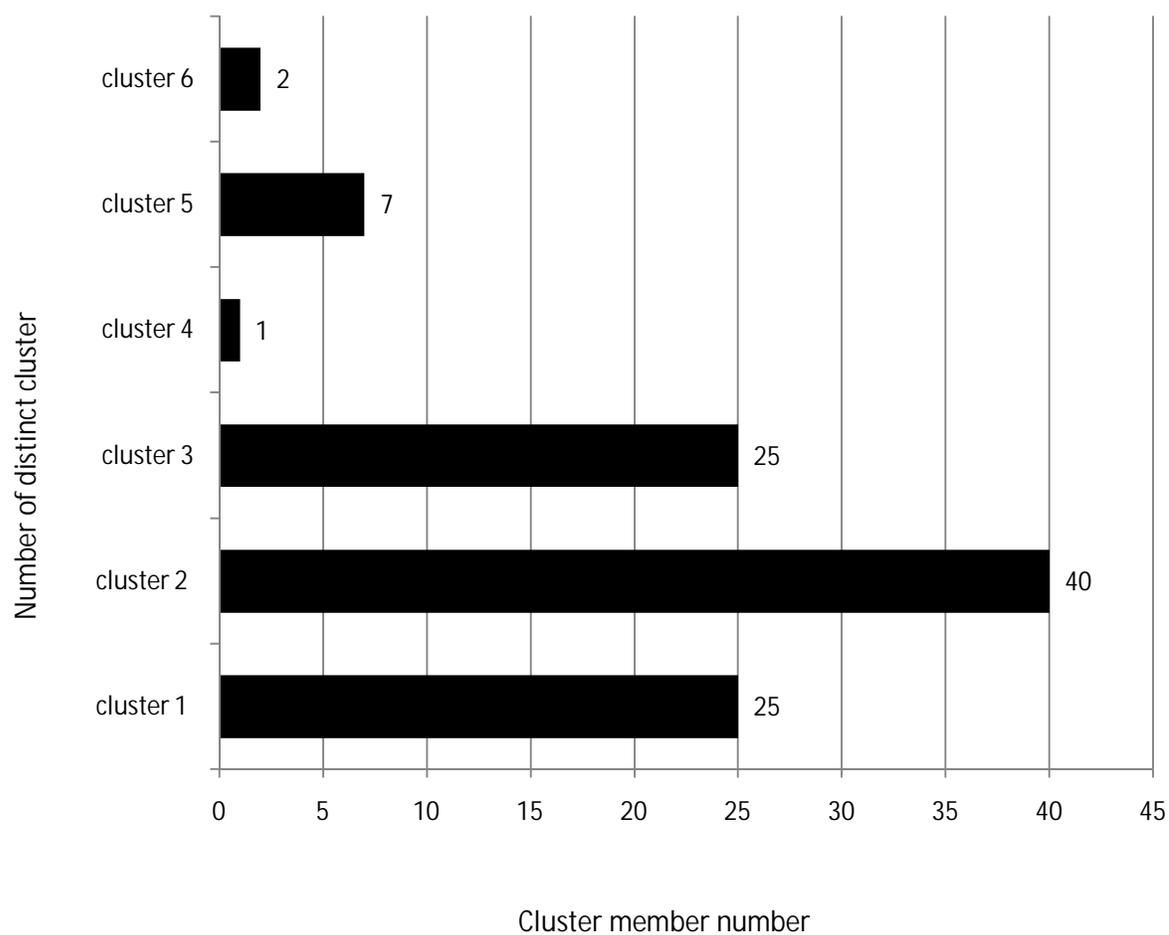


Figure 4.7: (Molecular docking conformations clustering of 1DIF-Ganoderic acid B)

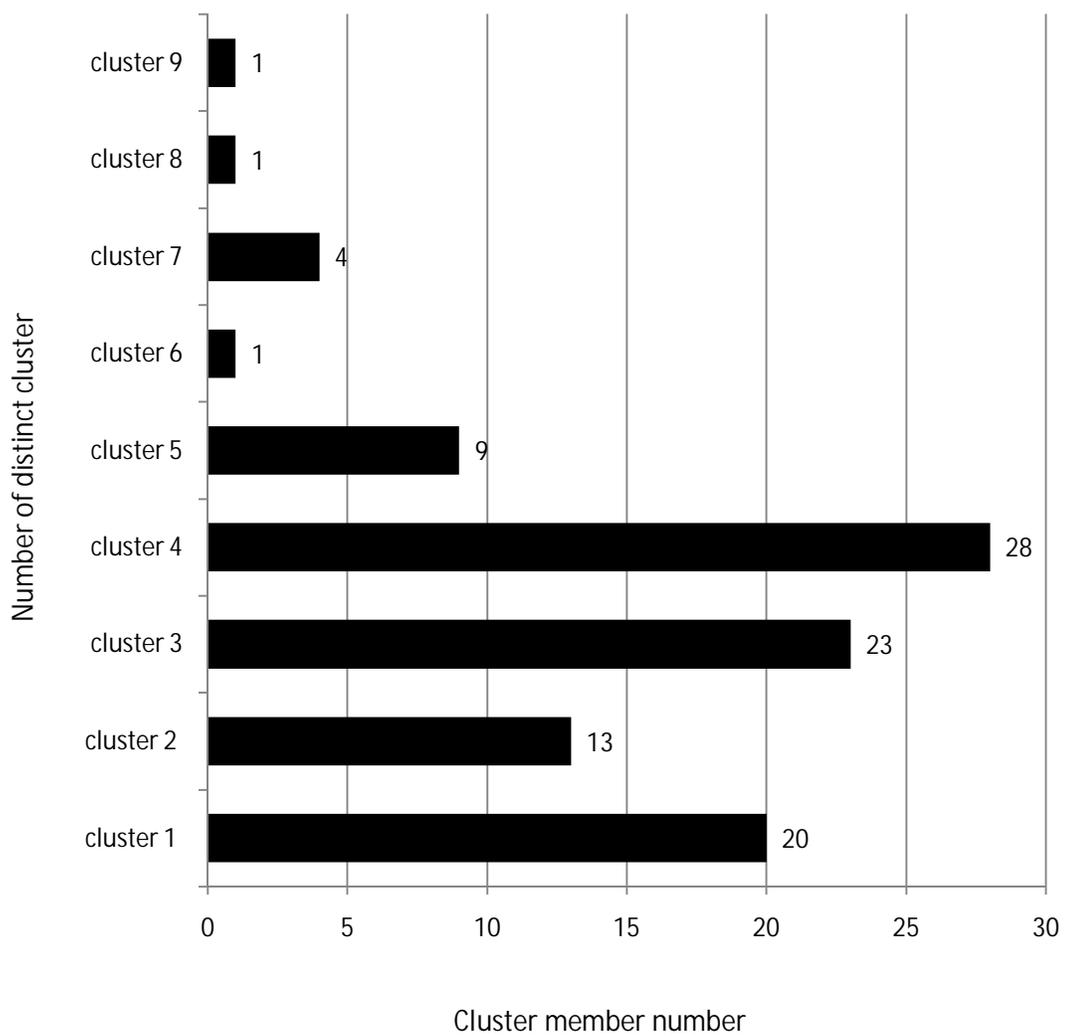


Figure 4.8: (Molecular docking conformations clustering of 1DIF-Ganoderic acid C1)

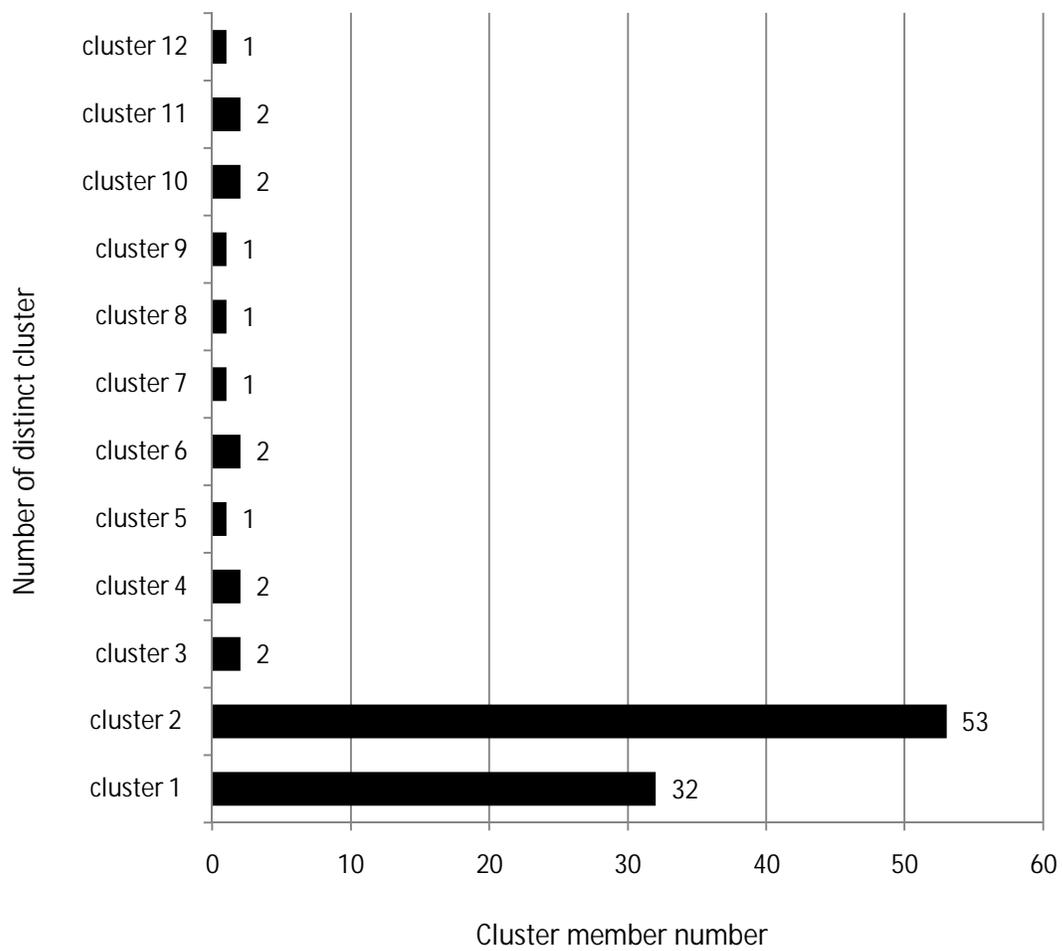


Figure 4.9: (Molecular docking conformations clustering of 1DIF-Ganoderic acid H)

Table 4.6: Summary of dominant cluster result for each Ganoderic acid docked to 1DIF

Compound	Dominant cluster mean estimated G(Kcal mol ⁻¹)	Dominant cluster lowest estimated G(Kcal mol ⁻¹)	Number of cluster	% dominant cluster	1DIF native ligand (A85) G _{obs} (Kcal mol ⁻¹)
Ganoderic acid	-7.53	-8.63	15	45	
Ganoderic acid B	-6.99	-7.97	6	40	-14.59
Ganoderic acid C1	-8.13	-8.53	9	28	
Ganoderic acid H	-6.59	-7.57	12	53	

4.3 Correlating Molecular Docking Result to Experimental Results

Mean ΔG of dominant molecular docking clusters from 1HVR and 1DIF and IC_{50} values from el-Mekkawy and co-workers (el-Mekkawy et al. 1998) are shown in Table 4.7. Better ΔG values in both molecular dockings compared to experimental ΔG values were observed

Table 4.7: Summary of mean ΔG from dominant molecular docking clusters and IC_{50} values of 1HVR and 1DIF

Compound	HIV-PR IC_{50} (mM) (el-Mekkawi et al. 1998)	Converted values of IC_{50} to ΔG (Kcal mol ⁻¹)	dominant cluster mean ΔG 1HVR (Kcal mol ⁻¹)	dominant cluster mean ΔG 1DIF (Kcal mol ⁻¹)
Ganoderic acid	0.19	-3.71	-8.43	-7.53
Ganoderic acid B	0.17	-3.77	-8.83	-6.99
Ganoderic acid C1	0.18	-3.74	-8.63	-8.13
Ganoderic acid H	0.20	-3.68	-7.02	-6.59

4.4 1HVR and Ganoderic acid B Molecular Interaction

1HVR and its native ligand (XK2) interacting residues is shown in Figure 4.12. Four hydrogen bonds were observed in 1HVR-XK2 complex depicted in Figure 4.12 and 4.13. 1HVR-Ganoderic acid B model 34 complex is shown in Figure 4.14. Four hydrogen bonds were also observed in 1HVR-Ganoderic acid B model 34 complex as shown in Figure 4.15. Similarly, four hydrogen bonds were observed in 1HVR-Ganoderic acid model 31 complex.

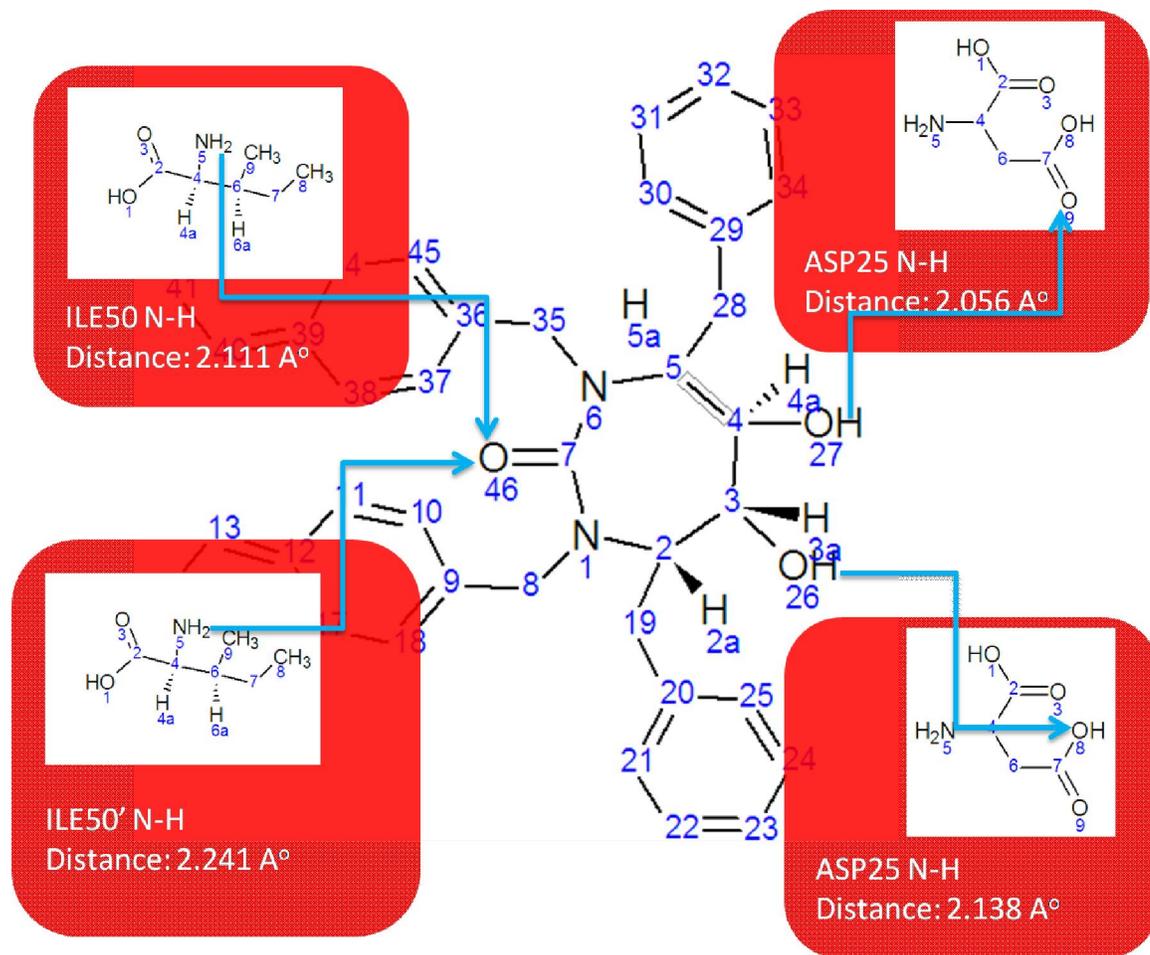


Figure 4.11: (Summary of 1HVR-XXK2 hydrogen bond interactions)

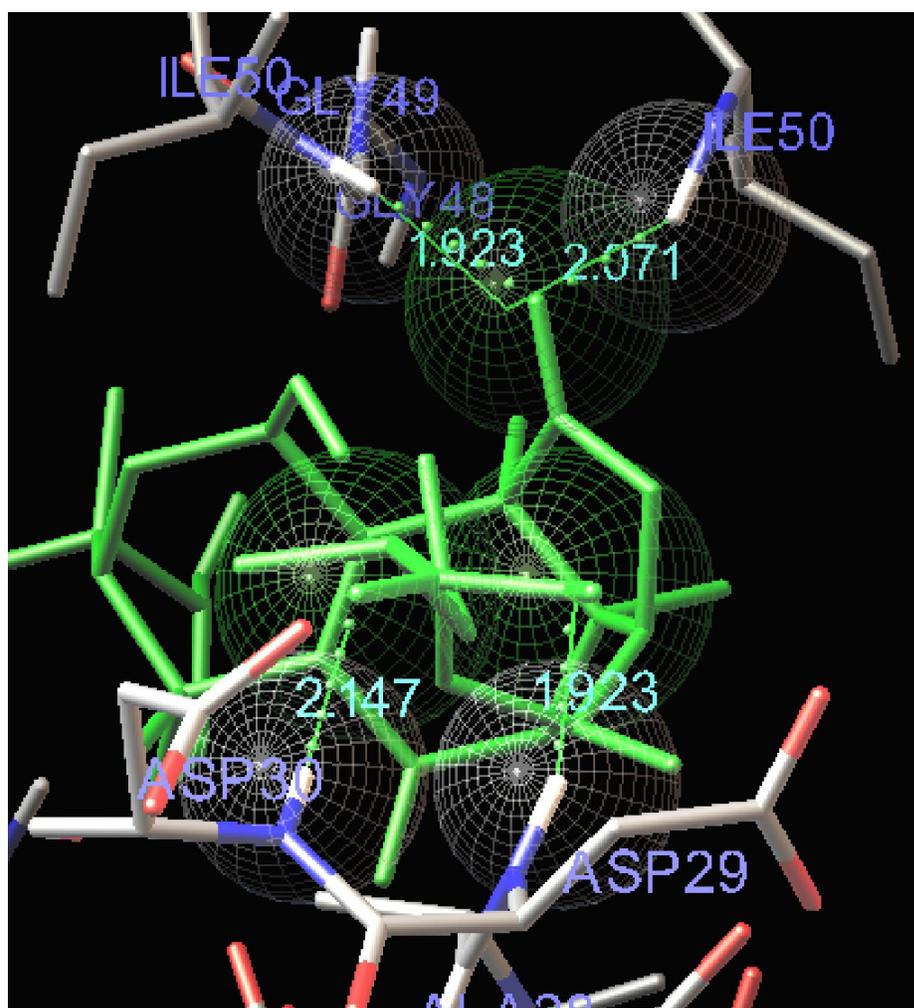


Figure 4.12: (Three dimensional depiction of Ganoderic acid model 34(green), ILE50, ILE50', ASP29 and ASP30 residues of 1HVR (spheres and green lines) interactions.

Numbers in turquoise color are distance between the interacting atoms)

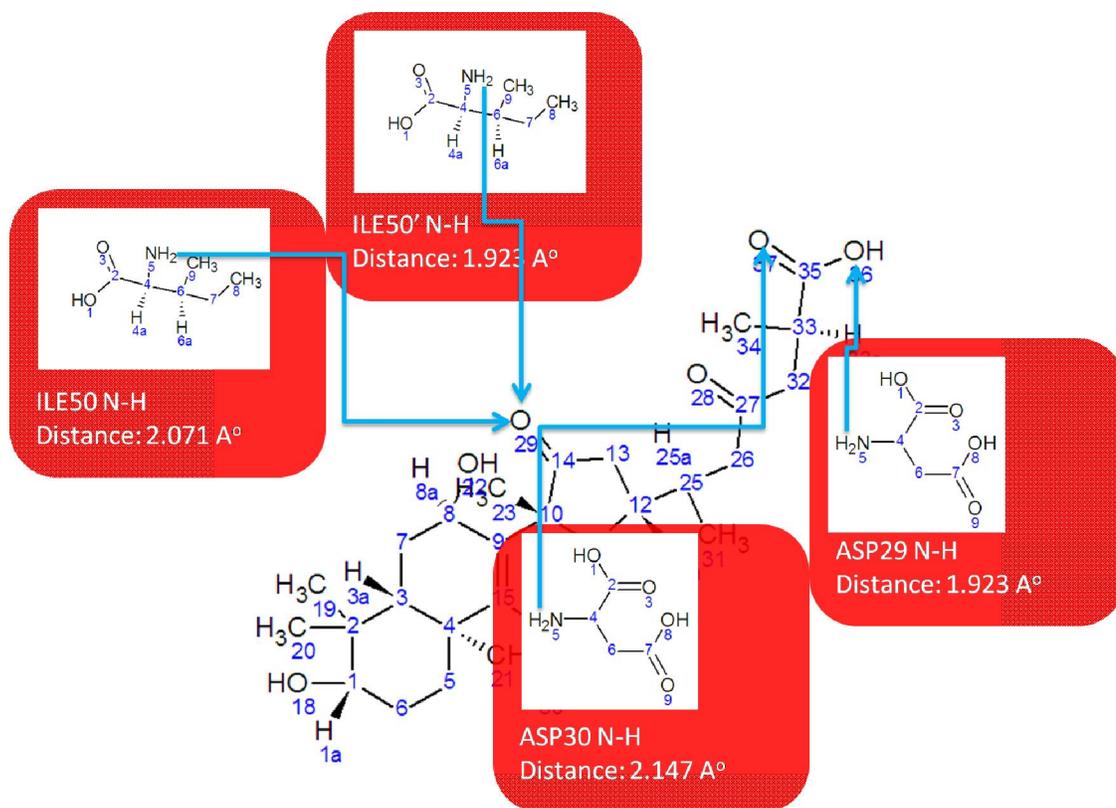


Figure 4.13: (Summary of 1HVR-Ganoderic acid B model 34 hydrogen bond interactions)

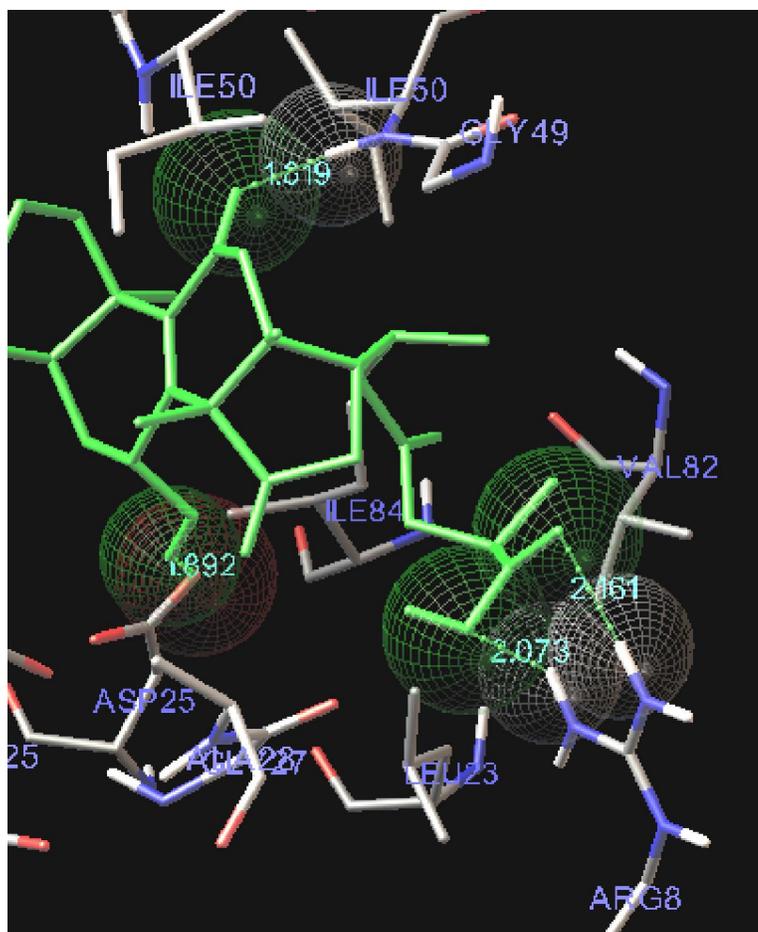


Figure 4.14: (Three dimensional depiction of Ganoderic acid model 31(green), ILE50, ARG8 and ASP25 residues of 1HVR (spheres and green lines) interactions. Numbers in turquoise color are distance between the interacting atoms.)

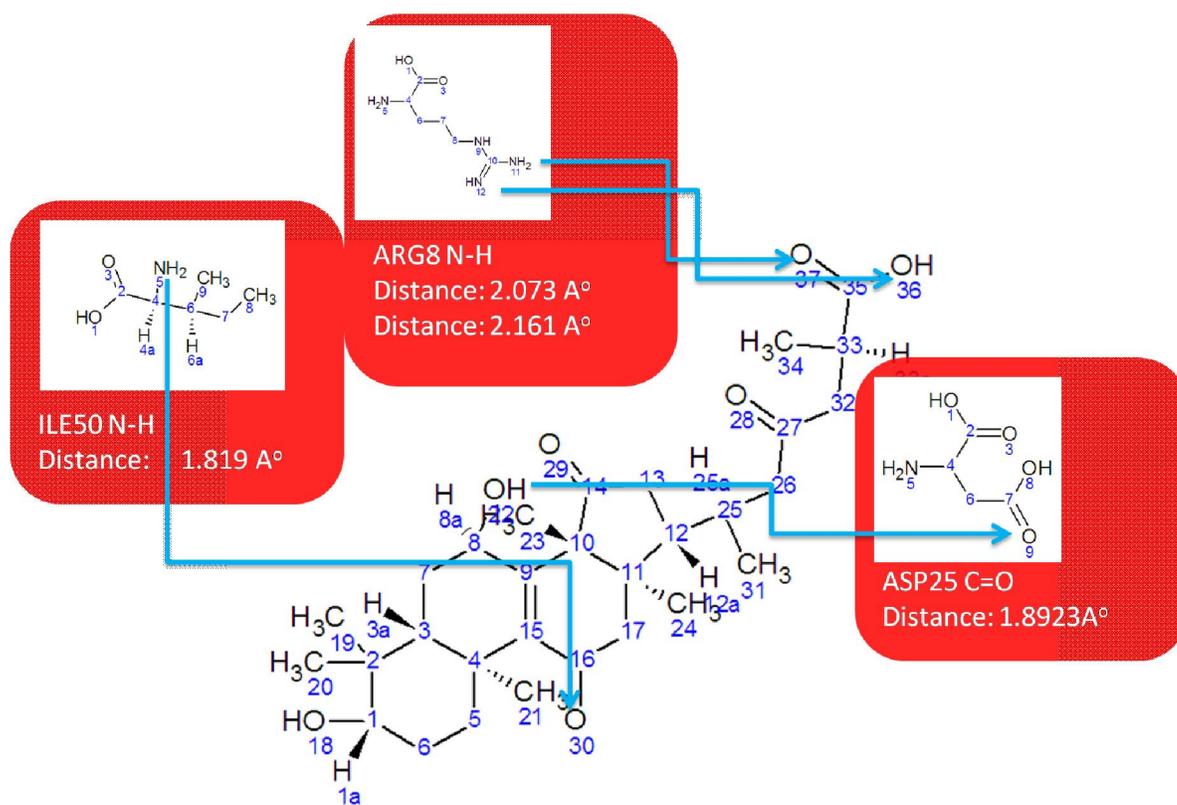


Figure 4.15: (Summary of 1HVR-Ganoderic acid B model 31 hydrogen bond interactions.)

CHAPTER 5

DISCUSSION

5.1 Calibration by Control Docking Studies

It is crucial to verify parameters that are implemented in molecular docking so that *in silico* conditions can be simulated as closely as possible to *in vivo* conditions. This is particularly important during the search space assignment of target's active site. Correct orientation and conformation can only be achieved in the right site. Control docking aims to verify these parameters by using deviation from crystal structure as the benchmark point. Root mean square deviations (RMSD) of the docked conformation were obtained by comparing and calculating differences in atomic coordinates between the docked conformation and the available crystal structure. It is desirable and a common practice in molecular docking to obtain RMSD value below 2 Å in control docking studies.

5.1.1 1HVR Control Docking

Redocking of original ligand (XK2) to 1HVR returned a decent result, conformations in cluster 1 which have close resemblance with the original ligand (RMSD less than 2 Å) dominated the result as shown in Figure 4.1 and Table 4.2. AutoDock was able to produce majority of conformations with low RMSD in the given parameters. AutoDock's ability to produce low RMSD conformations is a good indicator for reliable parameters. Parameters used in the redocking process were able to point AutoDock to the correct search space (grid box is well placed) and it also indicated the number of energy evaluations and Genetic Algorithm runs were sufficient to produce conformations which resembles crystal structure closely. To further verify reliability of the parameters, free binding energy for the original ligand (XK2) were calculated and compared against mean estimated free binding energy

the clusters. Free binding energy of the ligand was termed as observed free binding energy change (G_{obs}) equation (Morris et al., 1998) and calculated using the following equation:

$$\Delta G_{obs} = RT \ln K_i$$

Where R is the gas constant, $1.987 \text{ cal K}^{-1} \text{ mol}^{-1}$, T is absolute temperature in, 298.15 K, an K_i is inhibition constant, 0.31 nM (K_i for XK2 was obtained from PDB, <http://www.pdb.org/pdb/home/home.do>). Comparison of G is shown in Table 5.1:

Table 5.1: G comparison between the calculated 1HVR-XK2 and 1HVR control docking

Cluster	Mean estimated G (Kcal mol ⁻¹)	lowest estimated G (Kcal mol ⁻¹)	1HVR and XK2 G_{obs} (Kcal mol ⁻¹)
Cluster 1	-13.06	-14.62	
Cluster 2	-10.09	-11.21	
Cluster 3	-9.20	-10.58	
Cluster 4	-9.58	-9.58	
Cluster 5	-8.09	-8.50	
Cluster 6	-7.91	-8.40	
Cluster 7	-7.98	-7.98	-12.97
Cluster 8	-5.55	-6.93	
Cluster 9	-6.56	-6.91	
Cluster 10	-6.12	-6.12	
Cluster 11	-5.99	-5.99	
Cluster 12	-5.64	-5.64	
Cluster 13	-5.56	-5.56	
Cluster 14	-3.26	-3.26	

Again, strong resemblance between crystal structure G_{obs} and mean estimated G was observed. Cluster 1 G was found to be very close to the calculated G_{obs} , -13.06 Kcal mol⁻¹ and -12.97 Kcal mol⁻¹, respectively. Low cluster number, low RMSD value and similar mean G to the G_{obs} indicated the parameters used in 1HVR redocking were suitable to reproduce original ligand conformation and 1HVR redocking objectives were achieved.

5.1.21DIFControl Docking

As shown in table 5.2, low cluster number was not obtained in control docking of 1DIF in run 1 and 2 indicated by the huge number of clusters on each runs. Even though population size and energy evaluation numbers were increased significantly compared to 1HVR control docking, Autodock were unable to produce conformations that structurally resembles 1DIF original ligand (A85). None of the conformations in both runs were within the 0-2 Å range and a high number clusters were observed. Comparison of cluster's mean G with G_{obs} also showed significant differences where the mean cluster's G were underestimated compared to the G_{obs} (Table 5.2). Slight improvement were observed in run 2 indicated by decrease in cluster number and both lowest mean G and lowest G .

Additional docking runs were not able to improve 1DIF control docking results was due to the high torsion number of A85. Complexity of conformation search increased exponentially along with the increase of torsion numbers simply due to the increase in the number of Genetic Algorithm variables/genes. As shown in Table 5.2, torsion number of 1DIF ligand was more than twice as much of that of 1HVR's. Computing time was also increased significantly as search complexity increased. AutoDock took only five hours to complete 1HVR control docking while nearly seventy hours of computing time were not able to produce low cluster number in 1DIF control docking.

With the current limited computing facility capacity, optimizing parameters to more than 100×10^6 energy evaluations is beyond practical, let alone implementing those parameters during compounds of interest docking. Fortunately, all four compounds of interest number of torsions were almost similar to 1HVR as shown in Table 3.2. Based on the current control docking results, it is logically expected that implementation of docking parameters similar to 1HVR control docking or further improvement of parameters based

on 1HVR control docking will be able to produce low cluster number for molecular docking compounds of interest to both 1HVR and 1DIF.

Parameters from 1HVR were implemented for compounds of interest docking to 1HVR and parameters from 1DIF run 1 control docking were implemented for compounds of interest docking to 1DIF.

Table 5.2: Summary and comparison of 1DIF control docking runs and 1HVR cluster 1.

Control docking run	Ligand number of torsion	GA population size	Number of GA runs	Number of energy evaluations	Lowest mean (G)	lowest estimated (G)	Number of cluster	RMSD value less than 2 A°	Computing time	G _{obs}
1DIF run 1	21	300	100	10x10 ⁶	-8.50	-9.58	85	none	29h 45m 44.61s	
1DIF run 2	21	300	100	25 x10 ⁶	-9.01	-10.30	80	none	69h 58m 23.14s	-14.59
1DIF run 3	21	300	100	50 x10 ⁶					Running	
1DIF run 4	21	300	100	100 x10 ⁶					Running	
1HVR cluster 1	10	150	100	2.5 x10 ⁶	-13.06	-14.62	14	49	5h 09m 27.06s	-12.97

Computing time was calculated in system running Windows®7, Intel®Core™Duo CPU E7600 @ 3.06 GHz and 3.07 GHz, 3GB RAM

5.2 Compounds of Interest Molecular docking

5.2.1 Molecular Docking of Compound of interest to 1HVR

As summarized in Table 4.5, the given parameters were able to produce good docking for Ganoderic acid A and B as indicated by the large percentage of dominant cluster, 42% and 78%, respectively. On the other hand, Ganoderic acid C1 and H dominant cluster has smaller percentage, 26% and 25 %, respectively. Ganoderic acid B in particular took the centre stage in this docking result because the compound scored the highest percentage of dominant cluster and lowest number of cluster. Molecular docking was able to reproduce similar conformations iteratively.

It is also interesting and worth mentioning that Ganoderic acid B has more torsion angles compared to Ganoderic acid C1 yet low cluster number was observed. As discussed in the previous section of this report, torsion angles were able to increase conformation search complexity exponentially. The fact that Ganoderic acid B was able to achieve convergence without owning the lowest torsions is another indication that naturally Ganoderic acid B interacts with the target better compared to other compounds tested in this docking experiment.

5.2.2 Molecular Docking of Compound of interest to 1DIF

Similarly to 1HVR docking, the given parameters were able to produce low cluster number for Ganoderic acid A, C, and H in 1DIF docking. Ganoderic acid H has the most percentage of dominant cluster followed by Ganoderic acid C and B, 53%, 45% and 40%, respectively. On the other hand Ganoderic acid C1 was not able to produce low cluster number,

the results are summarized in Table 4.6. Interestingly Ganoderic acid B has the lowest cluster number compared to other compounds. Even though Ganoderic acid B was not able to produce the lowest ΔG , the low cluster number indicates Ganoderic acid B performed better compared to the rest of the compounds. An interesting phenomenon which also emerged strongly in 1HVR docking result.

5.3 Correlating Molecular Docking Result to Wet Lab Results

It is important to understand that molecular docking is a simulation and there is always possibility for false positive. Correlating molecular docking result to wet lab results will serve as a good indicator to molecular docking results credibility. For this reason, mean ΔG from dominant cluster of molecular docking and ΔG values of compounds on interest obtained from el-Mekkawy and co-workers (el-Mekkawy et al. 1998) were compared. Significantly better ΔG values from both 1HVR and 1DIF molecular dockings were observed as shown in Table 4.7. Similar trend to experimental results where Ganoderic acid B scored the lowest ΔG were retained in molecular docking of 1HVR, the same trend was not observed in 1DIF molecular docking. These observation indicated better interaction were formed between 1HVR and Ganoderic acid B. Based on this observation, 1HVR and Ganoderic acid B molecular interaction were further studied in the next section.

5.4 Molecular Interactions of 1HVR and Ganoderic acid B

Docking of Ganoderic acid B and 1HVR returned better results both in cluster and correlation analysis compared to other docking as discussed in previous sections. To gain deeper information about these two molecules interactions, interactions were reviewed in the best two conformations. The first conformation was model 31 and the second conformation was model 34. Model 31 was the lowest scoring ΔG of all 1HVR-Ganoderic

acid B docking runs. On the other hand model 34 was the lowest scoring Gin cluster 2 the dominant cluster of 1HVR-Ganoderic acid B docking.

Four hydrogen bonds were observed in 1HVR-Model 34 complex as shown in Figure 4.12 and 4.13. ILE50 and ILE50' amide hydrogen formed two hydrogen bonds with Ganoderic acid B model 34 carbonyl of carbon atom position 14. Other hydrogen bonds were observed between amide hydrogen of ASP30 and Ganoderic acid B model 34 carbonyl of carbon atom position 35. The last hydrogen bond was between amide hydrogen of ASP29 and Ganoderic acid B model 34 hydroxyl of carbon atom position 36. Four hydrogen bonds were also observed 1HVR-Model 31 complex as shown in figure 4.14 and 4.15. ILE50 amide hydrogen formed hydrogen bond with Ganoderic acid B model 31 carbonyl of carbon atom position 16. ARG8 amide hydrogens formed two hydrogen bond with Ganoderic acid B model 31 carbonyl and hydroxyl of carbon atom position 35. The last hydrogen bond was between ASP25 carbonyl and Ganoderic acid B model 31 hydroxyl of carbon atom position 8.

Interestingly similar hydrogen bonding regions were also observed in the 1HVR-XK2 complex. Two hydrogen bonds from amide hydrogen of ILE50 and ILE50' interaction with XK2 carbonyl of carbon atom position 7 were observed. Another two hydrogen bonds were from XK2 hydroxyl of carbon atom position 3 and 4 interaction with ASP25 carbonyl and hydroxyl of carbon atom position 7 as shown in Figure 4.10 and 4.11. Interaction of ILE50 and ILE50' amide hydrogen, structural water molecule and carbonyl of ligand or inhibitor through hydrogen bond is a common and a very important feature found in HIV-1 protease (Lebon and Ledecq 2000). Further highlighting the importance of this ILE50-ligand interactions, Lam and coworkers based the entire rational design of their protease inhibitor study on this feature (Lam, Jadhav et al. 1994). The lack of structural water

molecule in Figure 4.10, 4.12 and 4.14 is due to molecular docking were run on simplified environment by removing all water molecules from the macromolecule, a common practice in molecular docking routine. Ganoderic acid B interactions with one of the most important feature of HIV-1 protease were observed in this report. This finding verifies the significance of 1HVR and HIV-1 protease in general is a suitable Ganoderic acid B target.

CHAPTER 6

CONCLUSION

Ganoderic acid B performed better in molecular docking compared to other compounds of interest, not only it outperformed other compounds in terms of cluster number, it also showed similar ΔG trend with experimental data obtained from el-Mekkawy and co-workers (el-Mekkawy et al. 1998). Molecular interactions study revealed Ganoderic acid B interactions with important residues of 1HVR, thus making 1HVR and HIV-1 protease in general suitable targets for this compound. The fact that Ganoderic acid B is a naturally occurring compound and was found to interact with one of the most important features of HIV-1 protease indicated a huge potential for HIV cure discovery based on this compound.