CHAPTER 6

COMPUTATIONAL DOCKING OF COMPETITIVE INHIBITOR TO

DEN-2 NS2B-NS3

6.1   Introduction

Computational docking studies of non-competitive inhibitors to DEN-2 NS2B-NS3 have been discussed in previous chapter (Chapter 5). A study carried out by Tan *et al.* (2005, 2006) reported 4-hydroxypanduratin A (Figure 6.1), isolated from *Boesenbergia rotunda*, to be most active in competitively inhibiting protease activity. Subsequent study involving computational docking of this compound to the homology model of DEN-2 NS2B-NS3 was performed (Lee *et al.*, 2006, 2007). In this study we report the use of the DEN-2 NS2B-NS3 protease crystal structure obtained from the PDB as the target protein for the docking study. In addition, computational calculations were performed using the hybrid quantum mechanics / molecular mechanics (QM/MM) method to understand the important modes of interactions between the inhibitor and the active site. Results obtained from this study are then compared to those reported by Lee *et. al.* (2007). Information generated from observations made may provide an in-depth insight into the mechanism of inhibition of this ligand on DEN-2 protease. This, undoubtedly, invaluable information will be a tremendous help in the design of new ligands for the inhibition of dengue virus replication, leading towards possible antiviral drug.

6.2   Materials and Methods

The protocols in this study are as illustrated in the flowchart in Figure 6.2.
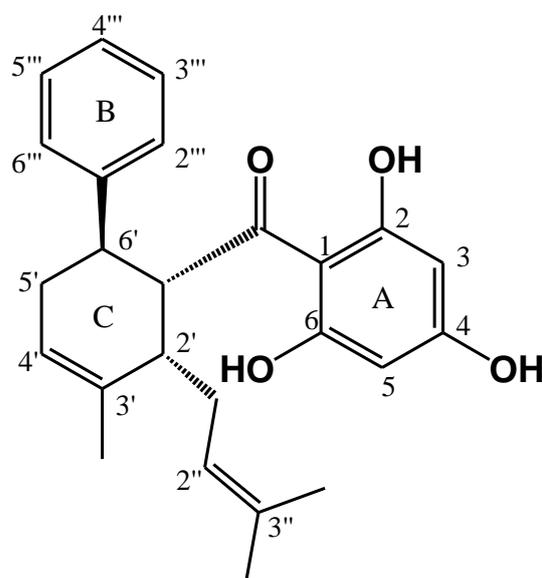
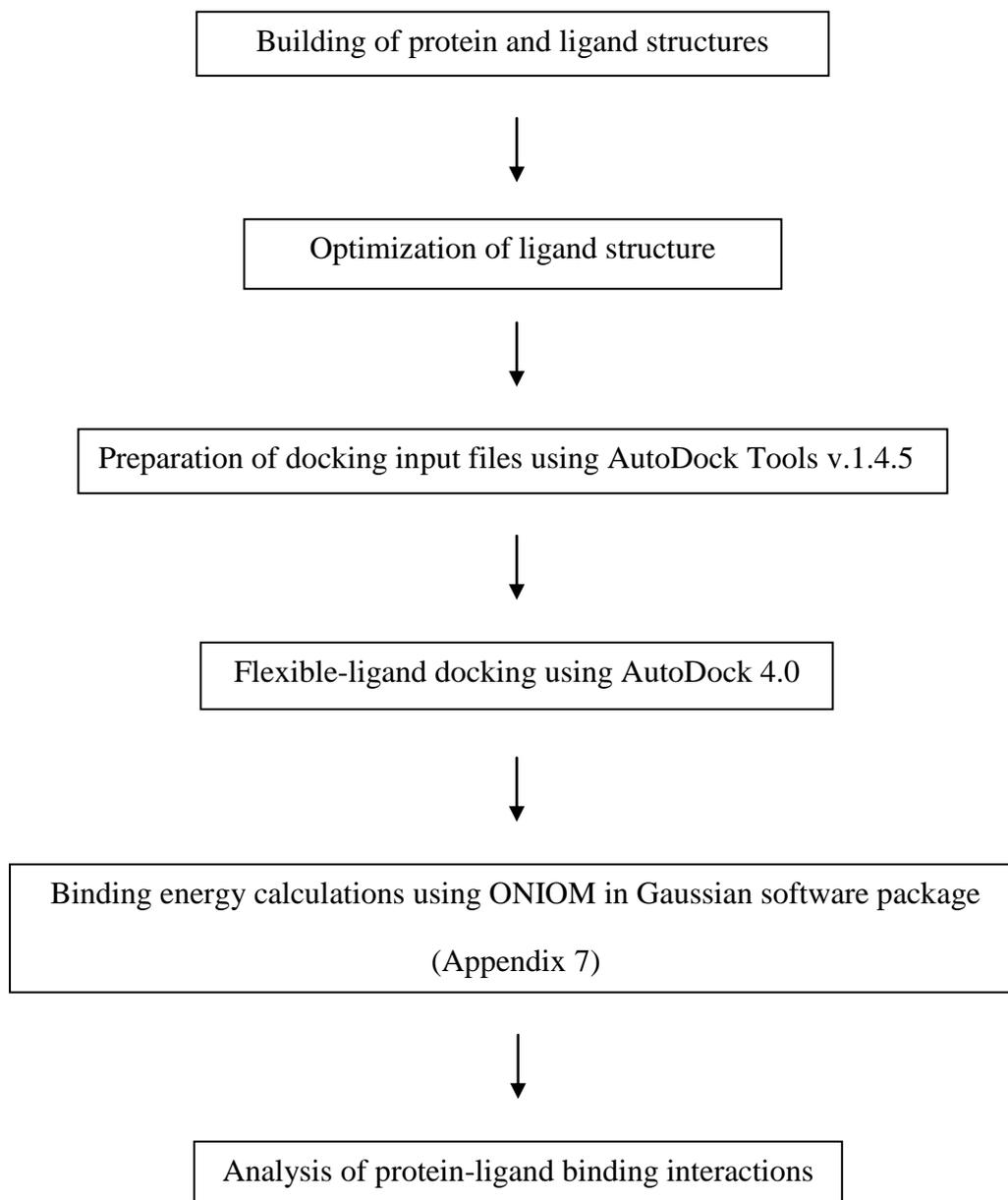Figure 6.1   Structure of 4-hydroxypanduratin A isolated from *Boesenbergia rotunda* (Tan, 2005; Tan *et al*.,2006).

Building of protein and ligand structures

↓

Optimization of ligand structure

↓

Preparation of docking input files using AutoDock Tools v.1.4.5

↓

Flexible-ligand docking using AutoDock 4.0

↓

Binding energy calculations using ONIOM in Gaussian software package

(Appendix 7)

↓

Analysis of protein-ligand binding interactions

Figure 6.2   Flowchart of protocols involved in the study of docking of 4-hydroxypanduratin A to DEN-2 NS2B-NS3.

6.2.1   Materials

6.2.1.1 Hardwares

Docking simulations in this study were performed on an IBM i386 computer, running on SUSE Linux 9.3 operating system, 2.40 GHz Intel® Pentium® 4 processor and 256 MB of RAM. Computational calculations using the ONIOM method were carried out on the CentOS-4.3 i386 operating system, 2.40 GHz Intel® Xeon® 4 processor and 1 GB of RAM. All other computational calculations and visualization were done on a notebook running on Microsoft® Windows® XP operating system, 795 MHz AMD Turion^{TM} 64 Mobile Technology and 480 MB of RAM.

6.2.1.2 Softwares

The ligand structure was built and optimized using the Hyperchem Pro 6.0 software (Hypercube Inc.) prior to docking simulations. Docking files were prepared using the AutoDock Tools v.1.4.5 software (http://www.scripps.edu/~sanner/python/adt) (Coon *et al.*, 2001; Sanner, 1999), an open-source software from MGL tools software package (http://mgltools.scripps.edu/). AutoDock 4.0 software package (Huey *et al.*, 2007; Morris *et al.*, 1998) was used for docking calculations. Viewerlite 4.2 (http://www.accelrys.com) visualization tool was used to display and analyse the structures and interactions. GaussView 3.09 software (Dennington II *et al.*, 2003) was used to prepare input files for QM/MM energy calculations. The ONIOM method (Dapprich *et al.*, 1999; Maseras & Morokuma, 1995; Matsubara *et al.*, 1996; Svensson *et al.*, 1996; Torrent *et al.*, 2002) as available in the Gaussian 03 rev. D.1 program (Frisch *et al.*, 2005) was used to perform QM/MM energy calculations.

6.2.2   Methods

6.2.2.1 Building 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. The ligand structure was built based on the crystal coordinates of panduration A (Tuntiwachwuttikul *et al.*, 1984) using the Crystal Builder module in the Hyperchem Pro 6.0 software (Hypercube Inc.). The 4-methoxy group was replaced with a hydroxyl group to yield the structure of 4-hydroxypanduratin A. The ligand structure was then optimized using Hyperchem Pro 6.0 with MM+ parameters using the steepest descent followed by conjugate gradient (Polak-Ribiere) algorithms (energy minimization termination conditions were set to maximum of 500 cycles or 0.1 kcal / Å mol RMS gradient).

6.2.2.2 Automated protein-ligand docking

Docking input files were prepared using AutoDock Tools v.1.4.5 software (http://www.scripps.edu/~sanner/python/adt) (Coon et al., 2001; Huey & Morris, 2007; Sanner, 1999). For the ligands, all hydrogen atoms were added, Gasteiger charges were computed, non-polar hydrogen atoms were merged and all rotatable bonds were made flexible; the number of torsional degree of freedom (TORSDOF) was set to 5 (corresponding to 5 rotatable bonds of the ligand). The ligand file was saved with the .pdbqt file extension, which is an AutoDock 4 specific file format; 'pdb' augmented by 'q', a charge, and 't', AutoDock atom types. For the protein molecule, all hydrogen atoms were added, Gasteiger charges were computed and non-polar hydrogen atoms were merged. The macromolecule file was saved with the .pdbqt file extension.

A grid parameter file (GPF) was prepared as input file for the AutoGrid 4 utility (Table 6.1). This file specifies the receptor around which the potential energies will be calculated, the types of maps to compute, and the location and extent of the maps. In general, one map is calculated for each atom type in the ligand plus an electrostatics map and a separate desolvation map. Default values were provided with the GPF generated and changes were made where necessary. In this study, the grid box with grid spacing of 0.375 Å and dimension of 60 x 60 x 60 points along the x, y and z axes, was built around the active site (Figure 6.3). During the grid calculation, a '*.glg' log file and the map files were written.

A docking parameter file (DPF) was prepared as input file for the AutoDock programme 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 6.2). The Lamarckian genetic algorithm (LGA) (Morris *et al.*, 1998) was selected as the search algorithm. A population size of 150 and 10,000,000 energy evaluations were used for 100 search runs.

All grid maps and docking calculations were performed using the AutoDock 4.0 software package. After the docking searches were completed, clustering histogram analyses were performed, based on RMSD (root mean square deviation) 0.5 Å. AutoDock 4.0 uses the united atom model to represent molecules, and the scoring function was calibrated using Gasteiger partial charges on both the ligand and the macromolecule. The force field used in AutoDock 4.0 (Huey *et al.*, 2007) is based on a comprehensive thermodynamic model that allows incorporation of intramolecular energies into the predicted free energy of binding. It also incorporates a charge-based

Table 6.1   The grid parameter file (GPF) containing the parameters and maps required by AutoGrid 4 utility in the AutoDock 4.0 software, for calculating grid maps where the ligand is 4-hydroxypanduratin A.

```
npts 60 60 60                          # num.grid points in xyz
gridfld 2FOMCa3.maps.fld               # grid_data_file
spacing 0.375                          # spacing(A)
receptor_types A C HD N NA OA SA       # receptor atom types
ligand_types A C HD OA                 # ligand atom types
receptor 2FOMCa3.pdbqt                 # macromolecule
gridcenter -3.706 -13.535 14.379       # xyz-coordinates or auto
smooth 0.5                             # store minimum energy w/in rad(A)
map 2FOMCa3.A.map                      # atom-specific affinity map
map 2FOMCa3.C.map                      # atom-specific affinity map
map 2FOMCa3.HD.map                     # atom-specific affinity map
map 2FOMCa3.OA.map                     # atom-specific affinity map
elecmap 2FOMCa3.e.map                  # electrostatic potential map
dsolvmap 2FOMCa3.d.map                 # desolvation potential map
dielectric -0.1465                     # <0, AD4 distance-dep.diel;>0,
                                         constant
```
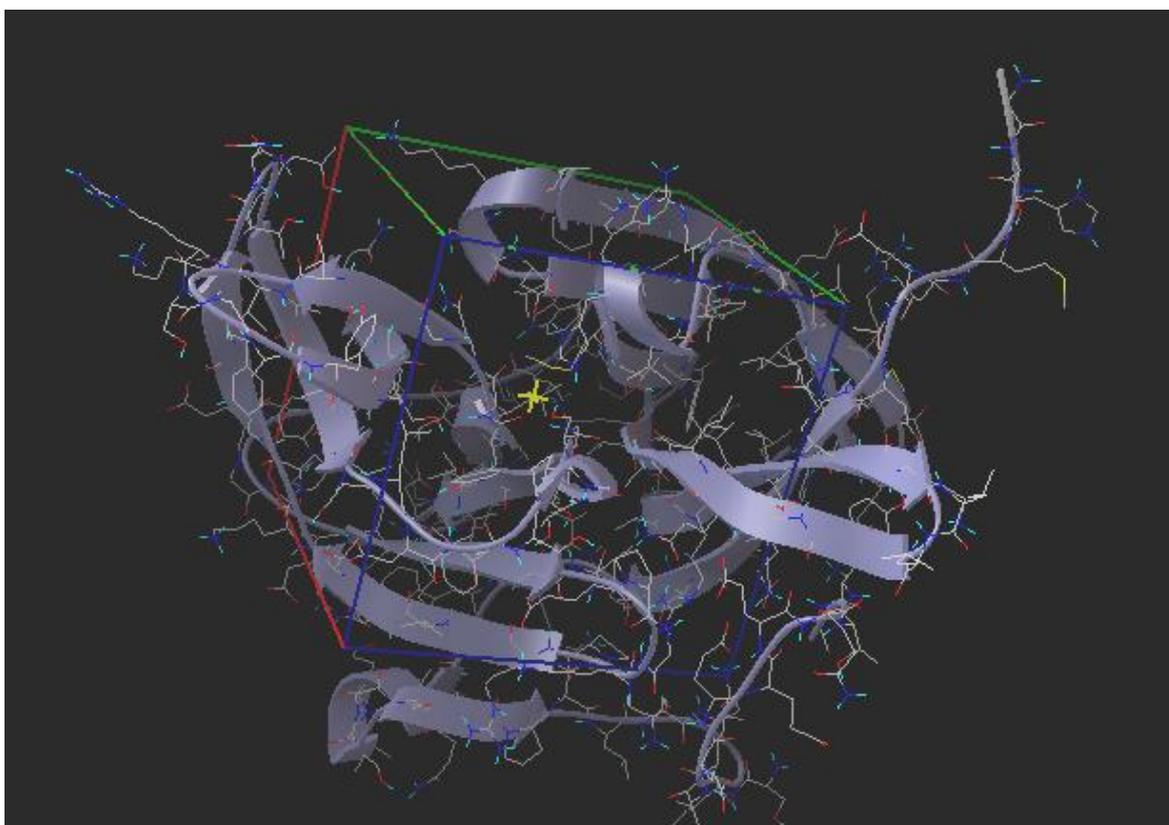
Figure 6.3   Setting up of grid box in the building of the grid parameter files prior to docking. The box surrounds the active site and is the defined grid space for docking. The solid ribbons represent the secondary structure of DEN-2 protease and the lines represent the amino acids, coloured according to the types of atom.

Table 6.2   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
intelec                               # calculate internal electrostatics
seed pid time                         # seeds for random generator
ligand_types A C HD OA                # atoms types in ligand
fld 2FOMCa3.maps.fld                  # grid_data_file
map 2FOMCa3.A.map                     # atom-specific affinity map
map 2FOMCa3.C.map                     # atom-specific affinity map
map 2FOMCa3.HD.map                    # atom-specific affinity map
map 2FOMCa3.OA.map                    # atom-specific affinity map
elecmap 2FOMCA3.e.map                 # electrostatics map
desolvmap 2FOMCA3.d.map               # desolvation map
move 4OHpanduratin.pdbqt              # small molecule
about -0.8461 -8.3824 -1.8898         # small molecule center
tran0 random                          # initial coordinates/A or random
quat0 random                          # initial quaternion
ndihe 8                               # 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 5 0.274000                    # torsional degrees of freedom and
                                         coefficient
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 10000000                 # 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_psw1                              # set the above pseudo-Solis & Wets
                                         parameters
compute_unbound_extended              # compute extended ligand energy
ga_run 100                            # do this many hybrid GA-LS runs
analysis                              # perform a ranked cluster analysis
```

method for evaluation of desolvation designed to use a typical set of atom types. The force field shows improvement in redocking simulations over the AutoDock 3.0.5 force field. The energy function implemented in the AutoDock 4.0 software is described in Appendix 6.

The final docked energy is calculated from the sum of the intermolecular energy and the internal energy of the ligand. However, only the minimum free energy of ligand binding is used for the result analysis.

6.2.2.3 Calculation of binding energies using quantum mechanics / molecular mechanics (QM/MM) methods

The methods involved in this part of the study are similar to those reported by Kuno *et al.* (2003), Nunrium *et al.* (2005) and Saen-oon *et al.* (2005). Two ligand poses from the preceeding automated docking job were selected based on the lowest free energy of binding and the high number of conformations in a cluster.

6.2.2.3.1   System setup

Two structural files (in PDB format) for this part of the study were obtained from the preceeding docking jobs and was opened as Gaussian input files ('.gif') in GaussView 3.09 software (Dennington II *et al.*, 2003). The two systems are named Structure I, where the inhibitor (4-hydroxypanduratin A) formed H-bond with Ser135, and Structure II, where the inhibitor formed H-bond with Asp75. The model systems (ligand-binding pocket complexes) used in the investigation of interactions between the inhibitor and the residues in the active site of DEN-2 protease were well defined. These

systems contained residues within the binding pocket surrounding the inhibitor with at least one atom interacting with any of the atoms of the inhibitor within the interatomic distance of 5 Å (Table 6.3). These residues were selected based on this criterion from Structure I and Structure II (from the docking results), and were mainly from the NS3 domain of the protease complex. Thus two model systems containing 4-hydroxypanduratin A bound into the binding site of 26 residues were generated, where the only differences between the two were the coordinates and geometries of the inhibitors.

The N- and C-terminal ends of the cut residues were capped with acetyl and methyl amino groups, respectively, from the adjacent amino acid residues. This is to retain, as close as possible, the peptide bond characters at both termini of the capped amino acid residues. The geometries and charges of the capping groups were retained as presented in the backbone geometries of the replaced residues (Figure 6.4). Table 6.4 shows the cut residues which were capped either on the N- or C-terminal ends. Hydrogen atoms were then added to the (crystal) structure of the model systems to generate the complete structure of the models. The atoms of the residues in the binding site were assigned with Amber charges. The geometries of the heavy atoms of the residues were fixed and the positions of the hydrogen atoms were optimized using the semi-empirical PM3 method. This structure was used as the starting geometry for the following calculations.

6.2.2.3.2   System optimization

Three stages of optimization were performed to prepare for the calculation of binding energy (BE) that followed.

Table 6.3   Residues in the binding sites of DEN-2 protease complex representing the model systems (in the presence of inhibitor) that are used in the QM/MM energy calculation. These residues have at least one atom interacting with any of the atoms of 4-hydroxypanduratin A within the interatomic distance of 5 Å. (Refer to section 6.2.2.2.1 of methods)

Residues	Shortest distance from an atom in the residue to the inhibitor (Å)
Gln35**	5.00
Ile36**	1.95
Gly37**	4.13
Met49*	4.60
Trp50*	3.70
His51*	2.36
Val52**	3.35
Val72*	4.57
Asp75*	1.97
Leu128*	2.45
Asp129# *	5.65
Phe130**	3.11
Ser131**	3.45
Pro132*	1.89
Gly133**	2.87
Thr134**	4.53
Ser135*	4.14
Tyr150*	3.10
Gly151*	1.85
Asn152*	2.35
Gly153*	1.75
Val154*	2.93
Val155*	3.33
Tyr161*	1.66
Val162*	3.92
Ser163*	4.79

\# Asp129 is added to the model system to conserve the connection between the amino acids within the chain.

* Based on Structure I of the docking results (ligand formed H-bond with Ser135).

** Based on Structure II of the docking results (ligand formed H-bond with Asp75).

137

Figure 6.4   Capping groups of the N- and C-terminal ends for the cut amino acid residues.

Table 6.4 Residues in the model systems which are capped to "cut" the bonds with other amino acids in the chain.

Residues with capped N-terminal ends	Residues with capped C-terminal ends
Gln35	Gly37
Met49	Val52
Val72	Val72
Leu128	Ser135
Tyr150	Val155
Tyr161	Ser163

(i) *Optimization of the ligand-pocket complexes.* From the starting geometries of the complexes (refer section 6.2.2.3.1), the coordinates of all amino acid residues were fixed, except those of Asp75, Ser135 and the inhibitor, which were optimized using PM3 method.

(ii) *Optimization of the binding pocket.* From the starting geometries of the complexes (refer section 6.2.2.3.1), the atoms of the inhibitor were deleted, leaving only those of the binding pocket. The coordinates of Asp75 and Ser135 were optimized using PM3 method, while the coordinates of the rest of the amino acid residues were fixed.

(iii)*Optimization of the ligand.* From the starting geometries of the complexes (refer section 6.2.2.3.1), the amino acid residues of the binding pocket were deleted, leaving only those of the inhibitor. The geometry of the inhibitor was optimized using PM3 method.

An example of the optimization job command lines is shown in Table 6.5. For all of the above optimization jobs, the Merz-Kollman-Singh (MK) (Singh & Kollman, 1984) approximate charges were applied during the geometry optimization microiterations.

6.2.2.3.3   ONIOM2 calculations

Two-layered ONIOM (ONIOM2) method was used to investigate the interactions between 4-hydroxypanduratin A and the amino acid residues in the binding site. Kuno *et al.* (2003) and Saen-oon *et al.* (2005) reported the use of ONIOM2 and ONIOM3 (three-layered ONIOM) in their studies. However, in the present study, only

Table 6.5   An example of the input file containing the command lines for the Gaussian optimization job performed using ONIOM method.

```
%chk=11_4OHpandu2_oniom2opt2.chk
%mem=64MW
%nproc=2
# opt=small oniom(pm3:pm3) geom=connectivity pop=mk

optimize Asp75HAF + inhibitor; ONIOM2; 11_4OHpandu2_oniom2opt2.com

-2 1 -1 1 -1 1
 C                 0    0.334533     5.518862     2.888590 H
 C                 0   -0.180700     4.623269     1.938032 H
 C                 0   -0.813554     3.435211     2.343636 H
 C                 0   -0.946494     3.139463     3.717105 H
 C                 0   -0.406165     4.024947     4.671262 H
 :
 :
 :
```

- The line beginning with # is the *route section*, which specifies the kind of job to run as well as the specific theoretical method and basis set which should be used.
- The line following the blank line after the route section is the *title section* for the job, which provides a description of the calculation for the job output and archive entry; it is not used by the program.
- The section that comes up after the blank line following the title section makes up the *molecule specification section*. The first line of this section gives the charge and spin multiplicity for the molecule/system. As for ONIOM2 calculation, the format for this input line is:
  charge/spin (entire system, low level); charge/spin (inner layer-high); charge/spin (inner layer-low)
  The remaining lines specify the element type, the command to freeze (-1) or unfreeze (0) element, the Cartesian coordinates (Å) and the symbol for the level of theory used in the ONIOM calculation (H, high; L,low).

(Foresman & Frisch, 1996; Frisch *et al.*, 2005)

ONIOM2 method was applied mainly due to the restricted computational resources available in our laboratory. Furthermore, Kuno *et al.* (2003) reported the use of both methods on their model system and, despite using different ONIOM methods, they found the intermolecular distances and the binding energies from both calculations to be quite similar (Nunrium *et al.*, 2005).

For the ONIOM2 calculations, the present model systems were partitioned into two layers, region **A** (inner) and region **B** (outer). The inner layer, or the interaction region (Figure 6.5, region **A**), was treated with high level calculations using the HF/6-31G(d,p) and B3LYP/6-31G(d,p) methods. The outer layer, or the environmental region (Figure 6.5, region **B**), was treated with low level calculations using the PM3 method. For Ser135 (inner layer), its bond with Thr134 (outer layer) was 'cut' to allow for the partitioning into the inner and outer layers. The resulting 'dangling' bonds were saturated with H atoms (link atoms) (refer Figure 6.6).

The total ONIOM energy (E^{ONIOM2}) of each model system was obtained from three independent energy calculations (Dapprich *et al.*, 1999):

$$E^{ONIOM2} [AB] = E[High,\mathbf{A}] + E[Low,\mathbf{AB}] - E[Low,\mathbf{A}] \qquad (1)$$

where $E[High,\mathbf{A}]$ is the energy of the inner layer calculated using high level of calculations, $E[Low,\mathbf{AB}]$ is the energy of the whole model system calculated using low level of calculations, and $E[Low,\mathbf{A}]$ is the energy of the inner layer calculated using low level of calculations.
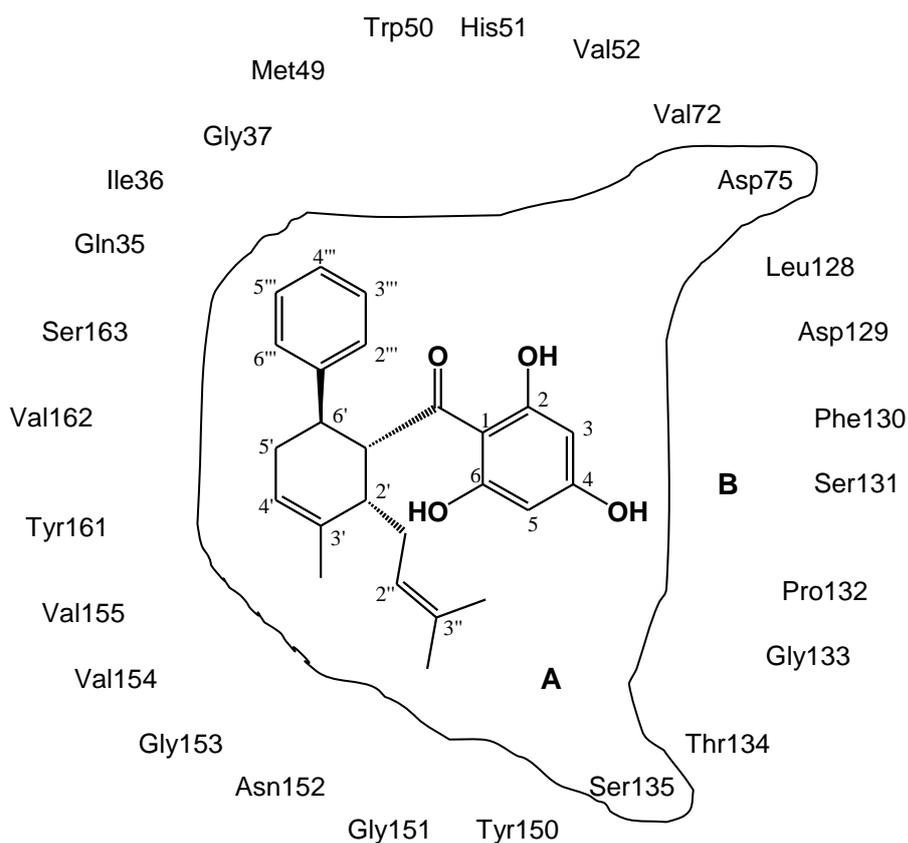
Figure 6.5   Schematic 2D diagram of the model system for 4-hydroxypanduratin A bound to DEN-2 protease binding site consisting of 26 amino acid residues. Layer partitioning is shown for ONIOM2 calculations. **A** is the inner layer (high level calculations) and **B** is the outer layer (low level of calculations). The arrangement of the residues shown is not representative of their actual positions in 3D space.
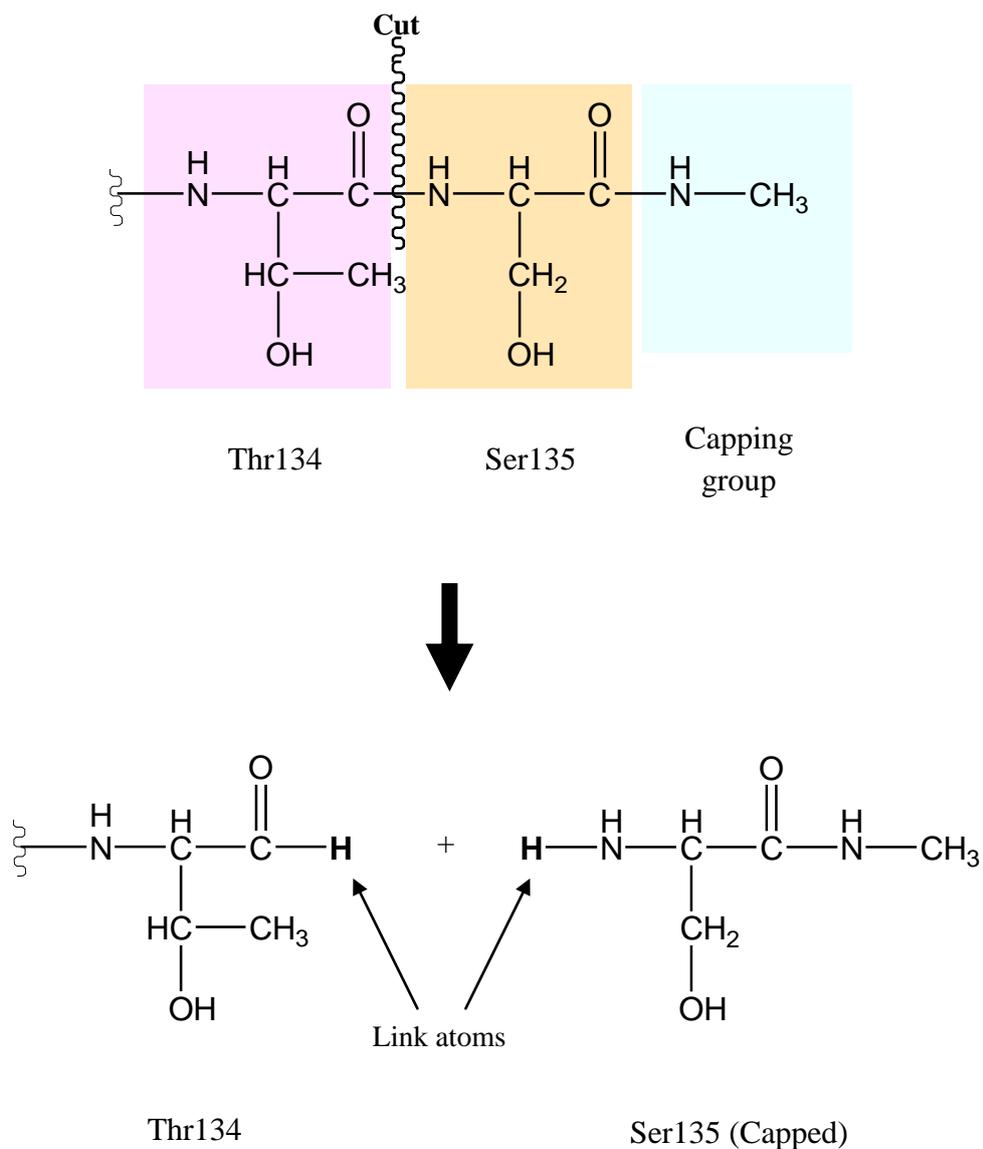
Figure 6.6   Schematic illustration of the cutting of peptide bond between Thr134 (outer layer) and Ser135 (inner layer) for energy calculations using ONIOM2 method.

For binding energy (BE) calculation, the following equation was applied (Saen-oon *et al*., 2005):

$$BE = E[Cpx]_{opt} - E[P]_{opt} - E[L]_{opt} \qquad (2)$$

where $E[Cpx]_{opt}$ is the energy of the optimized ligand-pocket complex, $E[P]_{opt}$ is the energy of the optimized binding pocket, and $E[L]_{opt}$ is the energy of the optimized ligand. The derivation of equation (2) is given in Appendix 7.

6.2.2.4 Analysis of results

Details of the applications used in the analysis of results obtained from the docking experiments were described previously in section 5.2.2.4 of Chapter 5.

6.3   Results

6.3.1   Automated protein-ligand docking

The ligand studied, 4-hydroxypanduratin A, was optimized prior to docking to DEN-2 NS2B-NS3 and the calculated single point energies are shown in Table 6.6. The minimized energy shown in Table 6.6 is for the structure in which the optimization process has reached the convergence criteria (see materials and methods).

The clustering histogram extracted from the docking output log file ('*.dlg') is appended as Appendix 8. The top two clusters have very similar minimum free energy of binding (ΔG_{bind}) values with a difference of only 0.01 kcal / mol. Hence, these two docked conformations of the ligand, from rank 1 (run #20) and rank 2 (run #11), were selected for further investigation. Table 6.7 showed the energy and constant values of

Table 6.6   Single-point energy values of 4-hydroxypanduratin A calculated using AM1 method (semi empirical) available in Hyperchem Pro 6.0 software (Hypercube Inc.).

Before minimization		After minimization		Δ Energy (kcal/mol)
Energy (kcal/mol)	Gradient (kcal/mol Å)	Energy (kcal/mol)	Gradient (kcal/mol Å)	
-6037.92	18.87	-6045.56	23.46	-7.64

Table 6.7 Results of the top two conformers generated from the automated flexible-ligand docking of 4-hydroxypanduratin A to DEN-2 NS2B-NS3, calculated using AutoDock 4.0.

	Run #20	Run #11
Experimental IC_{50} (µg/ml)	40.00	40.00
Experimental inhibition constant, K_i (µM)	21	21
Min. estimated free energy of binding, ΔG_{bind} (kcal/mol)	-7.47	-7.46
Estimated inhibition constant, K_i (μM)	3.37	3.43
Final intermolecular energy, ΔG_{inter} (kcal/mol)	-7.54	-7.36
Final total internal energy, ΔG_{intra} (kcal/mol)	-1.41	-1.59
Torsional free energy, ΔG_{tor} (kcal/mol)	1.37	1.37

The IC_{50} and K_i values were obtained from experiments performed by Tan (2005).

the two conformers obtained from the '*.dlg* file. Figure 6.7a illustrated the superimposition of the two conformations which indicated the two structures to be very similar and that they bind to the active site of the protein. Figures 6.7b and c showed the H-bonds formed between the conformers and the catalytic triad. The figures clearly illustrate that conformer from run #20 forms H-bonds with Ser135, and conformer from run #11 with Asp75. For further clarification on the probability of which residue of the catalytic triad, either Asp75 or Ser135, would actually have H-bond interaction with the ligand, QM/MM calculations using the ONIOM method were performed.

6.3.2   Calculation of binding energies using QM/MM methods

From this point onwards, the system in which the ligand formed H-bond with Ser135 is termed Structure I, and the system in which the ligand formed H-bond with Asp75 is termed Structure II (refer materials and methods, section 6.2.2.3.1). The results for the energies of the optimized components of the system under study are as tabulated in Table 6.8. The binding energies are summarized in Table 6.9. Results obtained from this study showed the binding of 4-hydroxypanduratin A to the active site of DEN-2 protease to be highly mediated by the initial binding to Ser135 of the catalytic triad with calculated binding energy of about -26 kcal/mol (Table 6.9).

6.4   Discussions

6.4.1   Automated protein-ligand docking

Before 4-hydroxypanduratin A was subjected to docking experiment, it was initially optimized under the conditions as described in methods (section 6.2.2.1). This optimization was carried out to enable the ligand structure to be 'relaxed' enough before
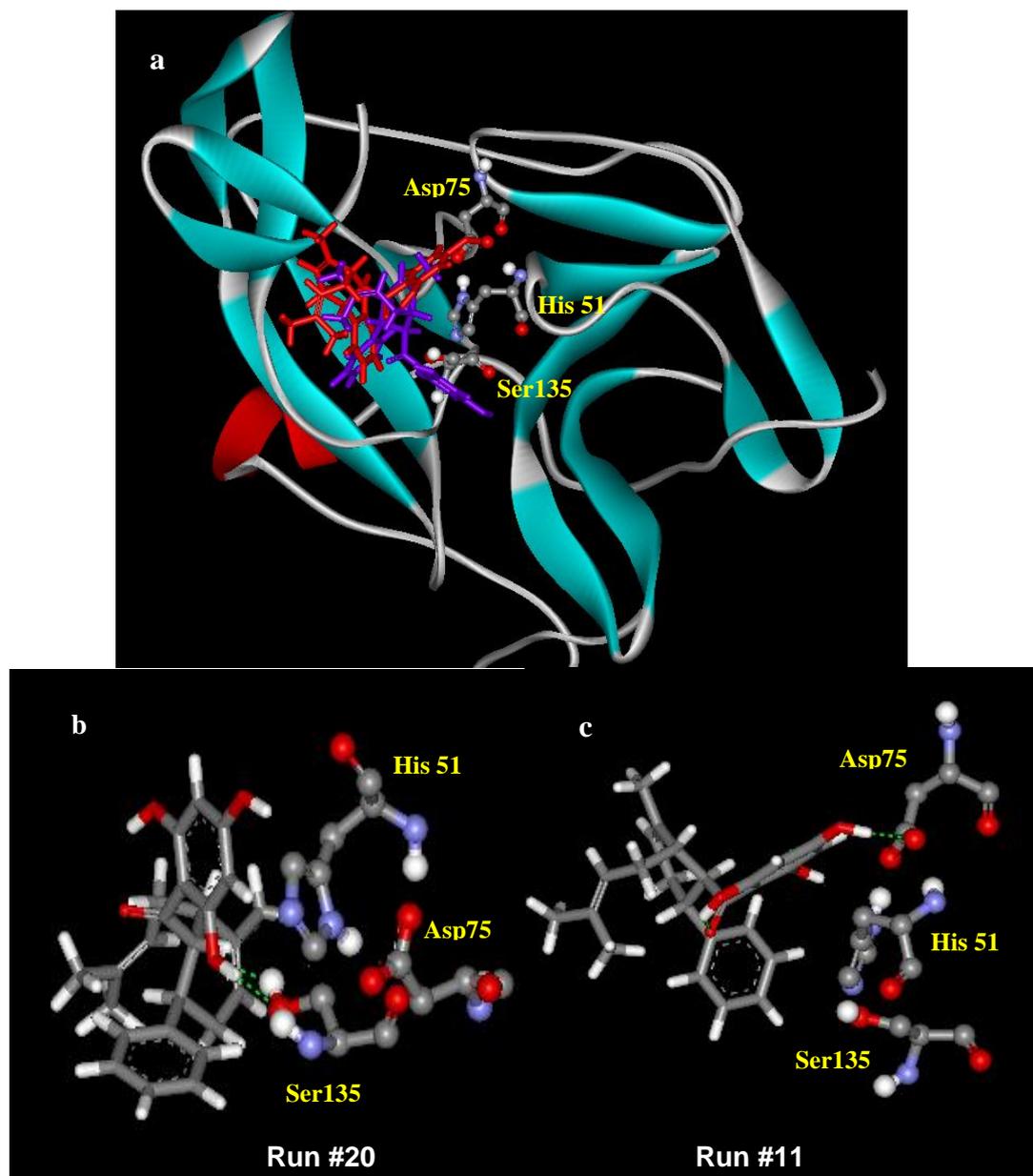
Figure 6.7   Computer generated models illustrating: (a) The superimposition of the selected conformers, from the docking of 4-hydroxypanduratin A into DEN-2 NS2B-NS3, at the binding site. The ligands are shown as sticks: run #20 (violet) and run #11 (red). DEN-2 protease is represented as ribbons. The catalytic triads labelled as His51, Asp75 and Ser135 are shown as balls and sticks. (b) and (c) The H-bonding interactions between the conformers and the catalytic triad. The H-bonds are indicated by green dashed lines.

Table 6.8   Calculated energies of the optimized systems (refer section 6.2.2.3.2) using ONIOM2 method, for (a) Structure I, where the ligand binds to Ser135, and (b) Structure II, where the ligand binds to Asp75, in the binding site via H-bonds.

(a)

Method	Energy (Hartree)			BE	
	$E[Cpx]_{opt}$	$E[P]_{opt}$	$E[L]_{opt}$	(Ha)	(kcal/mol)
HF/6-31G(d,p):PM3	-2020.3968	-757.9170		**-0.0407**	**-25.5464**
B3LYP/6-31G(d,p):PM3	-2032.6987	-762.2573		**-0.0427**	**-26.8252**
HF/6-31G(d,p)			-1262.4391		
B3LYP/6-31G(d,p)			-1270.3987		

(b)

Method	Energy (Hartree)			BE	
	$E[Cpx]_{opt}$	$E[P]_{opt}$	$E[L]_{opt}$	(Ha)	(kcal/mol)
HF/6-31G(d,p):PM3	-2020.6229	-758.1392		**-0.0243**	**-15.2813**
B3LYP/6-31G(d,p):PM3	-2032.9296	-762.4796		**-0.0303**	**-19.0416**
HF/6-31G(d,p)			-1262.4593		
B3LYP/6-31G(d,p)			-1270.4196		

Binding energy, BE = $E[Cpx]_{opt}$ - $E[P]_{opt}$ - $E[L]_{opt}$

where, $E[Cpx]_{opt}$ = Energy of the optimized ligand-pocket complex
      $E[P]_{opt}$ = Energy of the optimized binding pocket
      $E[L]_{opt}$ = Energy of the optimized ligand

Table 6.9   Binding energies of ligand to the binding site of DEN-2 protease, with regards to the two systems investigated in this study.

Method	Binding Energy (kcal/mol)	
	Structure I	Structure II
HF/6-31G(d,p):PM3	-25.55	-15.28
B3LYP/6-31G(d,p):PM3	-26.83	-19.04

Structure I: The ligand conformation was obtained from run #20 of the docking output file.

Structure II: The ligand conformation was obtained from run #11 of the docking output file.

it was subjected to conformational changes during the docking process. It is presumed that starting with a ligand structure with relatively low single point energy (Table 6.6) would be more representative of the most likely structure of the compound in nature.

Figure 6.7a showed the superimposition of the two selected docked conformers. These two structures were observed to be very similar and are able to bind to the active site of the protein. Both the conformers carried negative internal energy (ΔG_{intra}) values, which could reflect their favourable conformations (see Table 6.7). In addition, the estimated K_i values observed were reasonably small (within the μM range) indicating the formation of stable enzyme-inhibitor complexes.

The high similarities in energy and K_i values for the two conformers, as shown in Table 6.7, inferred the possibility of two modes of binding (Figures 6.7b and c). These modes of binding refer to the H-bond interactions between the ligand with Ser135 and Asp75, respectively. As discussed earlier (section 2.2), the serine residue in the serine protease catalytic triad has been reported to bind to substrate, forming a tetrahedral oxyanion intermediate (Wilmouth *et al.*, 2001). It would then be expected, from this study, for the inhibitor to illustrate similar binding mode. Automated docking results, however, showed that besides Ser135, another possible binding interaction with the catalytic site could be mediated through H-bond with Asp75. Theoretically, this would be possible if the inhibitor was bound to Asp75 before the charge-transfer relay network (Bachovchin, 1985) can take place within the catalytic triad (as elaborated in section 2.2). Furthermore, Asp75 carries the carboxylic acid group as its side chain, rendering it to be more susceptible to H-bonding with any H-donor group. To ascertain which one of the two residues in the catalytic triad would form the H-bond interaction with the inhibitor, a higher level of computational calculations was carried out.

6.4.2   Calculation of binding energies using QM/MM methods

Several QM/MM approaches were performed in this part of the study. Some examples of the attempts were: (i) to include the whole protease in the QM/MM calculations by applying both the ONIOM2 and ONIOM3 methods; (ii) considering the effect from the surrounding residues onto the ligand binding interactions by limiting the selected residues to those listed in the methods section (6.2.2.3.1) and performed ONIOM3; and (iii) using a few other combinations of basis sets. However, only the results from the successful approach are presented in this thesis. Most unsuccessful attempts, however, were due to the very limited computational resource available in our lab during the course of this study.

Results from the ONIOM2 calculations, as shown in Tables 6.8 and 6.9, yielded in binding energy values which are lower for Structure I compared to Structure II. This leads to the conclusion that the binding of 4-hydroxypanduratin A to the active site of DEN-2 protease could be highly mediated by its binding to Ser135 of the catalytic triad rather than to Asp75.

6.4.3   Protein-ligand binding interactions

Figure 6.8 showed the binding of 4-hydroxypanduratin A to the active site of DEN-2 protease, where the active site is observed to be considerably exposed to the surface of the macromolecule. The ring structures of the ligand are shown to interact with the surrounding residues of the binding site. However, the olefinic side chain of the cyclohexenyl ring (ring C; Figure 6.1) hanged outside the binding pocket. This seemed to suggest that the side chain is not involved in forming interactions with the
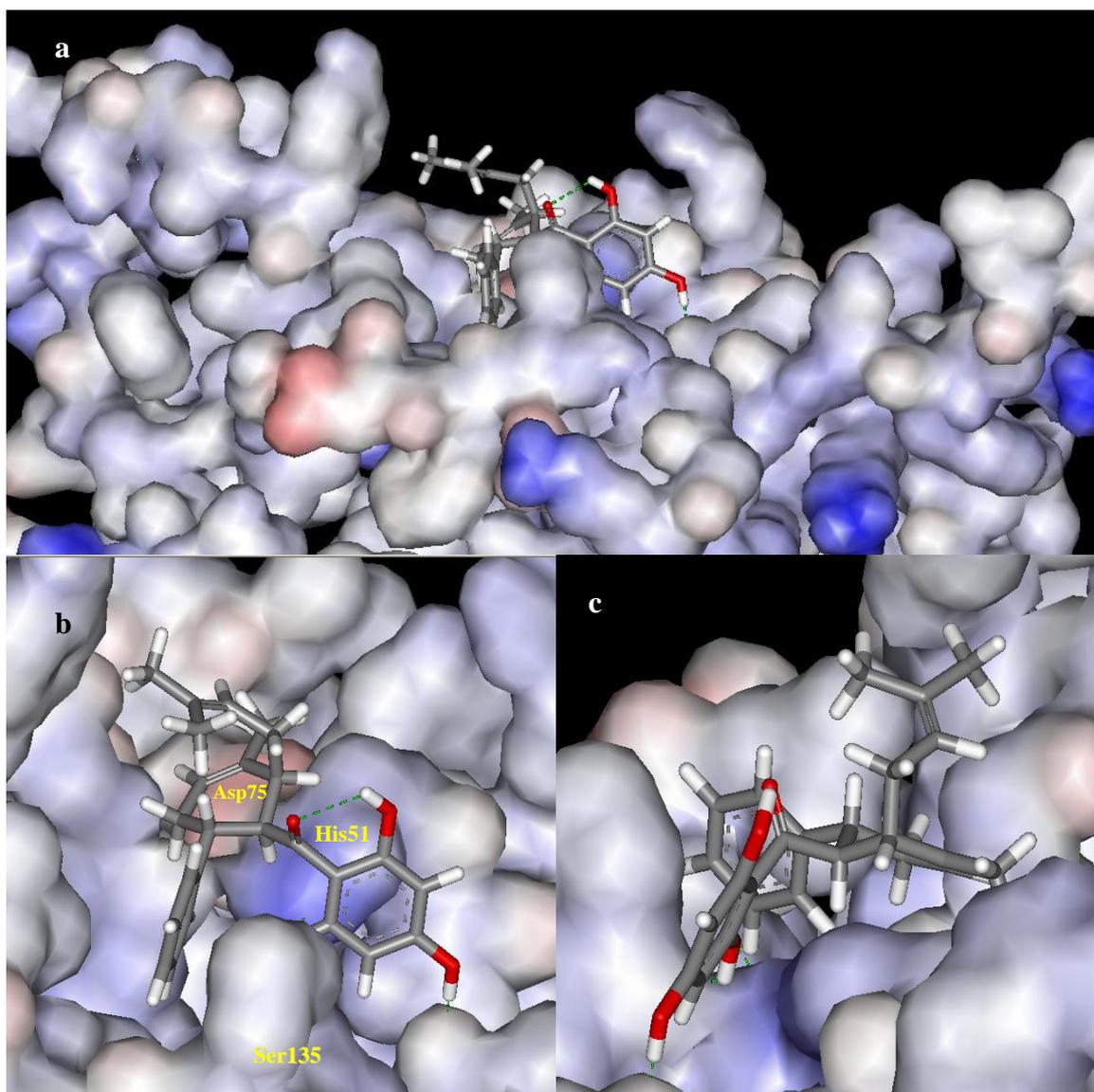
Figure 6.8   (a) Connolly surface representation of the binding site accommodating the ligand, 4-hydroxypanduratin A, shown as stick. (b) and (c) Views from different angles of the ligand in the binding site. The green dashed lines indicate H-bonds.

binding site. Perhaps, its bulky nature blocks the ligand from being embedded further into the pocket.

Figure 6.9a highlighted the position of the ligand with respect to the catalytic triad. The residues of the binding pockets that have various modes of interactions with the ligand are shown in Figures 6.9b - c. Table 6.10 listed all the residues of the binding pockets that interact with the ligand, compared to the results obtained from the docking work performed by Lee *et al.* (2007). In the study, they reported the docking of 4-hydroxypanduratin A on the homology model of DEN-2 protease (Lee *et al*., 2006). The common residues interacting with the ligand, obtained from this study and the study by Lee *et al.* (2007) are: Ser135 for H-bonding; His 51 and Pro132 for van der Waals interaction; and, Leu128 and Pro132 for hydrophobic interaction. The differences observed in the other residues involved in the interactions can be due to the difference in the protein fold of the crystallized protease and the homology model (refer Figure 6.10). Both studies, however, revealed the importance of the hydroxyl group on ring A since it can form H-bond interactions with the residues of the active site.

Erbel *et al.* (2006) reported the structural basis for the activation of the dengue and West Nile viruses NS3 protease where they highlighted the residues that make up the S1 pocket (for illustration of S1 pocket, refer to Figure 2.4) for inhibitor binding. These residues were shown in Figure 6.11. It can be observed from this figure that 4-hydroxypanduratin A blocked the entrance into the S1 pocket upon interaction with Ser135 (Figures 6.9a and b). This blockade may be enough to disable the entrance of substrate into the pocket. However, the limited number of H-bonding interactions between the inhibitor and the binding site (Table 6.10) may affect the equilibrium of the
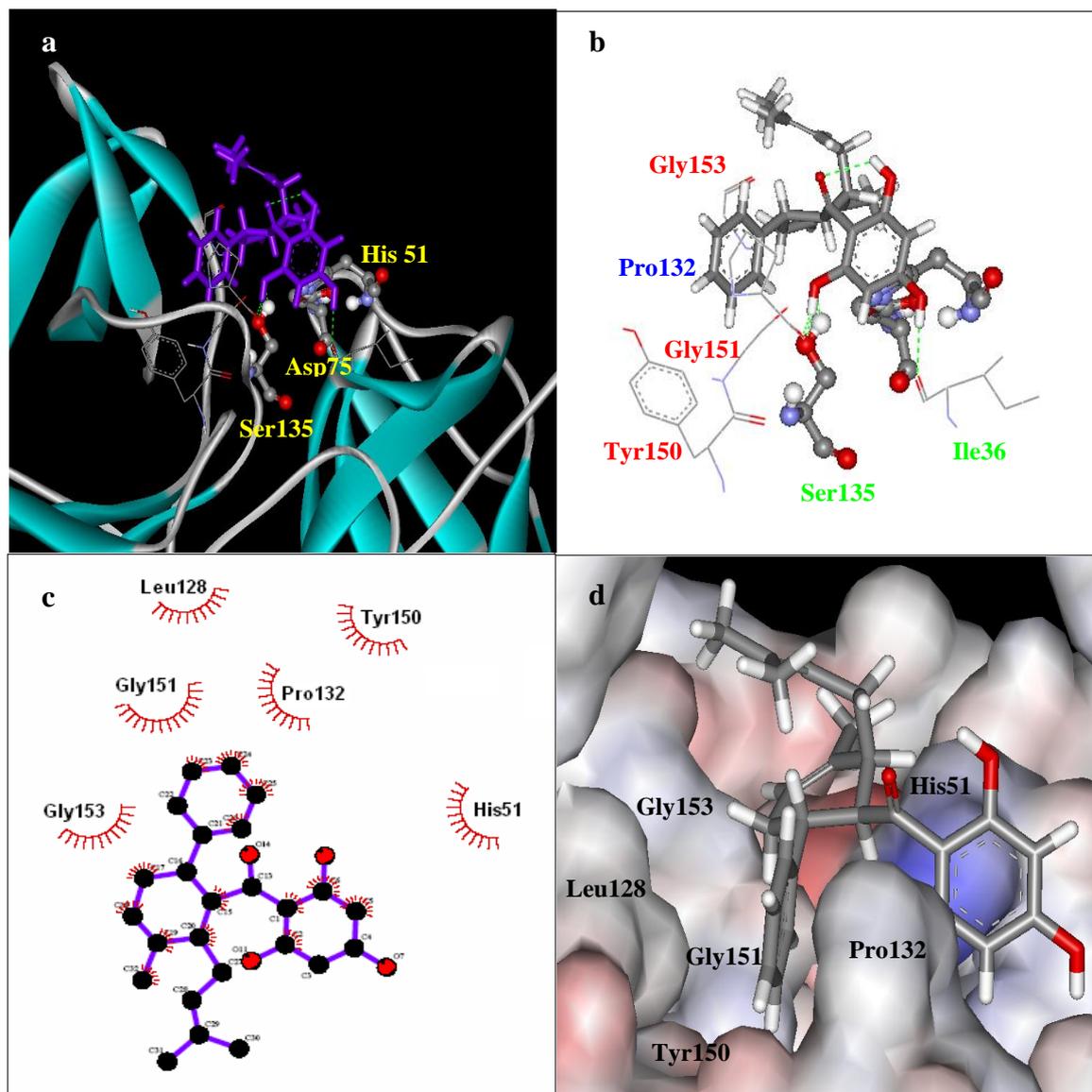
Figure 6.9   (a) View of 4-hydroxypanduratin A (violet) at the binding site of DEN-2 protease (ribbons). Residues interacting with the ligand are shown as lines. Catalytic triads are shown as balls and sticks, and are labelled. The green dashed lines indicate H-bonds. (b) Simplified view of 4-hydroxypanduratin A interacting with surrounding residues. Residues labelled in green interact with the ligand via H-bonds, those in red interact via van der Waals contacts, and 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 4-hydroxypanduratin A, 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 the residues (labelled in black) which exhibit hydrophobic interactions with the ligand.

156

Table 6.10   Residues in DEN-2 protease active site which are involved in the interactions with the docked conformer of 4-hydroxypanduratin A. Tabulated are the results obtained from this study, which used DEN-2 protease crystal structure, and those obtained from the study by (Lee *et al.*, 2006), which used DEN-2 protease homology model.

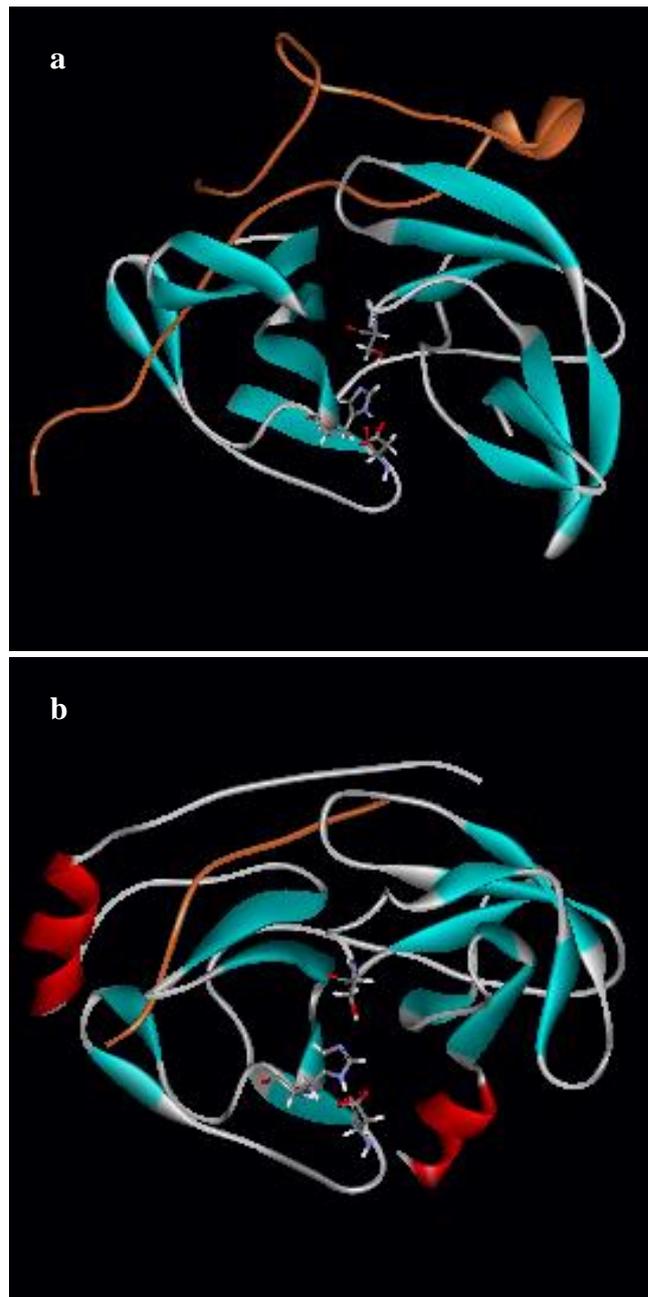|	Residues	
Mode of interactions	Crystal structure	Homology model
Hydrogen bonding	Ile36 Ser135 Pro132	Asp75 Ser135 Gly151
van der Waals	His51 Pro132 Gly151 Gly153	His51 Ser131 Pro132 Gly133 Tyr150 Asn152
Hydrophobic	His51 Leu128 Pro132 Tyr150 Gly151 Gly153	Val52 Leu128 Pro132 Val155

Figure 6.10   Diagrams to illustrate the difference in protein folding of DEN-2 protease, as obtained from (a) crystallized protease (D'Arcy *et al*., 2006; Erbel *et al*., 2006), and (b) homology modelling technique (Lee *et al.*, 2006). The catalytic triads are shown as sticks. For the NS3 protease domain, the regions in cyan are β-strands, regions in red are α-helices, and regions in grey are loops or coils. The strand in orange represents the NS2B domain.
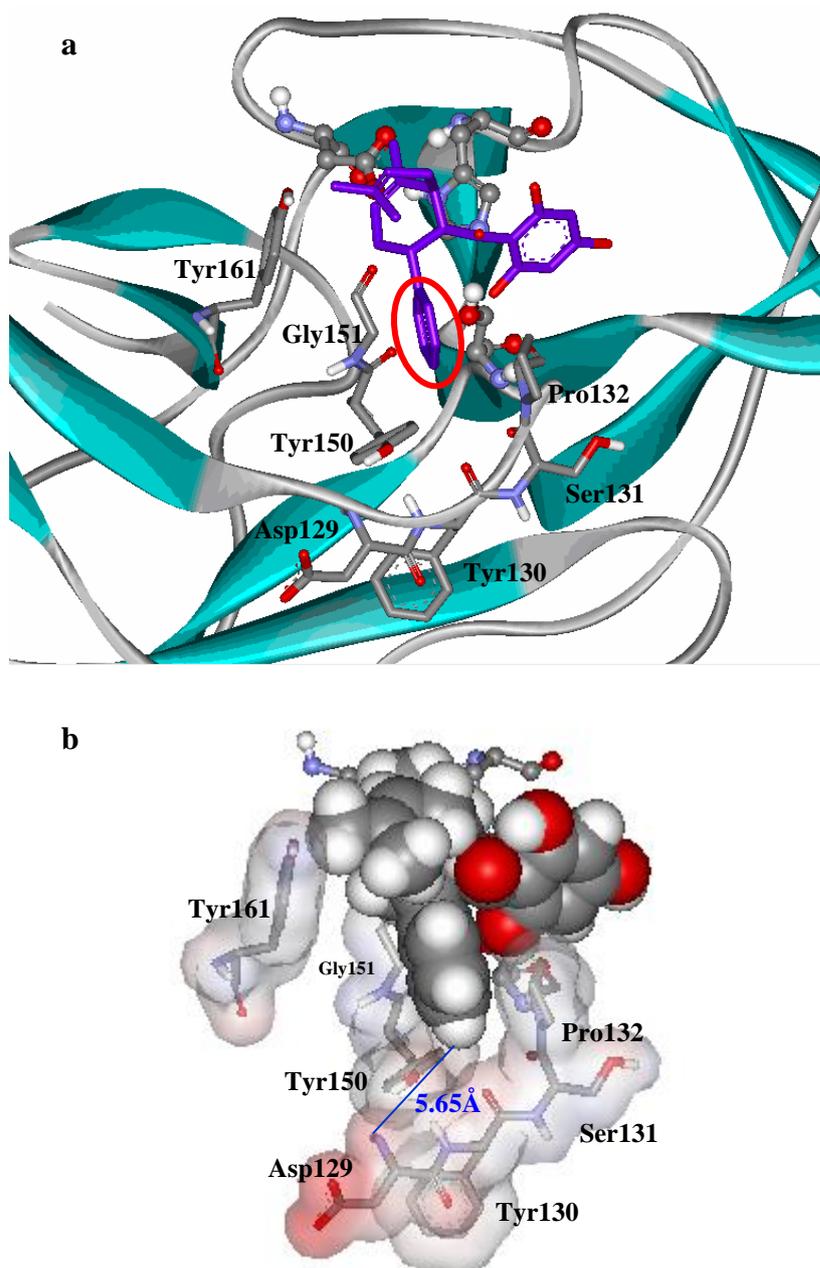
Figure 6.11   4-hydroxypanduratin A bound to the active site of DEN-2 protease. (a) The inhibitor is violet in colour, with the O atoms coded in red colour and the H atoms are not shown to improve visualization. Residues comprising the catalytic triad are shown as balls and sticks. Residues in sticks and labelled are the residues that make up the S1 pocket (Erbel *et al.,* 2006). The red circle highlights the position of ring B of the inhibitor (refer to text). (b) The inhibitor is represented in CPK format with colour coding for the type of element. The residues in the S1 pocket are shown as sticks embedded in transparent Connolly surfaces. The blue line indicates the shortest distance between an atom of the inhibitor and that of Asp129, which is located at the bottom of the S1 pocket.

unbound and bound enzyme-inhibitor complex (E + I ↔ E-I) in a manner where the

unbound could be formed more. Hence, the substrate can have a good chance to

compete with the inhibitor for the active site. The experimental K_i value of 21 μM (Tan

et al., 2006) is considered to be a good indication for stability of the enzyme-inhibitor

complex. However, it would be more desirable if an inhibitor with K_i value within the

nano molar range (nM) could be designed. For this, 4-hydroxypanduratin A provides a

good template for the design of a more potent inhibitor of DEN-2 protease.

Erbel *et al.* (2006) also reported that Asp129 is located at the bottom of the S1

pocket and stabilizes the positively charged side chain of the residue at P1 position of

the substrate. From Figure 6.11b, the shortest distance between an atom of the inhibitor

(H-4''' on ring B) and that of Asp129 (N atom on the amino group of the backbone) is

observed at 5.65Å, a distance which is considered to be far for a significant van der

Waals interaction to take place. Thus, to enhance the interaction, a side chain can be

added to the C-4''' position of the phenyl ring (ring B). It is suggested that this side

chain carries a functional group which can protrude into the groove of the S1 pocket

and create interactions with Asp129 specifically, and other surrounding residues

generally. These interactions should preferably be electrostatic interactions such as

ionic or H-bonding (for specific interactions with Asp129 which is acidic).

Furthermore, it can be seen from Figure 6.11 that ring B is orientated in such a way that

it faced Pro132, and is not directed towards the centre of the S1 pocket. This could

presumably be caused by the hydrophobic interactions between the ring B and Pro132.

In addition, the orientation may also be due to the rigidity that is imposed by the

intramolecular H-bond formed between the hydroxyl group on C-2 of the phenolic ring

(ring A; Figure 6.1) and the adjacent carbonyl group. Hence, it is suggested that the

carbonyl group can be reduced to an alkyl or substituted with other functional groups

which do not form intramolecular H-bond with the adjacent phenolic ring. This may allow some flexibility to enable the phenyl ring to orientate itself towards the centre of the S1 pocket. In addition, by removing the olefinic side chain on the cyclohexenyl ring (ring C), some degree of flexibility could be enhanced due to the reduction of steric hindrance by the side chain.

Lee *et al.* (2007) reported that a π-π type of aromatic interaction between the phenolic ring (ring A) and the pentacyclodiazo side-chain of His-51 was observed. In the present study, however, a similar interaction was not observed between the two moieties, as the orientation of the phenyl ring did not allow this interaction to take place.

6.5   Conclusion

In this study, 4-hydroxypanduratin A, a competitive inhibitor isolated from *Boesenbergia rotunda* L., was docked onto the active site of DEN-2 NS2B-NS3 protease using the AutoDock 4.0 software. Further energy calculations using ONIOM2 method on the Gaussian03 software confirmed the hypothesis of inhibitor-protein interaction which is mediated by the interaction of the inhibitor with Ser135 of the catalytic triad in the active site. This is in line with the mechanism of substrate binding to the active site of serine proteases, as reported by Garrett & Grisham (1997) and Wilmout et al. (2001).

Comparison of the results obtained from this study with those obtained by Lee *et al.* (2007) highlighted several differences in the binding interactions. One outstanding difference is the π-π type of aromatic interaction between the phenolic ring (ring A) and

the pentacyclodiazo side-chain of His-51 reported by Lee *et al.* (2007), which was not observed in this study. On the other hand, both studies revealed that the phenolic ring (ring A) to be an important pharmacophore for binding to the active site.

Some suggestions for the modification of certain structural features of 4-hydroxypanduratin A were made, with the aim to increase the binding efficacy. This may further reduce the K_i value of the inhibitor, which will lead to a more stable enzyme-inhibitor complex. Firstly, it is suggested that the olefinic side-chain of the cyclohexenyl ring be removed, as it did not contribute to any binding interaction with the binding site. Furthermore, its removal could reduce steric hindrance, and enhance the possibility of the ligand to enter the active pocket. In addition, this may enable the phenyl ring (ring B) to orientate itself into the centre of the S1 pocket of the binding site. Secondly, it is suggested that a side chain be added to the C-4''' position of the phenyl ring (ring B); preferably one which carries a functional group which can protrude into the groove of the S1 pocket creating interactions with the surrounding residues. Thirdly, it is also suggested that, in order to reduce the rigidity of the inhibitor and enable the phenyl ring (ring B) to be orientated towards the centre of the S1 pocket, the carbonyl group in a designed inhibitor be reduced to an alkyl or substituted with other functional groups which do not form intramolecular H-bond with the adjacent phenolic ring.

The structural modifications suggested above are important information for future studies in the attempt to design candidate drugs with potential as anti-viral agents against dengue virus infections.