

CHAPTER 5

COMPUTATIONAL DOCKING OF NON-COMPETITIVE INHIBITORS TO

DEN-2 NS2B-NS3

5.1 Introduction

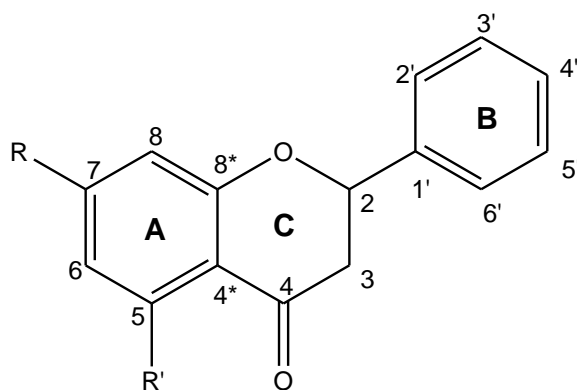
In spite of the efforts of many research groups (Diamond *et al.*, 2002; Hrobowski *et al.*, 2005; Putnak *et al.*, 2005; Tan *et al.*, 2006; Whitby *et al.*, 2005; Yin *et al.*, 2006), there is still no vaccine or antiviral drug currently available for dengue viral infections (Ray & Shi, 2006). Hence, there is an immense interest in search for antiviral therapeutic agents to fight the diseases caused by dengue viruses. Natural products have been the source leads for many antiviral therapeutics (Kitazato *et al.*, 2007). This approach continues to furnish investigators with new and interesting findings due to the molecular diversity of natural products. Between 1983 to 1994, seven out of ten synthetic agents approved by the Food and Drug Administration (FDA) for use as antivirals were based on lead compounds which were natural products (Spainhour, 2005).

Boesenbergia rotunda (L.) Mansf. Kulturpfl. belongs to the ginger family (Zingiberaceae) and it is a spice commonly used in Southeast Asia, especially in Malaysia, Thailand and Indonesia. Some reports on its medicinal properties have been documented and it is used traditionally, mainly for problems and diseases of women such as in post-partum protective medication, treatment for rheumatism, and as a tonic/lotion (Ibrahim & Rahman, 1989). Scientific reports on the anti-inflammatory and anti-HIV activities of extracts from this plant have also been reported (Tewtrakul *et al.*, 2003a; Tewtrakul *et al.*, 2003b; Tuchinda *et al.*, 2002). In addition, several compounds from *Boesenbergia rotunda* (L.) have also been shown to have non-competitive inhibitory activity against the ability of DEN-2 protease to cleave fluorogenic peptide substrates (Tan, 2005; Tan *et al.*, 2006).

In this study, we report the automated docking studies performed on the compounds isolated from *Boesenbergia rotunda* (L.) which exhibited non-competitive inhibitory activities towards DEN-2 NS2B-NS3. The subjects of this study are the flavanones: pinostrobin, pinocembrin and alpinetin, and their chalcone derivatives: pinostrobin chalcone, pinocembrin chalcone and cardamonin, respectively (Figure 5.1). This work aims at studying the interactions involved in the binding of these compounds (referred to as ligands hereafter) to DEN-2 NS2B-NS3 via computational docking methods and to gain insights into the experimental inhibition pattern (Tan, 2005). It is hoped that information obtained from this study will provide a better understanding of the mechanism of inhibition of DEN-2 protease, and aid in the design of antiviral drugs which could inhibit dengue virus replication. The work described in this chapter has been published (Othman *et al.*, 2008).

5.2 Materials and Methods

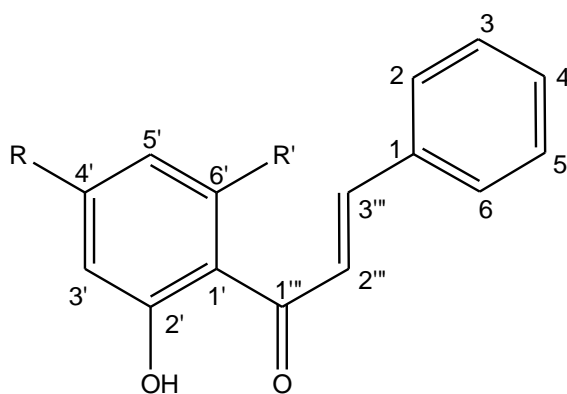
The protocols for this study are as illustrated in the flowchart in Figure 5.2. The purification techniques used to isolate the ligands involved in this study and the methods used in their biological screening for inhibition of proteolytic activity of DEN-2 NS2B-NS3 have been reported (Tan, 2005; Tan *et al.*, 2006). The fluorogenic substrate used in these studies was Boc-Gly-Arg-Arg-MCA, and DEN-2 NS3 protease has been shown to cleave this substrate.



R = OMe, R' = OH : Pinostrobin (1)

R = OH, R' = OH : Pinocembrin (2)

R = OH, R' = OMe : Alpinetin (3)



R = OMe, R' = OH : Pinostrobin chalcone (4)

R = OH, R' = OH : Pinocembrin chalcone (5)

R = OH, R' = OMe : Cardamonin (6)

Figure 5.1 Structures of the compounds extracted from *Boesenbergia rotunda* L.

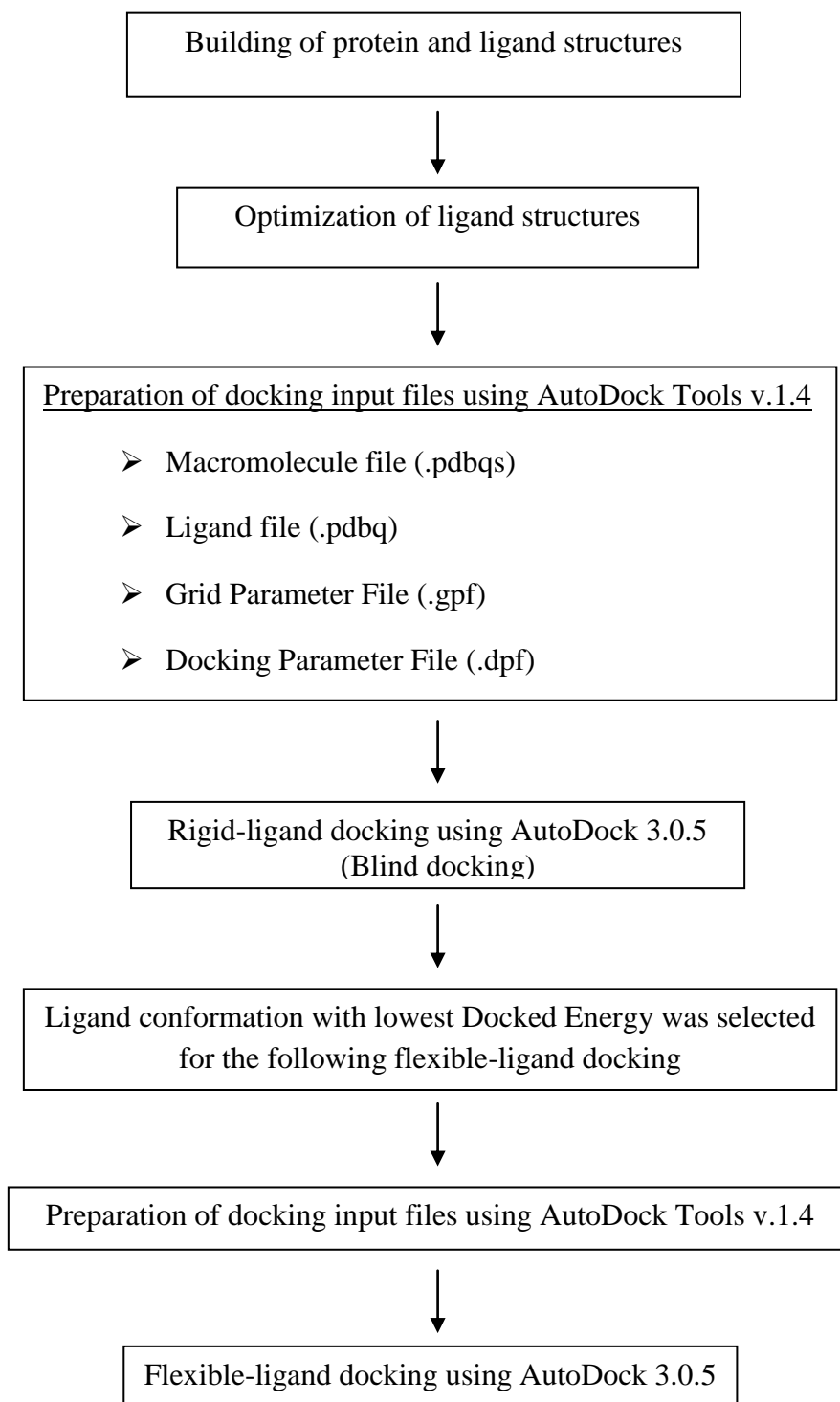


Figure 5.2 Flowchart of protocols involved in the study of docking of non-competitive inhibitors to DEN-2 NS2B-NS3.

5.2.1 Materials

5.2.1.1 Hardwares

Docking simulations in this study were performed on an IBM i386 computer, running on SUSE Linux 9.3 operating system, Intel® Pentium® 4 processor 2.40 GHz and 256 RAM. All other computational calculations and visualization were done on a notebook running on Microsoft® Windows® XP operating system, AMD Turion™ 64 Mobile Technology with speed of 795 MHz and 480 MB of RAM.

5.2.1.2 Softwares

The Hyperchem Pro 6.0 software (Hypercube Inc.) was used to build and optimize ligand structures prior to docking simulations. This software was also used to generate the 3-dimensional (3D) isosurface plots of the electrostatic potentials of the ligands. Docking files were prepared using the AutoDock Tools v.1.4 software (<http://www.scripps.edu/~sanner/python/adt>) (Coon *et al.*, 2001), an open-source software from MGL tools software package (<http://mgltools.scripps.edu/>). AutoDock 3.0.5 software package (Morris *et al.*, 1998) was used for docking calculations. Viewerlite 4.2 (<http://www.accelrys.com>) visualization tool was used to display and analyse structures and interactions, and LIGPLOT program (Wallace *et al.*, 1995) was used to map out hydrophobic interactions.

5.2.2 Methods

5.2.2.1 Building and optimization of protein and ligand structures

The three-dimensional structure of DEN-2 NS2B-NS3 was retrieved from the

Protein Data Bank (Berman *et al.*, 2000) with accession code 2FOM. Chlorine atoms, water and glycerol molecules were removed. Crystal coordinates of the flavanones were obtained from the Cambridge Crystallographic Data Centre and related publications (Brown *et al.*, 2006; Jiang *et al.*, 2001; Shoja, 1989), and the structures were built using the Crystal Builder module in the Hyperchem Pro 6.0 software (Hypercube Inc.). Structures of the chalcones were built through modification of the flavanones using the Hyperchem Pro 6.0 software. The ligand structures were then optimized using Hyperchem Pro 6.0 with MM+ parameters (default setting) using the steepest descent and conjugate gradient (Polak-Ribiere) algorithms (energy minimization termination conditions, ie. the convergence criteria, were set to maximum of 500 cycles or 0.1 kcal / Å mol RMS gradient).

5.2.2.2 Automated rigid-ligand docking

Docking input files were prepared using AutoDock Tools v.1.4 software (<http://www.scripps.edu/~sanner/python/adt>) (Coon *et al.*, 2001). For the protein molecule, polar hydrogen atoms were added and non-polar hydrogen atoms were merged, Kollman charges and solvation parameters were assigned by default. The macromolecule file was saved with .pdbqs file extension. For the ligands, Gasteiger charges were added, non-polar hydrogen atoms were merged and all bonds were made non-rotatable. The ligand files were saved with the .pdbq file extension.

Grid parameter files (GPF) were prepared as input files for the AutoGrid utility to calculate grid (or potential) maps. The maps generated depend on the types of ligand atoms specified (Figure 5.3). Essentially, in the GPF (Table 5.1), the following parameters are specified:

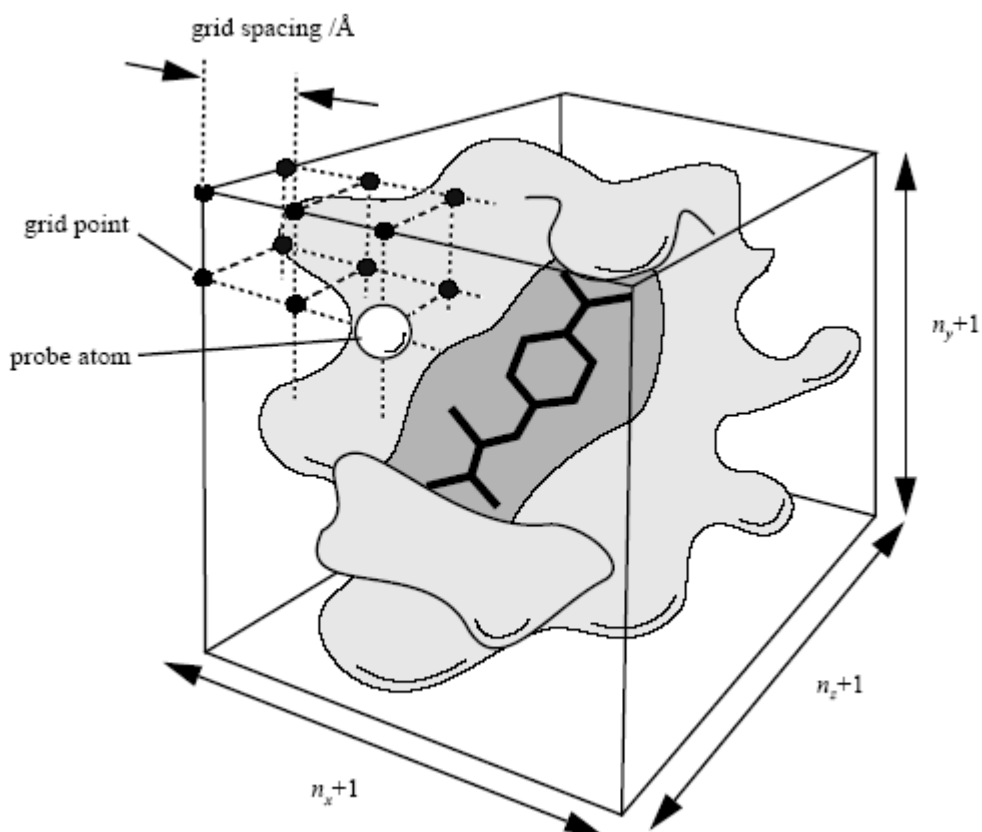


Figure 5.3 The main features of a grid map. The ligand can be seen in the centre of the grid map, buried inside the active site of the protein. A grid map consists of a three dimensional lattice of regularly spaced points, surrounding (either entirely or partly) and centered on some region of interest of the macromolecule under study. Typical grid point spacing varies from 0.2 Å to 1.0 Å, although the default is 0.375 Å. Each point within the grid map stores the potential energy of a ‘probe’ atom or functional group that is due to all the atoms in the macromolecule. (Figure adapted from AutoDock version 3.0.5 User’s Guide).

Table 5.1 An example of the grid parameter file (GPF) containing parameters required for calculating grid maps.

```

receptor 2FOMCa3.pdbqs # macromolecule
gridfld 2FOMCa3.maps.fld # grid_data_file
npts 128 100 116 # num.grid points in xyz
spacing 0.375 # spacing(A)
gridcenter -2.572 -16.384 16.826 # xyz-coordinates or auto
types CAOH # atom type names
smooth 0.5 # store minimum energy w/in rad(A)
map 2FOMCa3.C.map # atom-specific affinity map
nbp_r_eps 4.00 0.0222750 12 6 # C-C lj
nbp_r_eps 3.75 0.0230026 12 6 # C-N lj
nbp_r_eps 3.60 0.0257202 12 6 # C-O lj
nbp_r_eps 4.00 0.0257202 12 6 # C-S lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
sol_par 12.77 0.6844 # C atomic fragmental volume, solvation
parameters
constant 0.000 # C grid map constant energy
map 2FOMCa3.A.map # atom-specific affinity map
nbp_r_eps 4.00 0.0222750 12 6 # A-C lj
nbp_r_eps 3.75 0.0230026 12 6 # A-N lj
nbp_r_eps 3.60 0.0257202 12 6 # A-O lj
nbp_r_eps 4.00 0.0257202 12 6 # A-S lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
sol_par 10.80 0.1027 # A atomic fragmental volume, solvation
parameters
constant 0.000 # A grid map constant energy
map 2FOMCa3.O.map # atom-specific affinity map
nbp_r_eps 3.60 0.0257202 12 6 # O-C lj
nbp_r_eps 3.35 0.0265667 12 6 # O-N lj
nbp_r_eps 3.20 0.0297000 12 6 # O-O lj
nbp_r_eps 3.60 0.0297000 12 6 # O-S lj
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
sol_par 0.00 0.0000 # O atomic fragmental volume, solvation
parameters
constant 0.236 # O grid map constant energy
map 2FOMCa3.H.map # atom-specific affinity map
nbp_r_eps 3.00 0.0081378 12 6 # H-C lj
nbp_r_eps 1.90 0.3280000 12 10 # H-N hb
nbp_r_eps 1.90 0.3280000 12 10 # H-O hb
nbp_r_eps 2.50 0.0656000 12 10 # H-S hb
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
sol_par 0.00 0.0000 # H atomic fragmental volume, solvation
parameters
constant 0.118 # H grid map constant energy
elecmap 2FOMCa3.e.map # electrostatic potential map
dielectric -0.1146 # <0, distance-dep.diel;>0, constant
#

```

1. the center of the grid map (in cartesian coordinate),
2. the number of grid points in each of the x -, y - and z - directions,
3. the grid spacing (in Å),
4. the grid map filename,
5. the non-bonded parameters for each pairwise-atomic interaction,
6. the electrostatic potential energy grid map,
7. the dielectric function flag.

AutoGrid calculates the non-covalent energy of interaction between the receptor and a probe atom that is located in the different grid points. AutoGrid builds as many files as the number of probe atoms used. These probes can be the atoms that are present in the ligands which will be docked onto the receptor, generating the corresponding affinity grid maps. Default values were provided with every GPF generated and changes were made where necessary. In this study, blind docking was performed where the grid box, with grid spacing of 0.375 Å and dimension of 128 x 100 x 116 points along the x , y and z axes, was centered on the macromolecule (Figure 5.4a). All rotatable bonds of the protein were made non-rotatable. During the grid calculation, a '*.glg' log file and the map files were written.

Docking parameter files (DPF) were prepared as input files for the AutoDock utility to perform the docking process. AutoDock uses the full set of grid maps built by AutoGrid (*.glg) to guide the docking process of the ligands through the lattice volume. Default values were generally used with some changes made necessary for the study undertaken (Table 5.2). The Lamarckian genetic algorithm (LGA) (Morris *et al.*, 1998) was selected as the search algorithm. A population size of 150 and 250,000 energy evaluations were used for 100 search runs.

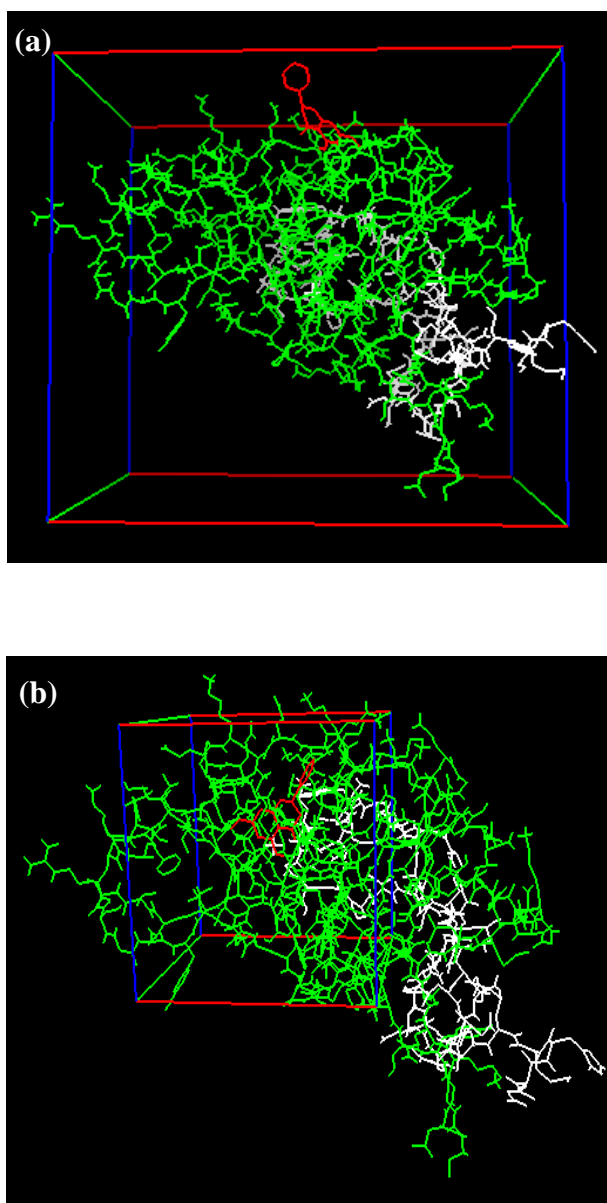


Figure 5.4 Setting up of grid boxes in the building of the grid parameter files prior to (a) rigid-ligand docking, where the box is centred on the macromolecule (blind docking); and (b) flexible-ligand docking, where the box is centred on the ligand. (See section 5.2.2 for details). The ligand is shown in red colour, while the green colour denotes the NS3 chain of DEN-2, and the white colour denotes the NS2B chain of DEN-2.

Table 5.2 An example of the docking parameter file (DPF) containing parameters required for the generation of different ligand poses and calculation of interaction energies between the ligand and the binding site.

```

outlev 1 # diagnostic output level
seed pid time # seeds for random generator
types CAO#H # atom type names
fld 2FOMCa3.maps.fld # grid_data_file
map 2FOMCa3.C.map # atom-specific affinity map
map 2FOMCa3.A.map # atom-specific affinity map
map 2FOMCa3.O.map # atom-specific affinity map
map 2FOMCa3.H.map # atom-specific affinity map
map 2FOMCa3.e.map # electrostatics map
move pstrobnrgd2.out.pdbq # small molecule
about -1.8941 -6.3764 -0.4856 # small molecule center
tran0 random # initial coordinates/A or random
quat0 random # initial quaternion
ndihe 0 # number of active torsions
dihe0 random # initial dihedrals (relative) or random
tstep 2.0 # translation step/A
qstep 50.0 # quaternion step/deg
dstep 50.0 # torsion step/deg
torsdof 0 0.3113 # torsional degrees of freedom and
# coeffiecent
intnbp_r_eps 4.00 0.0222750 12 6 # C-C lj
intnbp_r_eps 4.00 0.0222750 12 6 # C-A lj
intnbp_r_eps 3.60 0.0257202 12 6 # C-O lj
intnbp_r_eps 3.00 0.0081378 12 6 # C-H lj
intnbp_r_eps 4.00 0.0222750 12 6 # A-A lj
intnbp_r_eps 3.60 0.0257202 12 6 # A-O lj
intnbp_r_eps 3.00 0.0081378 12 6 # A-H lj
intnbp_r_eps 3.20 0.0297000 12 6 # O-O lj
intnbp_r_eps 1.90 0.3280000 12 10 # O-H hb
intnbp_r_eps 2.00 0.0029700 12 6 # H-H lj
#
rmstol 0.5 # cluster_tolerance/A
extnrg 1000.0 # external grid energy
e0max 0.0 10000 # max initial energy; max number of retries
ga_pop_size 150 # number of individuals in population
ga_num_evals 250000 # maximum number of energy evaluations
ga_num_generations 27000 # maximum number of generations
ga_elitism 1 # number of top individuals to survive to
# next generation
ga_mutation_rate 0.02 # rate of gene mutation
ga_crossover_rate 0.8 # rate of crossover
ga_window_size 10 #
ga_cauchy_alpha 0.0 # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0 # Beta parameter Cauchy distribution
set_ga # set the above parameters for GA or LGA
sw_max_its 300 # iterations of Solis & Wets local search
sw_max_succ 4 # consecutive successes before changing rho
sw_max_fail 4 # consecutive failures before changing rho
sw_rho 1.0 # size of local search space to sample
sw_lb_rho 0.01 # lower bound on rho
ls_search_freq 0.06 # probability of performing local search on
# individual
set_sw1 # set the above Solis & Wets parameters
ga_run 100 # do this many hybrid GA-LS runs
analysis # perform a ranked cluster analysis

```

All grid maps and docking calculations were performed using the AutoDock 3.0.5 software package. After the docking searches were completed, clustering histogram analyses were performed, based on RMSD (root mean square deviation) of not more than 1.0 Å. The conformation with the lowest docked energy was chosen from the most populated cluster, and put through the following steps. The energy function implemented in the AutoDock 3.0.5 software is described in Appendix 3.

5.2.2.3 Automated flexible-ligand docking

The methods involved were similar to the above (section 5.2.2.2), except that in the ligand file preparation, torsional degrees of freedom were redefined by allowing rotatable bonds to remain flexible. In addition, the grid box size was reduced to a dimension of 60 x 60 x 60 numbers of points, centred on the ligand (Figure 5.4b). The grid box encompassed the region surrounding the selected ligand conformation resulting from the previous docking experiment. The number of energy evaluation was increased to 10 million.

5.2.2.4 Analysis of results

The conformations from the docking experiments were analyzed using Viewerlite 4.2 (<http://www.accelrys.com>), which also identified the H-bonding and van der Waals interactions between the protease and the ligands. A hydrogen bond was measured when the distance between the proton on the donor atom and the acceptor atom was equal to or less than 2.5 Å, and the bond angle is in the range 90° – 180°. Van der Waals interactions were assessed using the van der Waals surface plot application.

Hydrophobic interactions between the protein residues and the ligands were determined using the LIGPLOT (Wallace *et al.*, 1995) program. The HBPLUS (McDonald & Thornton, 1994) program in LIGPLOT calculates (by default) non-bonded / hydrophobic contacts, with cut-off of 2.9 – 4.3 Å. The LIGPLOT algorithm reads in the 3D structure of the ligands from the PDB files, together with the protein residues, and ‘unrolls’ each object about its rotatable bonds, flattening them out onto the 2D page.

Hyperchem Pro 6.0 software was used to calculate single point energy of the ligands with the most favourable conformations using the semi empirical AM1 method. The electrostatic potentials of the ligands were displayed as contour maps plotted in the form of 3D isosurfaces. By default, Mulliken charges were used in the calculations to obtain the electrostatic potentials.

5.3 Results

5.3.1 Automated rigid-ligand docking

The ligands were optimized prior to docking to DEN-2 NS2B-NS3 and the single point energies calculated are tabulated in Table 5.3. The minimized energies shown are for structures in which the optimization process reached the convergence criteria (see methods, section 5.2.2.1).

The clustering histograms extracted from the output log files (*.dlg) for the blind docking of the rigid ligands are appended as Appendix 4. The estimated free energy of binding (ΔG_{bind}) and final docked energy (E_{dock}) of the selected conformations are shown in Table 5.4. The best conformations were selected based on the lowest

Table 5.3 Single-point energy values of ligands calculated using AM1 method (semi empirical) available in Hyperchem Pro 6.0 software (Hypercube Inc.).

Compounds	Before minimization		After minimization		Δ Energy (kcal/mol)
	Energy (kcal/mol)	Gradient (kcal/mol Å)	Energy (kcal/mol)	Gradient (kcal/mol Å)	
Pinostrobin	-3085.55	248.02	-3760.10	33.91	-674.55
Pinocebrin	-3357.14	87.56	-3492.10	36.27	-134.96
Alpinetin	-3593.49	86.25	-3758.76	35.89	-165.27
Pinostrobin Chalcone	-3363.85	178.20	-3745.35	35.41	-381.50
Pinocebrin Chalcone	-3273.73	87.70	-3479.38	36.42	-205.65
Cardamonin	-2864.79	365.84	-3749.61	34.77	-884.82

Table 5.4 Estimated free energy of binding and final docked energy calculated using AutoDock 3.0.5 for rigid and flexible-ligand docking to DEN-2 NS2B-NS3.

Compounds	No. of atoms	No. of free torsions		Min. estimated ΔG_{bind} (kcal/mol)		Final E_{dock} (kcal/mol)	
		Rigid	Flexible	Rigid	Flexible	Rigid	Flexible
Pinostrobin	34	0	3	-8.23	-8.30	-8.23	-8.82
Pinocembrin	31	0	3	-8.63	-9.25	-8.63	-9.57
Alpinetin	34	0	3	-9.17	-8.84	-9.17	-9.45
Pinostrobin Chalcone	34	0	6	-7.74	-7.79	-7.74	-8.71
Pinocembrin Chalcone	31	0	6	-7.40	-7.59	-7.40	-8.85
Cardamonin	34	0	6	-8.00	-7.55	-8.00	-9.04

For flexible-ligand dockings, the starting conformations were the best docked conformations obtained from rigid-ligand docking procedures (refer to materials and methods section).

docked energy and the relatively high number of conformations in a cluster.

The rigid docking of pinostrobin to DEN-2 NS2B-NS3 resulted in two favourable ligand poses, ie. rank 1 (run #30) and rank 2 (run #69) (refer Appendix 4a). Figure 5.5 illustrates the superimposition of the two conformations which showed the two structures to be very similar and bind to the same region on the protein. Reclustering of all the resulting conformations was then performed at 1.0 Å and resulted in the same pose from rank 1 (run #30) as the most favourable conformation (refer Appendix 4a). This conformation was then used for subsequent flexible-ligand docking. For the case of the other ligands, the selection process for the best docked conformation of each ligand was fairly straightforward (refer to Appendix 4).

5.3.2 Automated flexible-ligand docking

Except for pinostrobin chalcone, the selections of the best docked conformations of the ligands were based on the lowest docked energy and the relatively high number of conformations in a cluster (Appendix 5). For pinostrobin chalcone, two favourable ligand poses were obtained, ie. rank 1 (run #67) and rank 11 (run #65) (refer Appendix 4a). Figure 5.6 illustrates the superimposition of the two conformations which indicates that the two structures bind to the same region on the protein. Reclustering of all the resulting conformations was then performed at 1.0 Å and yielded the same pose to that from rank 11 (run #65) as the most favourable conformation (data not shown). Hence, this conformation was used for the analysis of binding interactions.

Results for the flexible-ligand docking to DEN-2 NS2B-NS3 are summarised in Table 5.5. The ligands are arranged according to decreasing inhibition activity

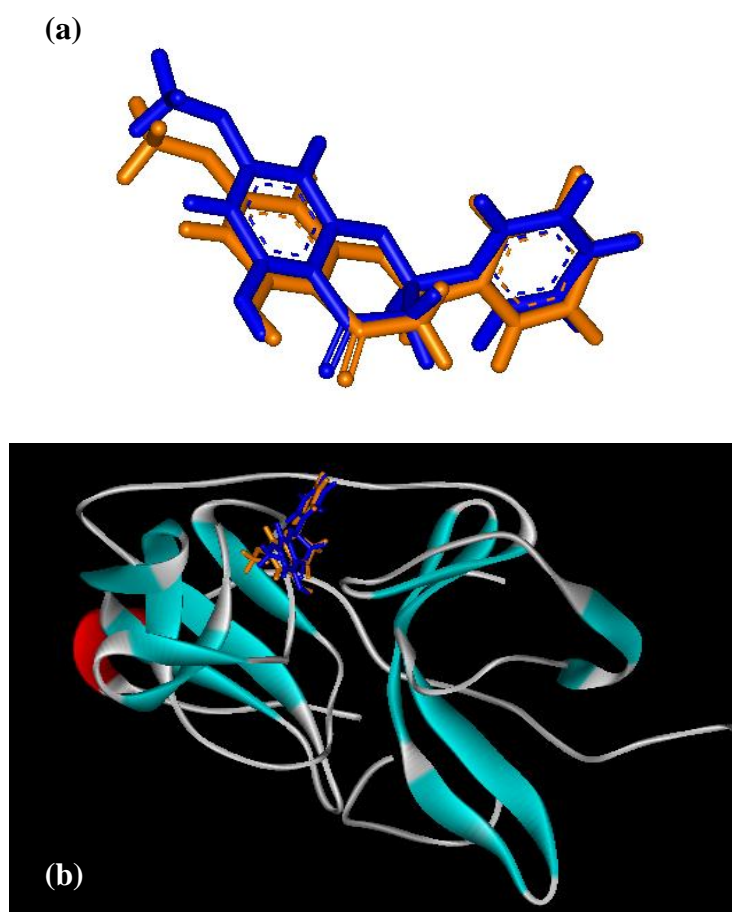


Figure 5.5 (a) Superimposition of pinostrobin conformations as a result of rigid-ligand docking to DEN-2 NS2B-NS3. The ligands are shown as sticks. Pose from rank 1 (run #30) is in orange colour and pose from rank 2 (run #69) is in blue colour (refer Appendix 4a). (b) The two conformations bind to the same region of the protein.

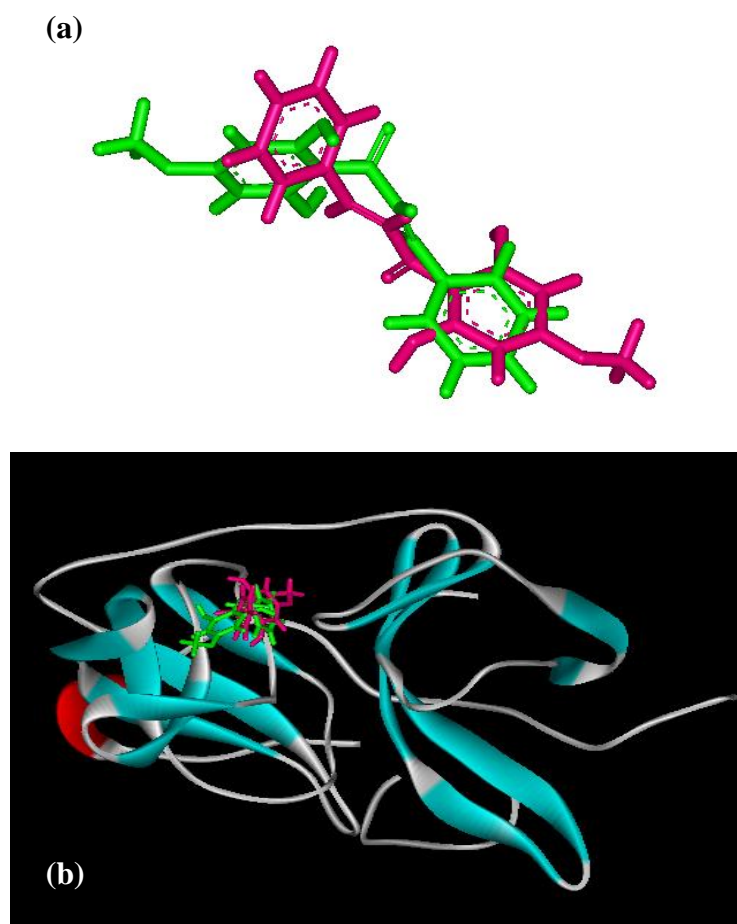


Figure 5.6 (a) Superimposition of pinostrobin chalcone conformations as a result of flexible-ligand docking to DEN-2 NS2B-NS3. The ligands are shown as sticks. Pose from rank 1 (run #67) is in green colour and pose from rank 11 (run #65) is in pink colour (refer Appendix 5d). (b) The two conformations bind to the same region of the protein.

Table 5.5 Results of automated flexible-ligand docking to DEN-2 NS2B-NS3 calculated using AutoDock 3.0.5.

	Pinostrobin	Cardamonin	Alpinetin	Pinocebrin chalcone	Pinocebrin	Pinostrobin chalcone
Experimental IC ₅₀ (μg/ml)	90.48	235.86	242.76	273.10	286.90	*
Min. estimated free energy of binding, ΔG _{bind} (kcal/mol)	-8.30	-7.55	-8.84	-7.59	-9.25	-7.79
Min. docked energy, E _{dock} (kcal/mol)	-8.82	-9.04	-9.45	-8.85	-9.57	-8.71
Estimated inhibition constant, K _i (μM)	0.83	2.92	0.33	2.71	0.17	1.95
Final intermolecular energy, ΔG _{inter} (kcal/mol)	-8.92	-8.80	-9.46	-8.53	-9.56	-9.03
Final internal energy of ligand, ΔG _{intra} (kcal/mol)	+0.10	-0.25	+0.02	-0.32	0.00	+0.33
Torsional free energy, ΔG _{tor} (kcal/mol)	+0.62	+1.25	+0.62	+0.93	+0.31	+1.25
RMSD (from ref structure) (Å)	0.61	3.15	0.60	2.05	0.88	4.60

* The IC₅₀ value for pinostrobin chalcone could not be determined from experiment due to its low inhibition activity (Tan, 2005).

The IC₅₀ values were obtained from experiments performed by Tan (2005). The ligands are arranged according to decreasing inhibition activity (increasing IC₅₀ values) from left to right. Estimated energy and K_i values shown were obtained from flexible docking of each ligand conformation which was initially the best docked conformation obtained from rigid docking procedures (refer to materials and methods section).

(increasing IC_{50} values) from left to right of the table. The IC_{50} (ligand concentration at 50% inhibition) values were obtained from experimental setup (Othman *et al.*, 2008; Tan, 2005). Figure 5.7 illustrates the superimposition of the best pose of each ligand at the binding site as a result of the flexible-ligand docking to the protease. Presumably, these ligands bind to a similar region of the protease resulting in non-competitive inhibition activities towards protease replication.

5.4 Discussion

5.4.1 Non-competitive inhibition of DEN-2 NS2B-NS3

Tan *et al.* (2005, 2006) reported the non-competitive inhibitory activities of these ligands (refer Figure 5.1) tested at three different concentrations (120, 240 and 400 $\mu\text{g/ml}$). Results indicated pinostrobin to be the most active inhibitor, with up to 85.4 % inhibition at 400 $\mu\text{g/ml}$ (Tan *et al.*, 2006; Othman *et al.*, 2008), while its chalcone derivative, pinostrobin chalcone, did not exhibit any activity. However, in the other two cases, chalcones were observed to be more active than their corresponding flavanones. The IC_{50} (ligand concentration at 50 % inhibition) values are shown in Table 5.5, arranged in decreasing activity from left to right. The IC_{50} value of pinostrobin (90.48 $\mu\text{g/ml}$) is 2.6 to 3 fold smaller than that of the other ligands, indicating it to be the most active non-competitive inhibitor.

5.4.2 Predicted free energy of binding, docking energy and K_i values

AutoDock has been reported to be the most popular docking program (Sousa *et al.*, 2006) due to its high accuracy and versatility. Through blind docking, Autodock is able to select the correct protein-ligand complexes, based on their energy, without

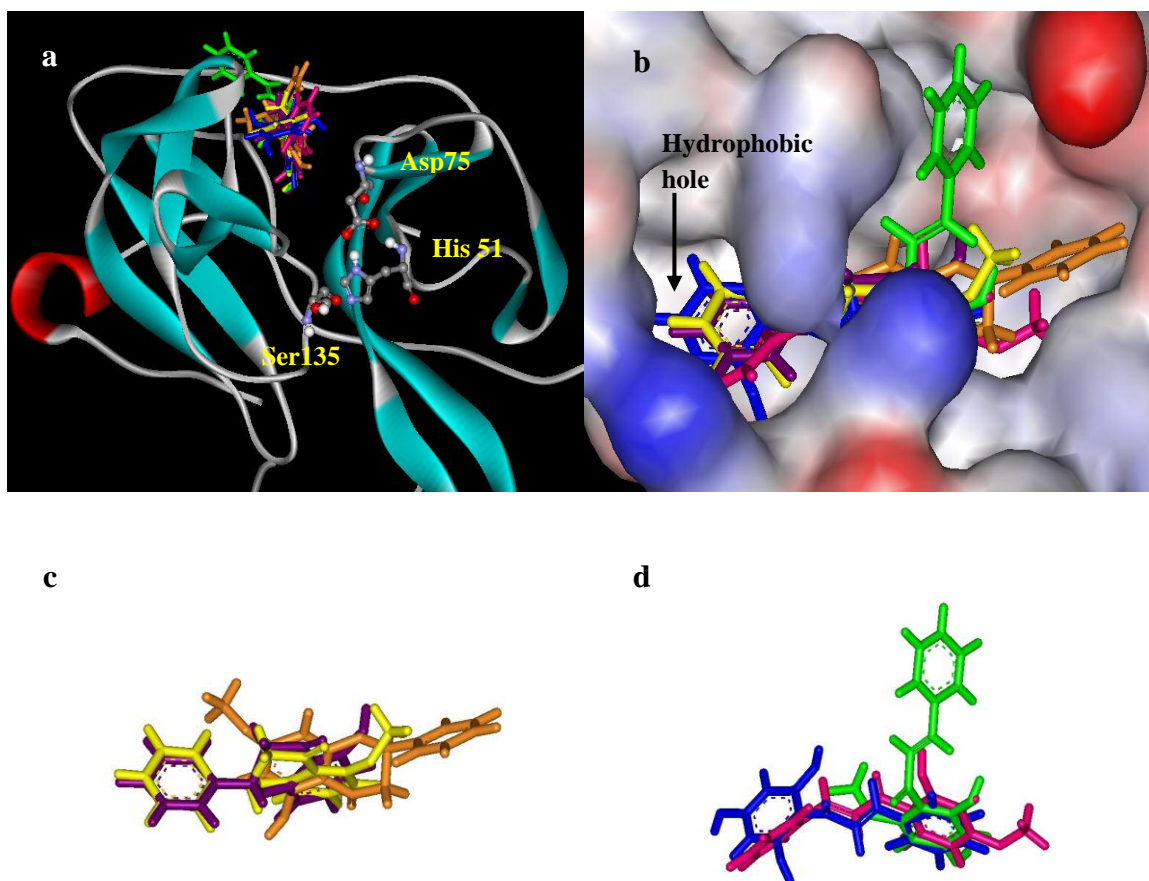


Figure 5.7 Computer generated models illustrating the superimposition of the flavanones: Pinostrobin (orange), pinocembrin (purple) and alpinetin (yellow), and the chalcones: pinostrobin chalcone (pink), pinocembrin chalcone (blue) and cardamonin (green), at the binding site. The ligands are shown as sticks. (a) All the ligands are superimposed at the binding site. DEN-2 protease is represented as ribbons. The catalytic triads labelled as His51, Asp75 and Ser135 are shown as balls and sticks. (b) Connolly surface representation of the binding site coloured according to electrostatic potential spectrum. (c) Superimposition of the flavanone poses as found in the binding site. (d) Superimposition of the chalcone poses as found in the binding site.

prior knowledge of the binding site (Hetényi & Spoel, 2002). Autodock has been proven to be efficient and robust in finding the binding pockets and binding orientations of the ligands, whether they are rigid (up to at least 30 heavy atoms) or flexible. In this study, rigid-ligand docking was initially performed using energetically-optimized ligands to different potential binding sites of the whole protein (blind docking). As the ligands involved in this study were non-competitive inhibitors, it was anticipated that they would dock to sites other than the active site, thus requiring blind docking. Once the binding site for non-competitive inhibition was identified, flexible-ligand docking was carried out at this site using Autodock. For this, a grid box of 60 x 60 x 60 number of points (grid spacing of 0.375 Å) was built around the binding region.

Table 5.4 shows the calculated docking energy (E_{dock}) and free energy of binding (ΔG_{bind}) for both the rigid and flexible-ligand docking simulations. On the whole, energy values obtained from flexible docking were lower than those from rigid docking. Allowing the bonds in the ligands to rotate (during flexible docking) generated different ligand conformations, enabling more refined search for preferred binding sites. However, the calculated energies and K_i (inhibitory constant) values did not follow the pattern of activity-ranking according to the observed IC_{50} values. In another docking study performed by our collaborator (data and results not shown here) (Othman *et al.*, 2008) using the Glide software (Friesner *et al.*, 2004), pinostrobin is also observed as the most active inhibitor with a high Glide score of -7.14 (Glide was able to identify the best inhibitor, but not the worst. Glide uses a Chemscore-derived algorithm to rank ligand binding; the more negative the score, the stronger the binding). Since cardamonin, alpinetin, pinocembrin chalcone and pinocembrin had IC_{50} values between 200 and 300 µg/ml, docking computations is not expected to discriminate between these ligands. This highlights the challenges associated with scoring compound

conformations and accuracy in predicting activity rank through computational modelling. Enthalpic and entropic effects drive ligand-binding processes and either of these effects could dominate specific interactions (Kitchen *et al.*, 2004). However, if binding induces conformational changes in the protein, as often seen in enzymatic interactions, using rigid binding sites limits the predictive ability to link experimental activity-ranking with calculated scoring functions.

In general, the estimated ΔG_{bind} for flavanones obtained in this study are lower than those of their chalcone derivatives (Table 5.5). This can be attributed to chalcones having more rotatable bonds, increasing the torsional free energy (ΔG_{tor}) and lowering binding affinity (higher K_i values). On the other hand, the calculated K_i values for all the ligands are reasonably small (within the μM range), indicating the formation of stable enzyme-inhibitor complexes.

Since there is no direct correlation between the ranking of the scoring functions and the pattern of activity-ranking (according to the observed IC_{50} values), as mentioned earlier, quantitative explanation of the experimental inhibitory activity of the ligands towards substrate-binding by the protease would require further structural insights. These insights are discussed in detail hereafter.

5.4.3 Binding site and ligand conformations

As expected, the automated docking placed all the ligands studied out of the active site of the protease. Surprisingly, however, the favourable binding site for all these ligands was confined to a specific region in the protease (Figures 5.7a and b). Presumably, in this case, the shape of the binding site complements the shapes of the

ligands. The region of the binding site can be divided into three parts (Figure 5.7b): the left constitutes a hydrophobic hole which accommodates the aromatic rings of the ligands, while the middle region allows placement of flavanone ring C or the chains carrying the enone groups of the chalcones. This enabled the ligands to interact with the surrounding residues. The region on the right is larger than the other two regions. This allowed room for the ligand to take up more poses. Figures 5.7c and d showed the superimposition of the flavanones and chalcones at their docked conformations, respectively, at the binding site. Of the flavanones, pinocembrin and alpinetin took up similar poses. Whilst pinostrobin lies on a similar axis as the other two ligands, it is oriented in the opposite direction from the other two flavanones, where the phenyl ring is placed on the opposite direction as opposed to those of pinocembrin and alpinetin.

The conformational differences between pinostrobin and the other flavanones could be seen more clearly in Figure 5.8 which illustrated the 3D isosurface plots of the electrostatic potential of the ligands. Similar plots were observed for alpinetin and pinocembrin. However, the potential map for pinostrobin (coloured green to indicate neutral to positive potential) is more elongated due to the presence of a methoxy group on ring A. The negative potentials (magenta) due to the carbonyl and hydroxyl groups on ring C and ring A, respectively, pointed towards the back of the plane of the model, while the ether group on ring C pointed more towards the front of the plane of the model. The reverse is observed with alpinetin and pinocembrin, where the negative potentials due to the carbonyl and hydroxyl/methoxy groups on ring C and ring A, respectively, pointed towards the front of the plane of the model, while the ether group on ring C is pointed more towards the back of the plane of the model. These observations inferred a different chemical environment requirement for the binding of pinostrobin compared to those for alpinetin and pinocembrin.

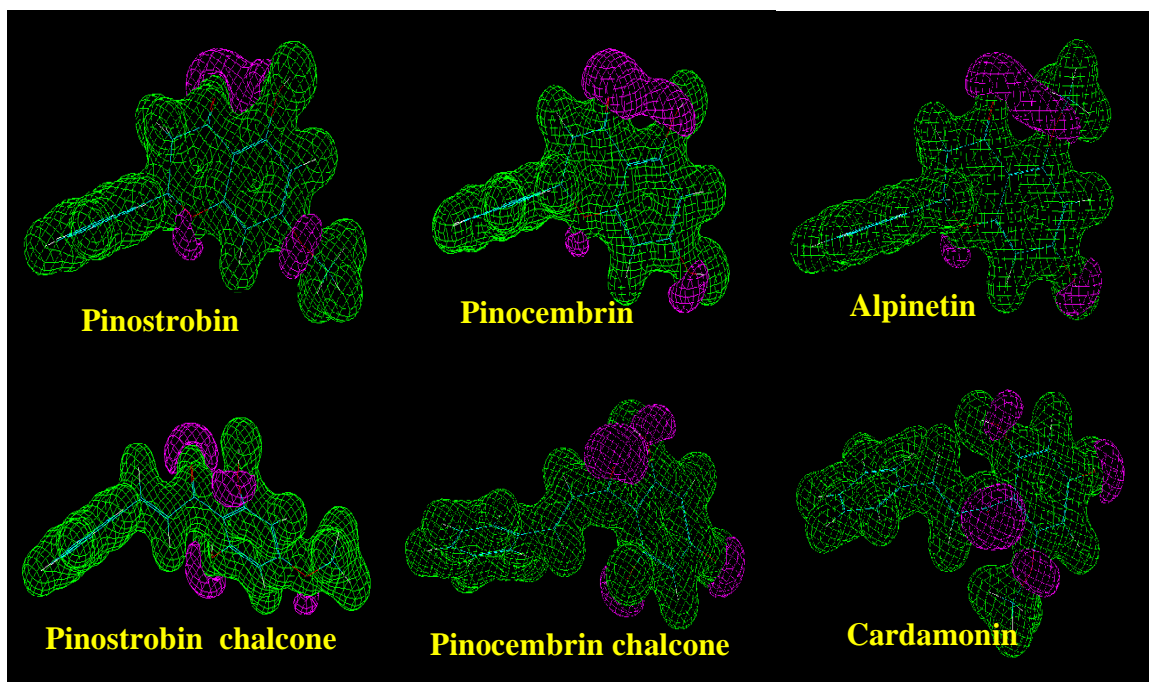


Figure 5.8 3D isosurface plots of the electrostatic potentials of the ligands, using Hyperchem Pro 6.0 software. Green surfaces indicate neutral to positive potentials while magenta surfaces indicate negative potentials.

For the chalcones, it can be seen that the orientation of cardamonin is on an axis almost perpendicular to those of pinostrobin chalcone and pinocembrin chalcone (Figure 5.7d). This could be due to the presence of the methoxy group on C6' of cardamonin, which made the molecule adopt a slightly distorted conformation, hence required an apparent different pose of binding. In addition, the shapes of the 3D isosurface plots of the electrostatic potential are different for each ligand (Figure 5.8). This could be attributed to the higher flexibility of the enone chains. This observation may also explain the requirement for a different (chemical) environment of the binding site by the ligands, hence the different orientations adopted by the ligands in the binding site.

5.4.4 Binding interactions

5.4.4.1 Pinostrobin

Table 5.5 showed pinostrobin to be the best inhibitor amongst all the ligands tested, with an IC_{50} value of 90.48 $\mu\text{g/ml}$. On the other hand, pinostrobin chalcone was observed to be inactive. Figures 5.9a and b illustrated the binding site for pinostrobin. The shape of the binding pocket complemented the pose of the ligand where the phenyl ring is protruded out towards the surface of the protein, while the rest of the molecule was embedded into the inner part of the protein. The surrounding residues involved in hydrogen bonding interactions with pinostrobin are Lys74, Leu149 and Asn152. The H-bonding interaction with Lys74 (Figures 5.9c and d) is not observed with the other flavanones and chalcones studied. The interaction between the hydroxyl H atom on ring A of pinostrobin with the backbone carbonyl O atom of Lys74 could account for the relatively high inhibition activity of pinostrobin. Since Lys74 is directly bonded to Asp75, the formation of H-bond between Lys74 and pinostrobin may directly induce

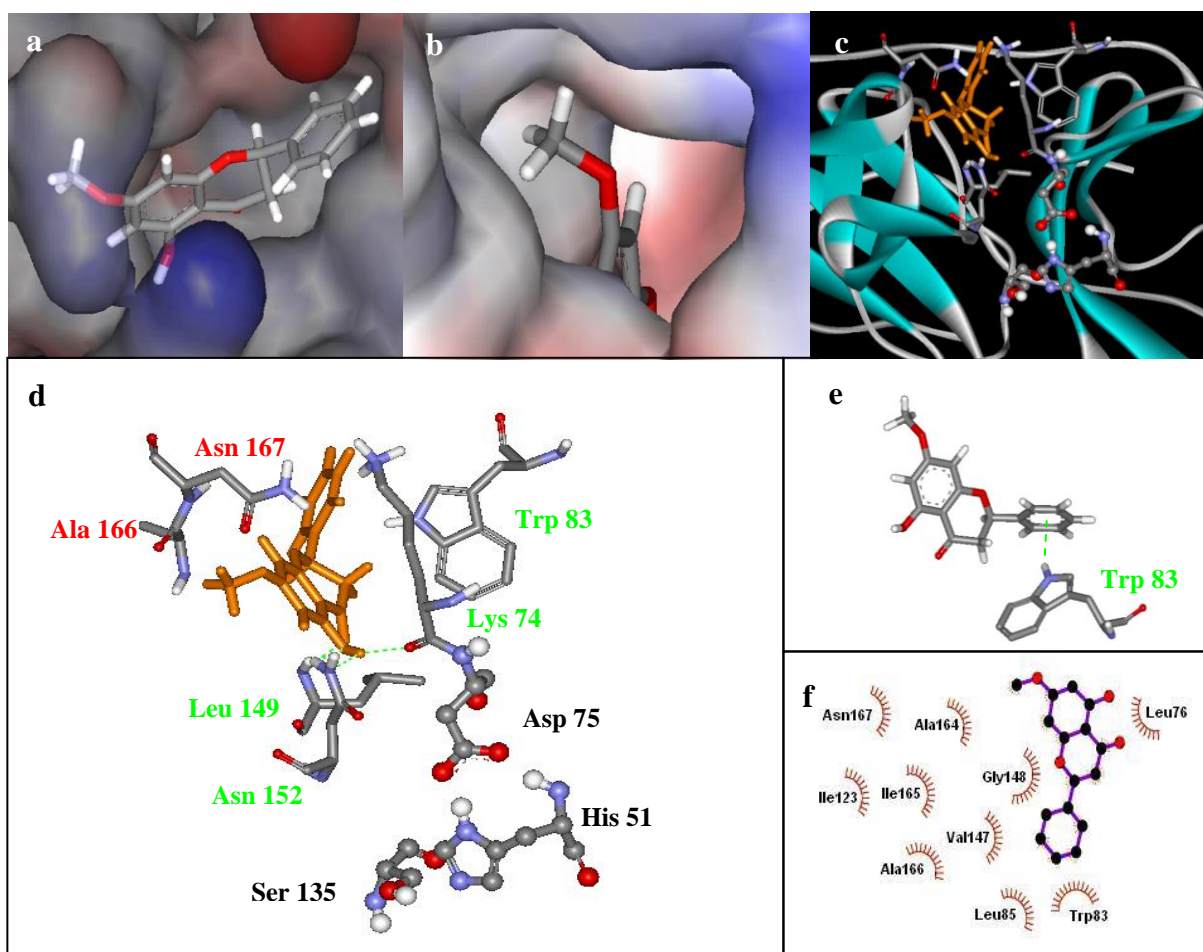



Figure 5.9 (a) Transparent Connolly surface representation of pinostrobin at the binding site. (b) View of pinostrobin in the binding site from a different angle. (c) View of pinostrobin (orange) at the binding site of DEN-2 protease (ribbons). Residues interacting with the ligand are shown as sticks. Catalytic triads are shown as balls and sticks. (d) Simplified view of pinostrobin interacting with surrounding residues. Residues labelled in green interact with the ligand via H-bonds, while those in red interact via van der Waals contacts. The catalytic triads are labelled in black. (e) Non-classical H-bonding between Trp83 and phenyl ring B of pinostrobin. (f) 2D schematic diagram of residues in the binding site which exhibit hydrophobic interactions with pinostrobin, obtained using the Ligplot program. Keys for the plot are: ●—● ligand bond;  non-ligand residue involved in hydrophobic contact; ● corresponding atom involved in hydrophobic contact.

conformational change in Asp75, in particular, or the catalytic triad region, in general. This would then disrupt the electron transfer process required for substrate binding at the active site, thus affecting the activity of the protease.

Studies have shown the formation of low-barrier H bond between Asp and His (of the catalytic triad) to facilitate the nucleophilic attack by the β -OH group of Ser on the acyl carbonyl group of substrates (Frey *et al.*, 1994; Hunkapiller *et al.*, 1976; Santis & Carloni, 1999). (Low-barrier H bond is considered to be strong H bond with bond energy of 12 to 24 kcal/mol (Frey *et al.*, 1994)). Besides Lys74, the residue Leu149 may also play a role in the activity of pinostrobin in two ways. First, the entry of the ligand into the active site could be blocked by Leu149 due to its position in the protease (Figures 5.9c and d). Second, Leu149 is observed to protrude towards the adjacent β -barrel which carried the residues Asp75 and His51. Erbel *et. al.* (2006) reported that the NS3 protease domain adopts a chymotrypsin-like fold with two β -barrels and that the catalytic triad is located at the cleft between the two β -barrels. Interaction between Leu149 with pinostrobin may cause a conformational change of the residue in order to reduce steric clashes. This would affect the spatial conformation of the surrounding residues, in particular the catalytic triad. Subsequently, the electron transfer process between Asp75 and His51 may be affected, reducing the capability of the active site to bind with the substrate. This process could further increase the inhibition capability of pinostrobin.

Figure 5.9f shows the residues involved in hydrophobic interactions with pinostrobin obtained by the Ligplot program. In addition to the hydrophobic interactions, Trp83 also exhibited non-classical hydrogen bonding interaction between the H atom on N ϵ 1 of its indole ring with the phenyl ring (B) of pinostrobin (Figure

5.9e), with the hydrogen bond acceptor in this case being the aromatic ring (Gervasio *et al.*, 2002). Brocchieri and Karlin (1994) reported that the interplanar angle, α (dihedral angle between the extended planes of interacting planar groups), between phenyl ring of the residue Phe and the aromatic ring of Trp to be favourable at $30^\circ < \alpha < 90^\circ$. They also reported that, generally, planar interactions of Trp involved mostly the five-atom ring which is capable of forming a hydrogen bond (involving its imino group), engaging π -cloud electrostatic interactions and also undergoing hydrophobic interactions. Electrostatic charges associated with phenyl ring include a weak negative charge about the centre of the aromatic ring and a weak positive charge projected at the ring periphery (Burley & Petsko, 1988). In this study, similar interactions were observed between Trp83 and pinostrobin where the bond distance between the H atom on N ϵ 1 of Trp83 and the centre of the ring of the ligand (Rcentroid) was 2.67 Å, and α was 60.7° (Figure 5.10).

5.4.4.2 Pinostrobin chalcone

Previous studies have shown pinostrobin chalcone to be inactive towards the DEN-2 serine protease (Table 5.5) (Tan, 2005). Investigation into the electrostatic interactions of pinostrobin chalcone with the binding site revealed the surrounding amino acid residues of the protease to be involved in the interactions via van der Waals and H-bond. However, they were confined to the C-terminal region of the protease, with residues ranging from Asn152 to Asn167 (Figure 5.11a). This binding mode does not seem to have any structural effect on the catalytic triad in promoting the disruption of electron transfer for the initiation of proteolytic processing. Presumably, binding activity of this ligand with its surrounding residues is confined to a region which does not impose any conformational change to the active site. Figure 5.11b illustrates the

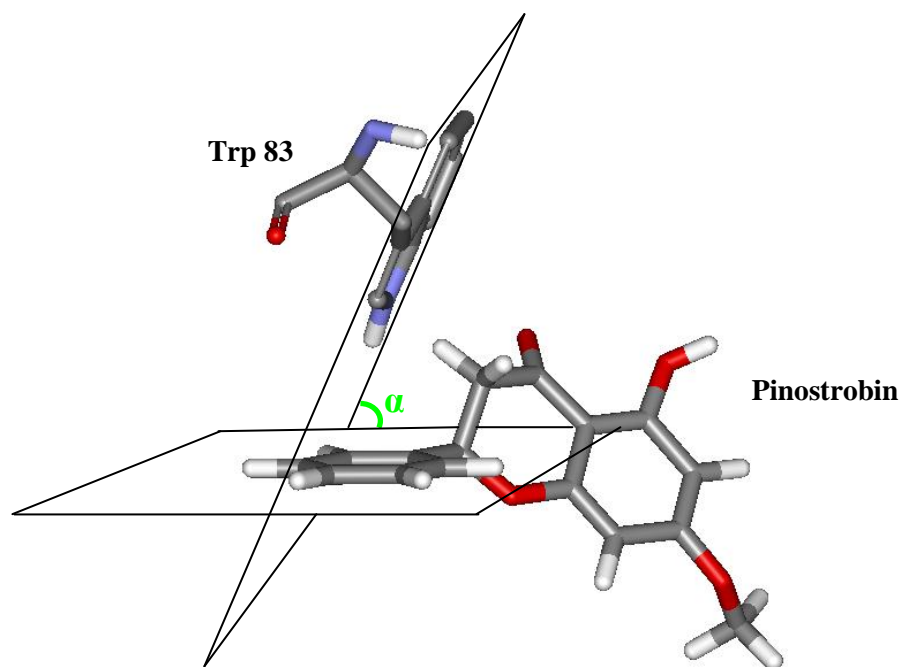


Figure 5.10 Diagram illustrating the interplanar angle (dihedral angle between the extended planes of interacting planar groups), α , between the phenyl ring B of pinostrobin and the aromatic ring of Trp 83 of the DEN-2 protease.

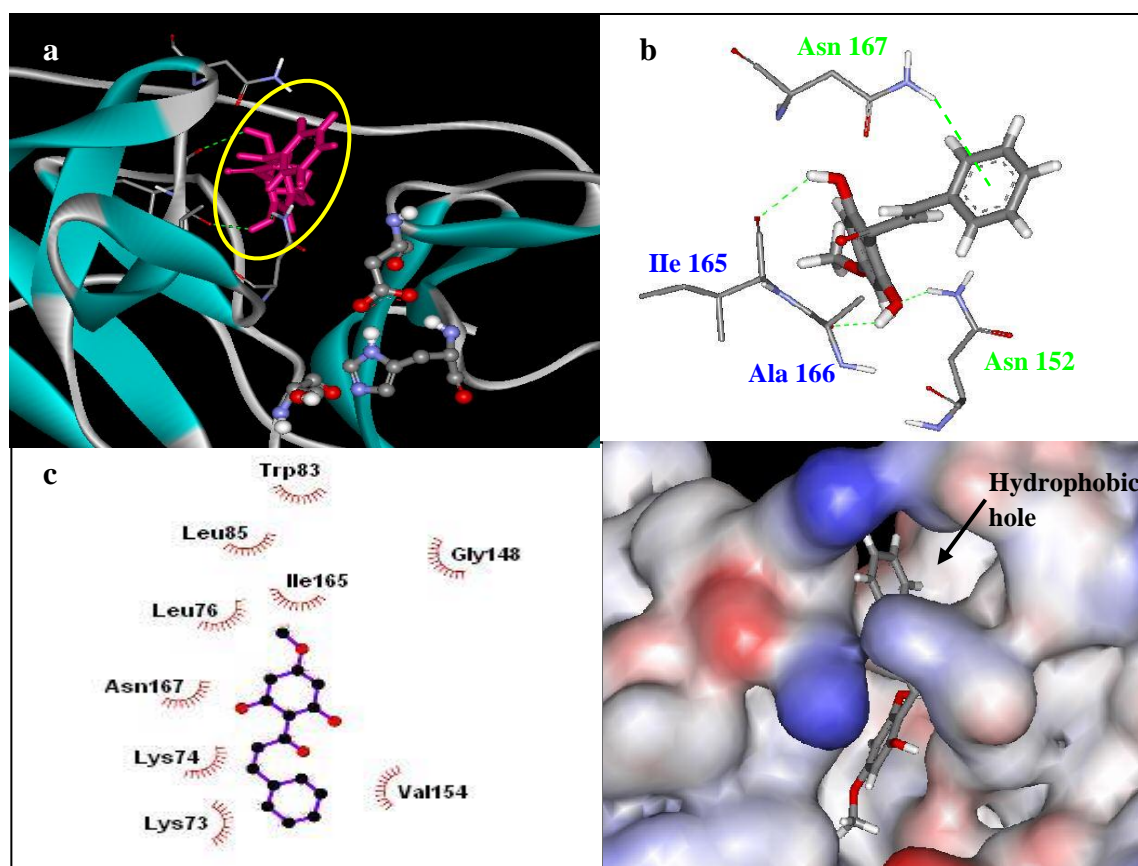



Figure 5.11 (a) View of pinostrobin chalcone (pink) at the binding site of DEN-2 protease (ribbons). Residues interacting with the ligand are shown as sticks. Catalytic triads are shown as balls and sticks. The yellow circle highlights the binding region of the ligand which is confined to the C-terminal region of the protease, away from the catalytic triads. (b) Simplified view of pinostrobin chalcone interacting with surrounding residues. Residues labelled in green interact with the ligand via H-bonds, while those in blue exhibit both H-bond and van der Waals interactions. (c) 2D schematic diagram of residues in the binding site which exhibit hydrophobic interactions with pinostrobin chalcone obtained using the Ligplot program. Keys for the plot are: ●—● ligand bond;  non-ligand residue involved in hydrophobic contact; ● corresponding atom involved in hydrophobic contact. (d) Connolly surface representation of pinostrobin chalcone at the binding site.

residues involved in H-bonding and van der Waals interactions with pinostrobin chalcone, where, a non-classical H-bond is also observed between Asn167 and the phenyl ring (B) of the ligand ($R_{\text{centroid}} = 3.64 \text{ \AA}$, $\alpha = 40^\circ$). Figure 5.11c showed the residues involved in hydrophobic interactions with the ligand obtained from the Ligplot program. In addition to these interactions, the shape of the binding site complemented the ligand's pose. From Figure 5.11d, the hydrophobic pocket which accommodated the phenyl ring (B) of the ligand showed the ring to not be placed directly in the centre of the pocket, rather, to be bent towards the left. This orientation could again be attributed to the non-classical H-bond interaction between the ring and Asn167.

5.4.4.3 The other ligands

Figures 5.12a to d showed the orientations of cardamonin, alpinetin, pinocembrin and pinocembrin chalcone in the binding site of the protease. Except for pinocembrin chalcone, all other ligands experience H-bond and van der Waals interactions with almost the same surrounding residues. The common residues forming H-bond with the three ligands are Leu149 and Asn152. Pinocembrin was also seen to form additional H-bond with Val147 and Ile165. In terms of the van der Waals interactions, cardamonin interacted with Ile165, alpinetin interacted with Val147 and Ile165, while pinocembrin with Asn167. Pinocembrin chalcone, however, did not demonstrate any van der Waals interaction with the surrounding residues, rather it interacted with Lys73 and Asn167 via H-bond instead.

The degree of inhibition offered by these four ligands was similar to each other, although both cardamonin and alpinetin exhibited slightly better activities than pinocembrin chalcone and pinocembrin (Table 5.5). The similarity in the degree of

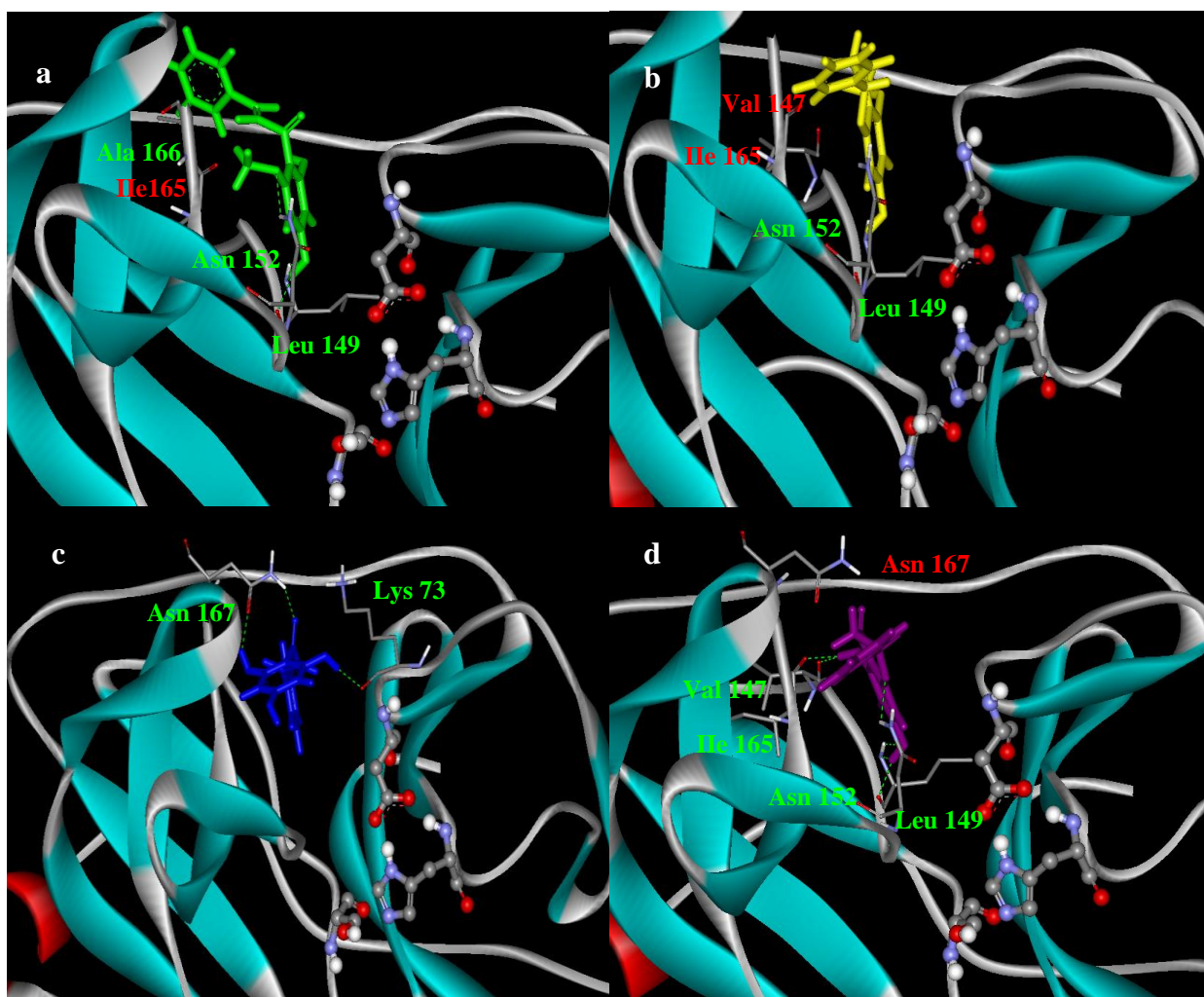


Figure 5.12 Models of ligands bound to DEN-2 protease (ribbons) at the binding sites. The ligands shown are (a) cardamonin (green), (b) alpinetin (yellow), (c) pinocembrin chalcone (blue), and (d) pinocembrin (purple). Residues interacting with the ligands are shown as sticks. Catalytic triads are shown as balls and sticks. Residues labelled in green interact with the ligands via H-bonds, while those in red interact via van der Waals contacts.

activities shown may be explained by the similar axis of orientation in the binding site (except for cardamonin; Figure 5.7; section 5.4.3), and similar mode of interactions as described above (except for pinocembrin chalcone; Figure 5.11). Unlike pinostrobin, these ligands do not form H-bond with Lys74 which may have resulted in the reduced activity observed. Nevertheless, Lys74 is still involved in hydrophobic interaction with alpinetin, pinocembrin and pinocembrin chalcone. Figure 5.13 illustrates the residues involved in hydrophobic interactions with the ligands using the Ligplot program. The observed activities of cardamonin, alpinetin and pinocembrin could be due to the interactions of these ligands with Leu149, as shown in Figures 5.12a, b and d, respectively. The effect of Leu149 on the substrate-binding capability of the protease upon its binding with the ligands was discussed earlier (refer section 5.4.4.1).

A different mode of interaction is observed with pinocembrin chalcone which did not bind to Leu149 to exhibit similar inhibition effect. Rather, it is believed that the H-bonding interaction with Lys73 to be the cause of the activity observed (Figure 5.12c). Lys73 is two residues downstream from Asp75 and interaction between the ligand with this residue may contribute directly to the conformational change of the active site, the extent of which, however, is not as great as that compared to the interaction between pinostrobin and Lys74. The fact that pinocembrin chalcone does not exhibit van der Waals interactions with the surrounding residues may be the cause for the lower activity observed with the pinocembrin chalcone.

5.4.5 Structure-activity relationship (SAR) analysis

The docking experiments performed in this study provided better structural insights and understanding into how the various ligands interacted with the DEN-2

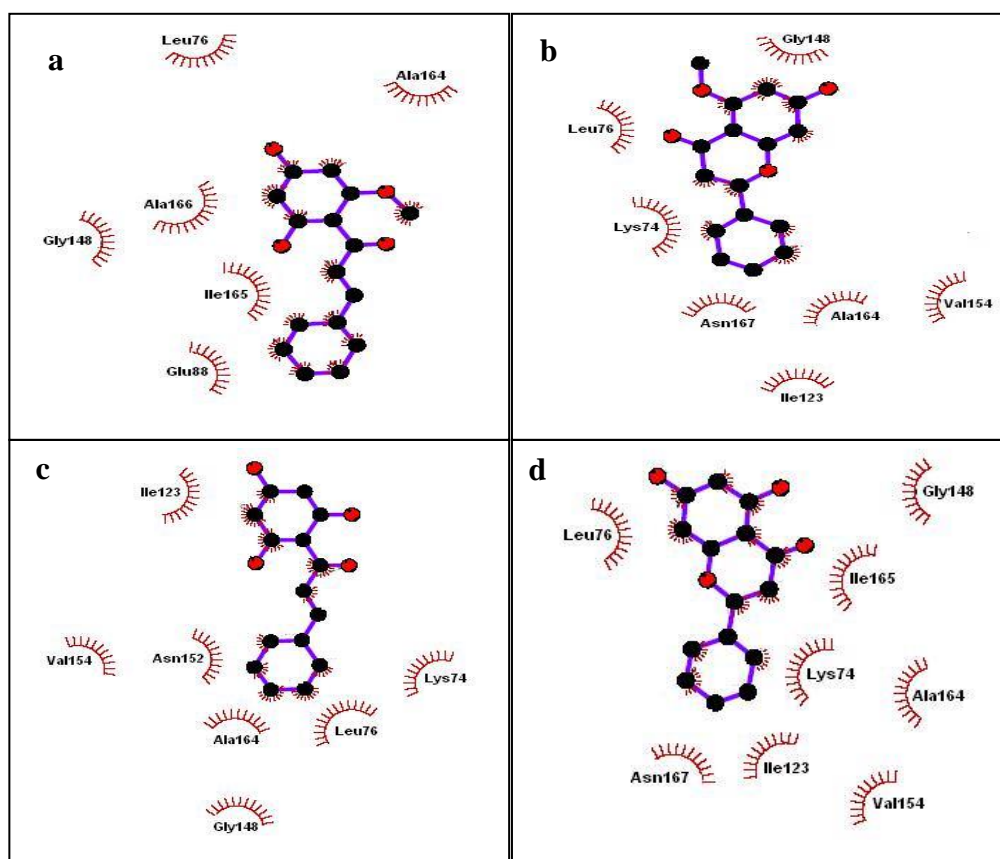





Figure 5.13 Schematic diagrams (2D) illustrating residues in the binding sites which are involved in hydrophobic interactions with: (a) cardamomin, (b) alpinetin, (c) pinocembrin chalcone, and (d) pinocembrin. Keys for the plot are:  ligand bond;  non-ligand residue involved in hydrophobic contact;  corresponding atom involved in hydrophobic contact.

serine protease which resulted in a non-competitive inhibition of the protease activity. Analyses of the structure-activity relationships between the ligands and the binding site have shed some light onto the important structural and conformational information which could be applied in designing new compounds as inhibitors to this serine protease.

This study affirmed that rigid conformation of the flavanones ensured ligand activities. Opening of the ring C, as in chalcones, would make the ligands more flexible. This led to an extended conformation which then required a different environment at the binding site (Figure 5.8). Higher flexibility of the chalcones also meant higher ΔG_{tor} compared to the corresponding flavanones, and hence, higher ΔG_{bind} . This resulted in chalcone binding to be less favourable to this site when compared to the corresponding flavanones (Tables 5.4 and 5.5). Pinocembrin showed the lowest values for the E_{dock} and ΔG_{bind} , followed by alpinetin and pinostrobin, indicating the parent structure of 7-hydroxyflavanone to have good binding capabilities with the protease.

The methoxy group at C7 on the ring A of pinostrobin may be important in creating preferable electrostatic potential surface of the molecule in 'search' of the optimal chemical environment within the binding site. In general, electron-donating group at this position would be preferable for the design of new inhibitory compounds. The hydroxyl group on the ring A of pinostrobin could be considered to be an important pharmacophore where it could be involved in forming H-bonds with the surrounding residues, for example between pinostrobin (C5) and Lys74, and between pinocembrin chalcone and Lys73 (the OH group is on C2' of ring A for pinocembrin chalcone). For pinocembrin chalcone, the number of electrostatic interactions involved was the lowest among all the ligands reported. Thus, pinocembrin chalcone is expected to exhibit a

much lower inhibitory activity than the other ligands (except pinostrobin chalcone). However, experimental data showed its activity to be comparable to those of pinocembrin, cardamonin and alpinetin. This could be attributed to its H-bonding interaction with Lys73 via the hydroxyl group, where Lys73 is positioned two residues downstream from Asp75. The observations obtained in this study with regards to the importance of the methoxy and hydroxyl groups in determining the activity of the ligands are in accordance with those postulated by Tan (2005).

The presence of the phenyl ring (B) in all the ligands is essential in contributing to the hydrophobicity of the ligands. In general, hydrophobic interactions between a ligand and its binding site contribute to the increase in entropy of the system. The entropic contribution is represented by the desolvation energy term of the ΔG_{bind} function implemented in AutoDock 3.0.5 (Morris *et al.*, 1998) (Appendix 3). In addition, as discussed previously, the phenyl ring of pinostrobin and pinostrobin chalcone also contributed to the non-classical H-bonding with Trp and Asn, respectively.

5.5 Conclusion

Three flavanones and three chalcones isolated from the plant *Boesenbergia rotunda* L. were docked onto the DEN-2 NS2B-NS3 protease using the AutoDock 3.0.5 software. Results obtained from this study are consistent with those observed experimentally. As expected, these ligands were observed to bind to sites other than the active site of the DEN-2 serine protease, illustrating the non-competitive inhibitory activities of most of the ligands. On the other hand, the calculated K_i values for these ligands were very small, indicating good binding of these ligands to the allosteric

binding site. The higher non-competitive inhibitory activity shown by pinostrobin compared to the other compounds could be accounted for by the H-bonding interaction of its ring (A) hydroxyl group with the backbone carbonyl of Lys74, which is bonded to Asp75 (one of the catalytic triad residues in the protease). This interaction was not observed with the other ligands. It was postulated that this mode of interaction could directly affect the conformation of the catalytic triad region, which would then disrupt the electron transfer process required for substrate binding at the active site, hence the lost of the proteolytic activity.

Docking experiments were performed on rigid ligands, as well as on flexible ligands, using rigid protein structures obtained from the PDB. SAR analysis provided some structural information which may be useful for the design of new compounds with potential inhibitory activities. These features are the rigid structure of flavanone, the presence of the C5 hydroxyl and C7 methoxy groups on ring A, and the phenyl ring (B) in the molecule.

The results obtained in this study, however, did not show major conformational changes that could occur after the ligand-protein binding process. To overcome such limitation, flexible-protein docking would be preferable. The development of computational strategies for this purpose is still in its infancy. Several methods have been reported and promising results have emerged from the application of combined methods such as the ensemble docking approach and an induced fit. Flexible-protein docking would be a way forward towards the deeper insight into the system which can aid in the design of new ligands in our effort to search for a therapeutic lead compound that would inhibit the viral replication activities.