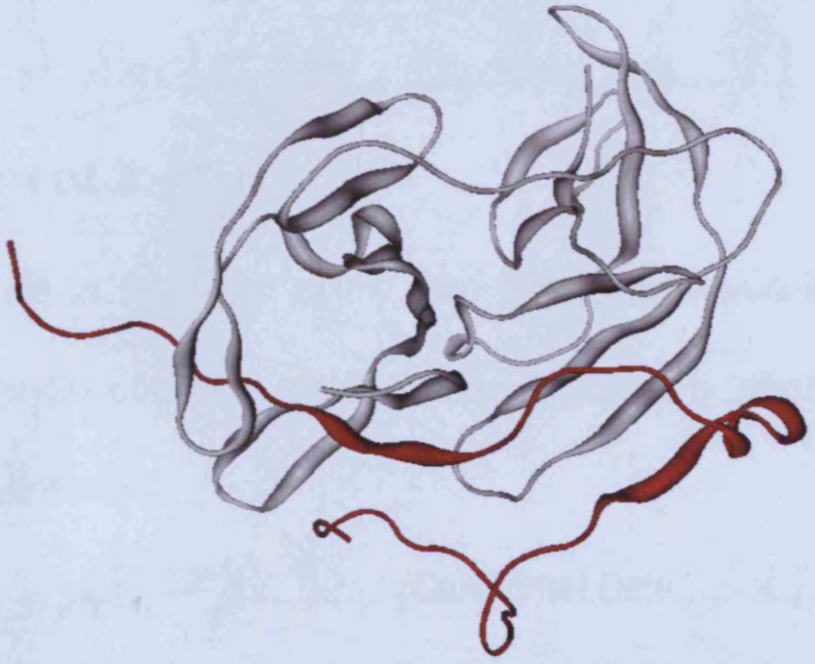
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Computer-Aided Drug Design And Synthesis of Novel Antivirals.



A thesis submitted to Cardiff University for the degree of doctor in philosophy in medicinal chemistry.

Mohammed Abdou Khedr

Welsh School of pharmacy, Cardiff University

October 2010

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed....M. Abdoel..... (Candidate) Date....18/10/2010

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed...M. Abdoel..... (Candidate) Date....18/10/2010

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed....M. Abdoel..... (Candidate) Date.....18/10/2010

Statement 3

I hereby give consent for my thesis, if accepted to be available for photocopying and for interlibrary loan, and for the title and summary to be made available to outside organisations.

Signed....M. Abdoel..... (Candidate) Date.....19/10/2010

Abstract

The Flaviviridae is a family of 66 viruses of which almost half have been associated with human disease. The most well-known members are: Hepatitis C virus (HCV), Dengue virus (DV), and West Nile virus (WNV). Diseases caused by these viruses are a global health problem that put an estimated 2.5 billion people at risk. At present, there are neither vaccines nor other treatments available to prevent or cure these diseases. Potential targets for the development of therapeutics against the virus are the viral protease and polymerase. The aims of this project are to design and synthesize compounds that can be used as inhibitors for these two key enzymes for Dengue. Structure-based drug design methods utilize knowledge of a three dimensional structure of an enzyme/receptor to develop small molecules able to bind to the desired target, generating a specific biological response. These computer-based methodologies are now becoming an integral part of the drug discovery process and, although the principles of molecular recognition are far from being completely understood, some marketed compounds (i.e. Zanamivir, Lopinavir) have been developed with the help the of successful application of structure-based design techniques. Different structure-based drug design approaches have been used to identify putative new inhibitors for the Dengue protease and polymerase. A pharmacophore query has been built based on the active site of the Dengue protease enzyme and then used for screening different databases for identification of potential inhibitors. For the polymerase, a fragment-based approach has been used to find the fragments that would interact more efficiently with a specific binding pocket on the enzyme. The virtual library obtained by linking the best scored fragment was then docked to identify the most promising structures to be synthesized. The identification of potent small molecules that bind to receptors and enzymes is one of the major goals of chemical and biological research.

Dedicated to my parents, my wife, and my lovely kids; Ammar and Jumana.

Acknowledgement

First of all I would like to express my sincere thanks to my supervisor **Dr. Andrea Brancale** to whom I am indebted for accepting me to work in his group and for his constant invaluable guidance and encouragement throughout my PhD.

I would like to acknowledge the **Egyptian Ministry of higher education** especially the **Mission department** for granting me this scholarship. I am grateful to the **Egyptian embassy** and the **Cultural Bureau in London** for their continuous help and support.

I would like to thank **Prof. Chris McGuigan** for his valuable advices in my first year viva. And **Dr. Andrew Westwell** for his important discussions in our group meeting.

My thanks to all the group members and working staff in Welsh school of pharmacy.

Finally, I would like to thank my family wholeheartedly especially my wife and my lovely kids; Ammar and Jumana.

List of Contents

Chapter 2: Molecular modeling in drug design.

Chapter 3: Docking and modification of Panduratin A and 4-hydroxy panduratin A. toward the inhibition of DV.

Chapter 4:Design of Dengue Virus NS3 Serine Protease inhibitors

List of Abbreviations.

Abbreviation	meaning
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
BVDV	Bovine viral diarrhea virus
BZ-NLe	Benzoylated L-nor Leucine
BR	Boesenbergia Rodunta
CS	Complementary sequence
CPE	Cytopathic effect
CSFV	Classical swine fever virus
DV	Dengue Virus
DHF	Dengue hemorrhagic fever
ER	Endoplasmic reticulum
FF	Force fields
Gly	Glycine
Gln	Glutamine
Glu	Glutamic acid
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
His	Histidine
HM	Homology modeling
JEV	Japanese Encephalitis virus
Lys	Lysine
Leu	Leucine
MbBBI	Mung bean Bowman-birk inhibitor
MD	Molecular dynamics
MM	Molecular Mechanics
MMFF	Merck Molecular Force Fields
MFS	Multifragment search
NS	Non-structural

Abbreviation	meaning
NMR	Nuclear magnetic resonance
NTP	Nucleoside triphosphate
Pn	Non prime side
Pn"	Prime side
Phe	Phenylalanine
Pro	Proline
prM	Precursor membrane protein
Pdb	Protein data bank
QSAR	Quantitative structure-activity relationship
RdRp	RNA dependant RNA polymerase
RMSD	Roat of mean square deviation
RF	Replicative form
Ser	Serine
+ssRNA	Positive single-stranded sense RNA
SF	Scoring functions
SLA	Stem Loop A
SFs	Super families
Tyr	Tyrosine
UTR	Untranslated region
UAR	Upstream of AUG region
Val	Valine
VS	Virtual screening
WNV	West Nile virus
YFV .	Yellow fever virus

Chapter 1

An Introduction to the Dengue Virus

1- Introduction.

The dengue virus (DV) is one of 53 viruses which are members of ***Flaviviridae*** family. It is a mosquito borne virus transmitted by the ***Aedes aegypti*** mosquito.[1] Classification according to the analysis of NS3 Helicase regions has been done and Flaviviridae viruses were found to be in three genera, [2] in which DV is grouped within the flavivirus genus.

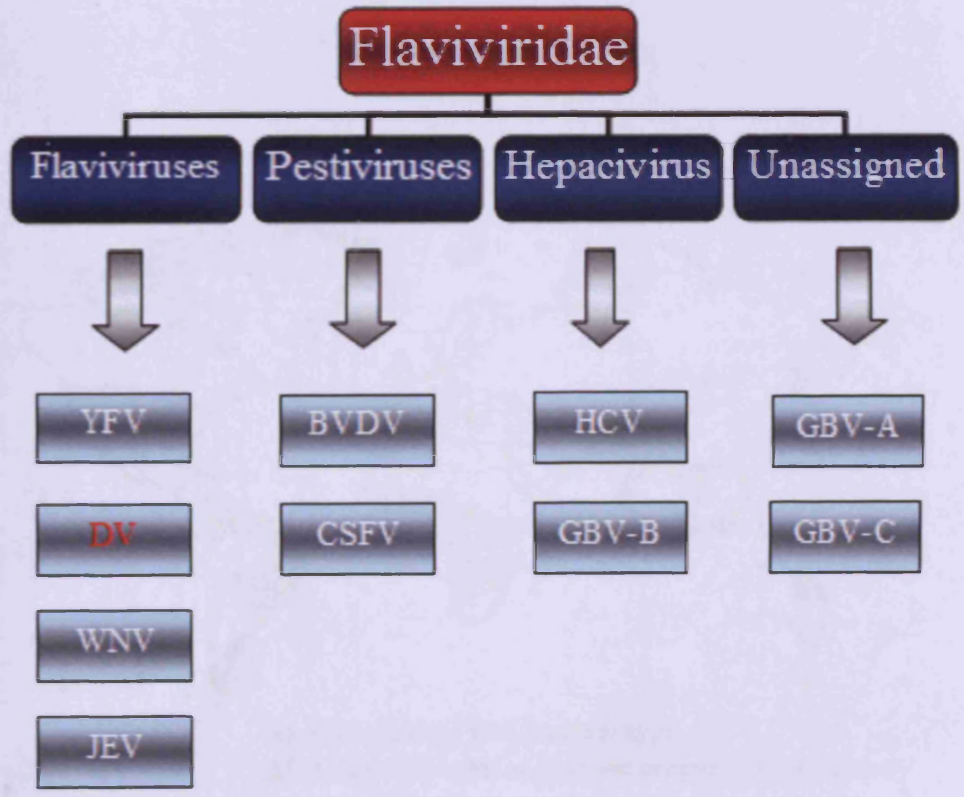


Figure 1.1 Classification of Flaviviridae family according to the analysis of NS3 Helicase. YFV: Yellow Fever Virus. DV: Dengue Virus. WNV: West Nile Virus. JEV: Japanese Encephalitis Virus. BVDV: Bovine Viral Diarrhea Virus. CSFV: Classical Swine Fever Virus. HCV: Hepatitis C Virus. GBV (A, B, C): Unassigned Viruses.

DV is a global health problem which started after the epidemics that occurred in the 18th, and the 19th centuries.[3] The problem with DV is its hemorrhagic fever (DHF) which has been the main cause of death of many children in different countries such as Manila, Philippines, Australia, Malaysia, Pakistan, Canada, Mexico beside other countries in south Africa. Another problem is the fact that it is highly distributed on a wide geographic area.[4-5] In 1997, dengue viruses and the ***Aedes aegypti*** mosquito had a worldwide distribution in the tropics (Figure 1.2), and over 2.5 billion people now live in areas where dengue is endemic.[6-7]

Figure 1.2 The world wide distribution of the dengue virus.[8]

More than 50 million people are infected with DV annually. From this number about 24000 people die every year and until now there is neither vaccination nor any chemotherapeutic agents that can inhibit the replication of DV.[9] An understanding of the viral replication cycle can help us in the identification of the important targets that can be used for the design of novel inhibitors of DV replication. In this chapter we will report a survey of the DV genome, and replication cycle in detail to aid in the identification of the important targets that can be used as a starting point for the design of inhibitors.

1.1- Dengue Virus Serotypes.

DV has four serotypes DV-1, DV-2, DV-3, and DV-4.[10] The complete genomic sequences for these serotypes was identified in 1990.[11-14]

1.2- Clinical manifestations and Symptoms.

Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF)/Dengue shock syndrome are the most important diseases that are caused by DV.[15-16] Fatal DHF is caused by DV-2 and DV-3.[17] The incubation period is from 2 to 14 days,[18-19] and the main symptoms are high fever, headache, lumbosacral pain, facial flushing, and conjunctival congestion. In severe cases of DF myalgia, anorexia, nausea, and vomiting with general weakness can be caused. In addition to DF the dengue virus can cause DHF which is considered to be a sever form of Dengue infections and known as Vascular Leak Syndrome.[20] The dengue virus may be the cause of other diseases such as hepatitis with severe hemorrhage, liver failure, cardiomyopathy, and encephalopathy.[21-22] Some patients with severe cases may develop a gastrointestinal hemorrhage and

shock may result due to plasma leakage.[20] Death from dengue is mainly due to DHF and shock is the second cause after DHF.

1.3- The *Aedes Aegypti* Life Cycle

A female mosquito becomes infected with dengue when biting an infected human who is viraemic (has enough dengue particles in his blood). Dengue can spread quickly and an infected person can transmit the virus to mosquitoes within 3-4 days of being bitten and can continue to do so for up to 12 days. It is important to know that the dengue virus is not spread directly from person to person. The ***Aedes Aegypti*** female mosquito bites only humans and animals for blood which is needed to mature its eggs. The eggs are laid separately to allow them to spread over large surfaces of water, if conditions are good. The eggs will have a better chance of survival. The freshly laid eggs are white in colour and soon turn black and convert to larvae. The young larvae feed on bacteria in water and grow rapidly. After a few weeks in the summer it enters the pupa Stage which is a short stage. They rapidly rise to the surface of water where the adult emerges. The female adult has the virus in its salivary glands.[23] (Figure 1.3). It is important to note that eggs can survive for long periods in a dry state (more than a year), the virus can be passed from adult mosquito to the eggs and in this case the virus is guaranteed survival until the next summer. The virus remains in the salivary glands of the mosquito and when it bites humans it injects saliva plus the virus into the wound where the anticoagulants contained in its saliva facilitate feeding. During the life cycle of the mosquito there is no intermediate host and it seems to be a closed system between man and mosquito

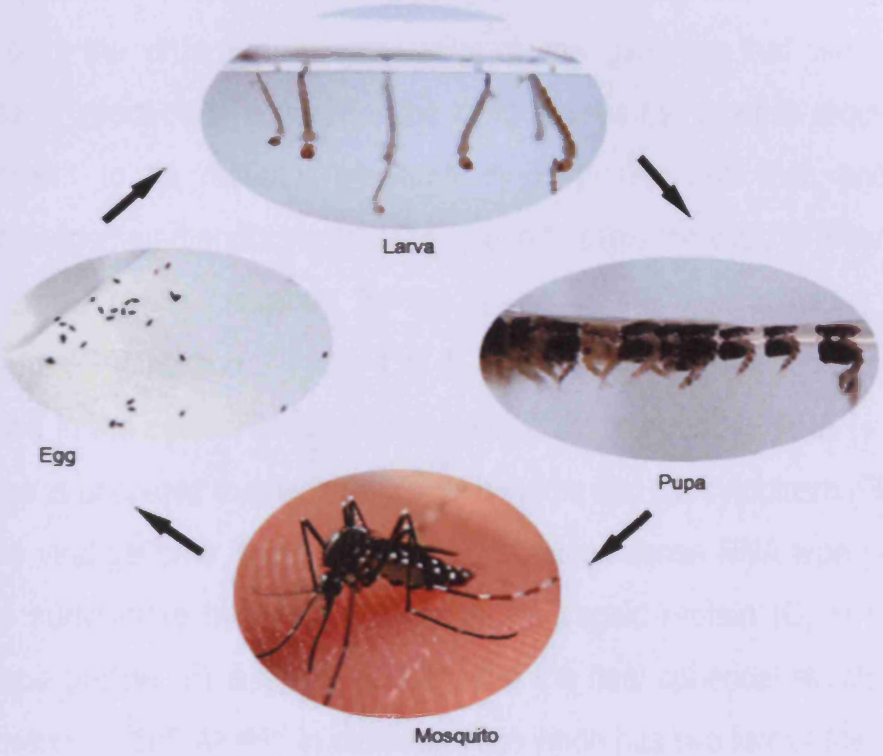


Figure 1.3 Life cycle of the ***Aedes Aegypti*** mosquito.[23]

1.4- Viral Replication Cycle.

Viral replication is the process by which the virus reproduces to form new viruses inside the host cells. The main aim of this process is to maintain the survival of the virus and the generation of new genomes that can form new viruses to infect other hosts. For the dengue virus the process requires virus attachment to its receptor mediated by E protein and then endocytosis (receptor-mediated endocytosis) takes place to allow the virus to enter the cell. Here acidic pH is important for the fusion of the viral envelope and the endosomal membrane (Figure 1.4: **I**). After that the nucleocapsid (cap) is released in the cytosol and the viral positive single-stranded RNA (+ ss RNA) genome is uncoated to release the viral genome into the cytoplasm (Figure 1.6: **II**). The viral genome is a positive single-stranded sense RNA type (+ss RNA) and is surrounded by a lipid bilayer of the capsid protein (C) in which the envelope protein (E) is embedded to form the final spherical structure of the virion which is 500 A^{o} [25-26] in diameter. The virion has two forms; the immature form in which the E protein is complexed with precursor membrane protein (prM) in the endoplasmic reticulum (ER) and upon moving to the trans-Golgi prM-E complex will be cleaved by the aid of cellular protease and furin enzymes to give the mature form of the virion in which a membrane protein M is complexed with E.[27] The cleavage of prM occurs at pH = 6 (Figure 1.5).[28] The M protein was found to have ion channels in its C-terminal that allow the permeability of important ions such as sodium, potassium, chloride and calcium, [29] while the E protein is important in the binding and fusion of the virus to the host cell.[30-31]

The 5' Untranslated region of the genome (5' UTR) will then be directed to the ribosome (Figure 1.4: **III**) in which the viral genome will be translated directly to produce a polyprotein (Figure 1.4: **IV**).[33] The polyprotein is processed by viral and host proteases (Figure 1.6) to produce structural (Figure 1.4: **V**) and non-structural proteins (Figure 1.4: **VI**).[32] The viral RNA-dependant RNA-polymerase (RdRp) uses the +RNA genome to produce a complementary – ss RNA strand that is used as a template for the production of new strands of + ss RNA genome (Figure 1.4: **VII**).[34] Then the Helicase enzyme will unwind the newly formed strands to separate between them and liberate the + ss RNA genome (Figure 1.4: **VIII**). After replication the nucleocapsid will be formed around the viral genome and be directed to endoplasmic reticulum where it is surrounded by the lipid envelope (Figure 1.4: **IX**). Then the new completed viruses are released (Figure 1.4: **X**). Figure1.6 illustrates the viral replication cycle.

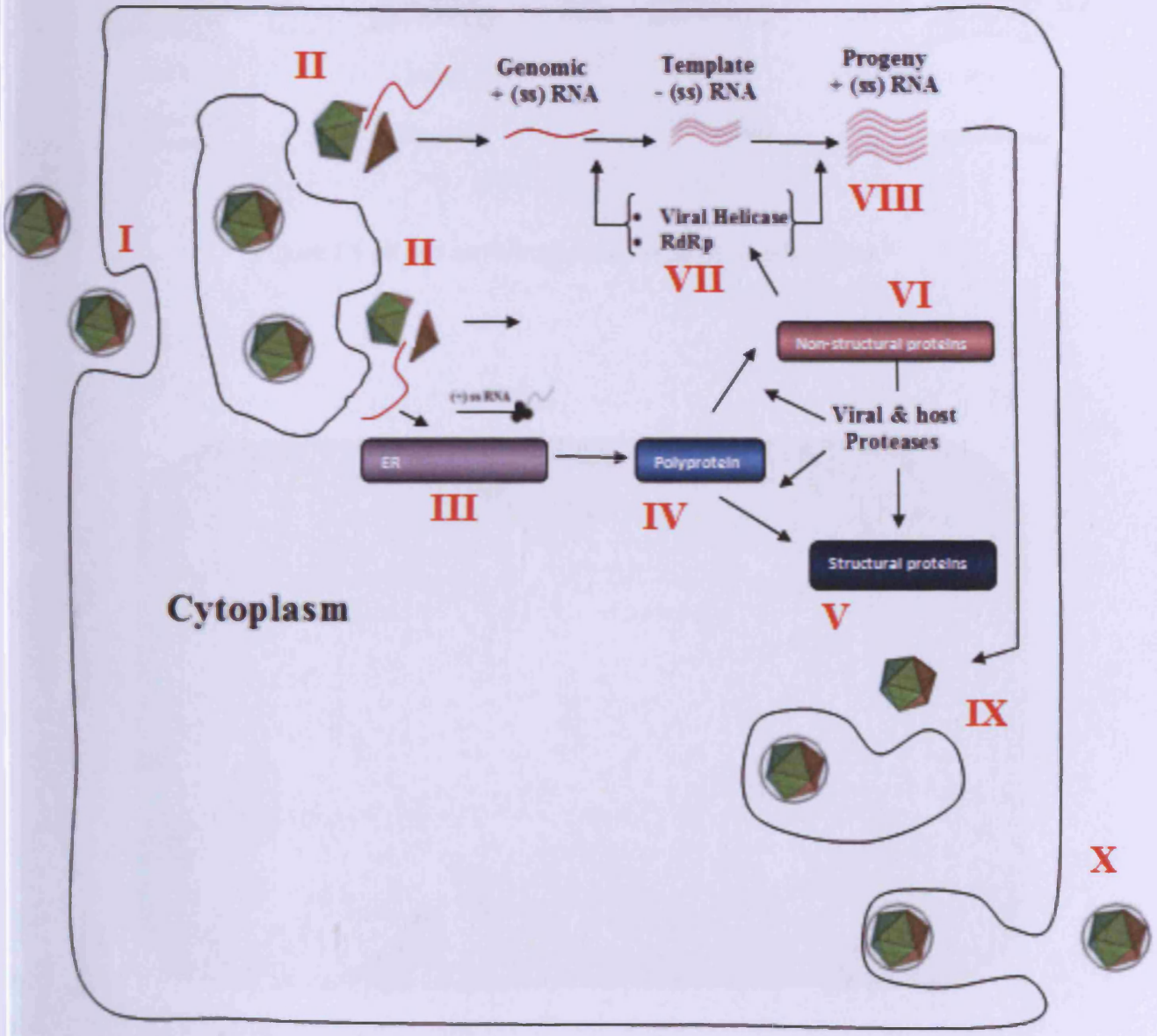


Figure 1.4 Replicative life cycle of the dengue virus genome.

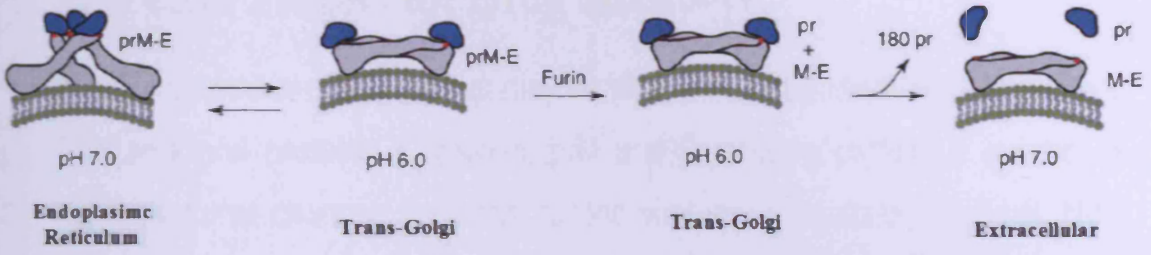


Figure 1.5 pH and conformational changes of the prM protein.[30]

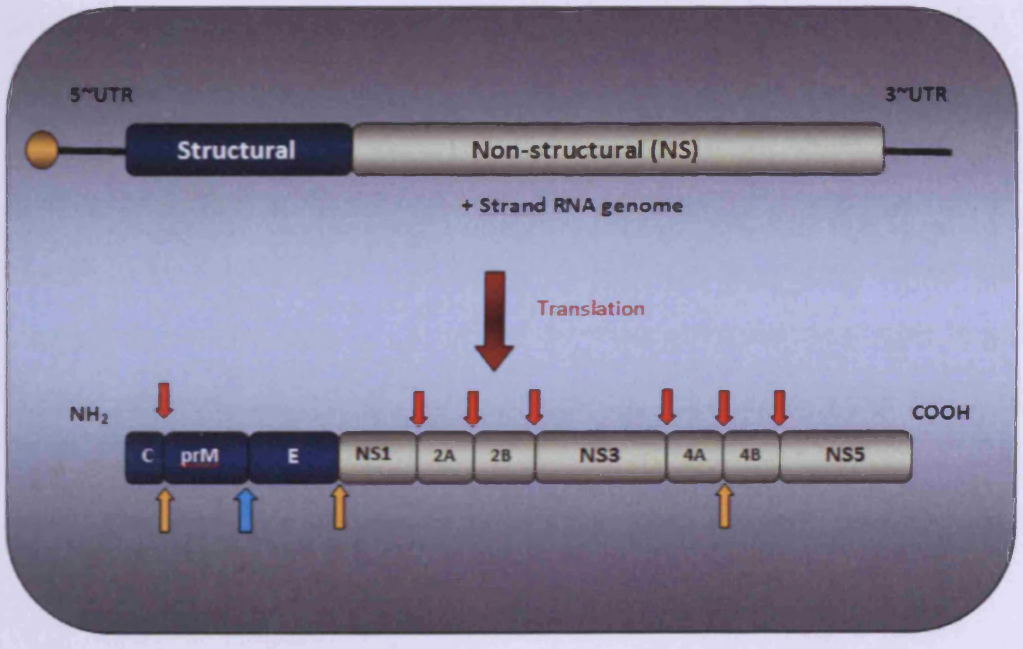


Figure 1.6 The processing of dengue polyprotein showing the cleavage sites of action of different enzymes.

1.5- Viral Targets for Drug discovery.

DV has a number of targets that may be used for the discovery of DV inhibitors. The structural proteins; C protein, prM and E proteins perform a number of conformational changes important for the viral entry, assembly and exit. Non-structural proteins found in NS3 and NS5 are also good targets. One example is the NS3 protease that is needed for the polyprotein processing and cleavage. HIV protease inhibitors are a good example of these kinds of drugs that are in clinical use.[36] RdRp is the second target found in NS5 protein. It may be the target of choice regarding the specificity due to its absence in host cells. RdRp plays an important role in viral genome replication.[37] NS3 helicase is a third target that is also, important for viral replication.[38-39] It is needed for the unwinding of the double-stranded RNA intermediate during the genomic replication (figure 1.7).

Identification of potential small molecules that can specifically inhibit the viral life cycle via the inhibition of one of the previously mentioned targets requires a detailed study of each target. In this research; protease, RdRp and helicase were selected for study and are discussed in the following sections.

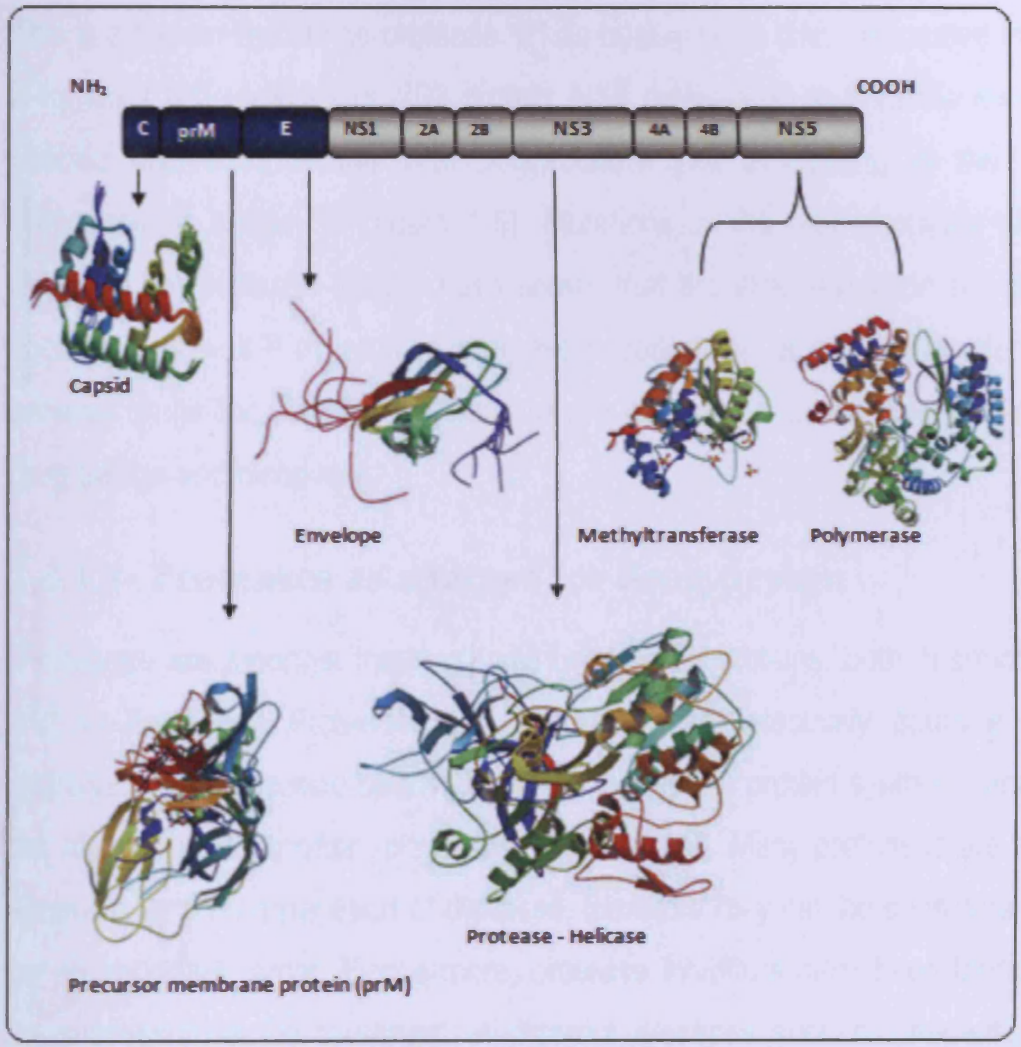


Figure 1.7 Potential drug targets present in the dengue virus genome.

1.5.1- NS3 Serine Protease.

This is a trypsin-like serine protease. It has been proven that it is located in the N-terminal 180 residues of NS3 protein. NS3 protease is responsible for site-specific cleavages in the viral polyprotein.[35] The processing of the viral polyprotein is shown in (figure 1.5). Mutations in the NS3 protease which eliminate the protease activity have shown that the viral replication has been abolished as well.[36] In addition protease inhibitors are a well known class of antiviral drugs for different viruses and are considered as an ideal target for drug design and discovery.

1.5.1.1- Proteases as a target for drug design

Proteases are amongst the most well understood proteins, both in structure and in function.[39] Proteases are enzymes that selectively catalyze the hydrolysis of polypeptide bonds. They are involved in protein synthesis and in the regulation of important physiological processes. Many proteases are also essential for the propagation of diseases, therefore they can be considered to be an important target. Furthermore, protease inhibitors have been found to be successful in the treatment of different diseases such as, hepatitis [40], herpes [41], various forms of cancer and human immunodeficiency virus (HIV).[42]

1.5.1.2- Nomenclature of Serine protease.

The protease enzyme is composed of a number of pockets (**S** pockets), stands for subsite. DV NS3 serine protease has **S1, S2, S3, S1", S2", S3"** pockets. These pockets are essential for the fitting of the substrate that is required to be cleaved by protease. On the other hand, the substrate has some side chains that are labeled by "**P**" letter. Every **P** side chain fits with a

specific pocket. For example **P1** fits into **S1** pocket, **P2** for **S2** and so on (Figure 1.8). Residues in the N-terminal direction from the amide bond that is cleaved (scissile bond) by the protease are called the nonprime side and are designated **Pn.** However, the residues in the C-terminal directions are referred to as the prime side and are labeled **Pn"** and the corresponding **Sn** and **Sn"** in the enzyme are designated for the N- and C-terminal respectively.[43]

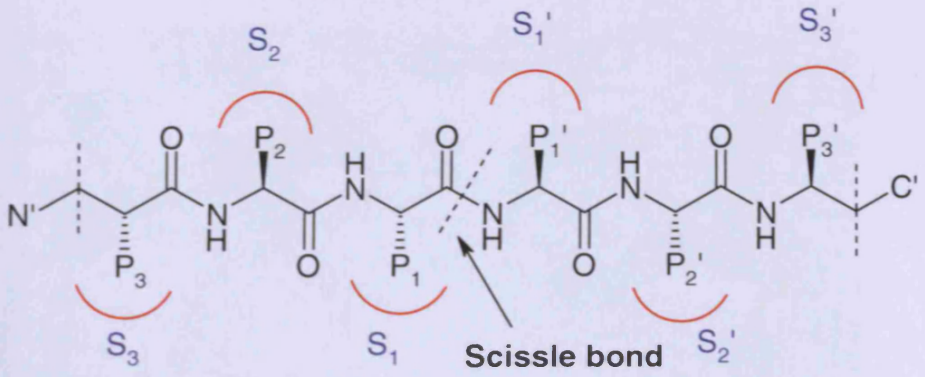


Figure 1.8 Schematic representation of the standard nomenclature of a protease/substrate complex.

There are four types of proteases; aspartic, cysteine, metallo, and serine proteases. In each type the proteases are often very similar, thus in designing a protease inhibitor it is important to consider possible cross inhibition of other proteases in the same subclass.[44] DV has a NS3 serine protease enzyme which is a trypsin-like serine protease. The proteolytic mechanism here is the same as all serine proteases that have a catalytic triad in their active site composed of Ser, His, and Asp residues in order to activate the cleavage and processing of the polyprotein.

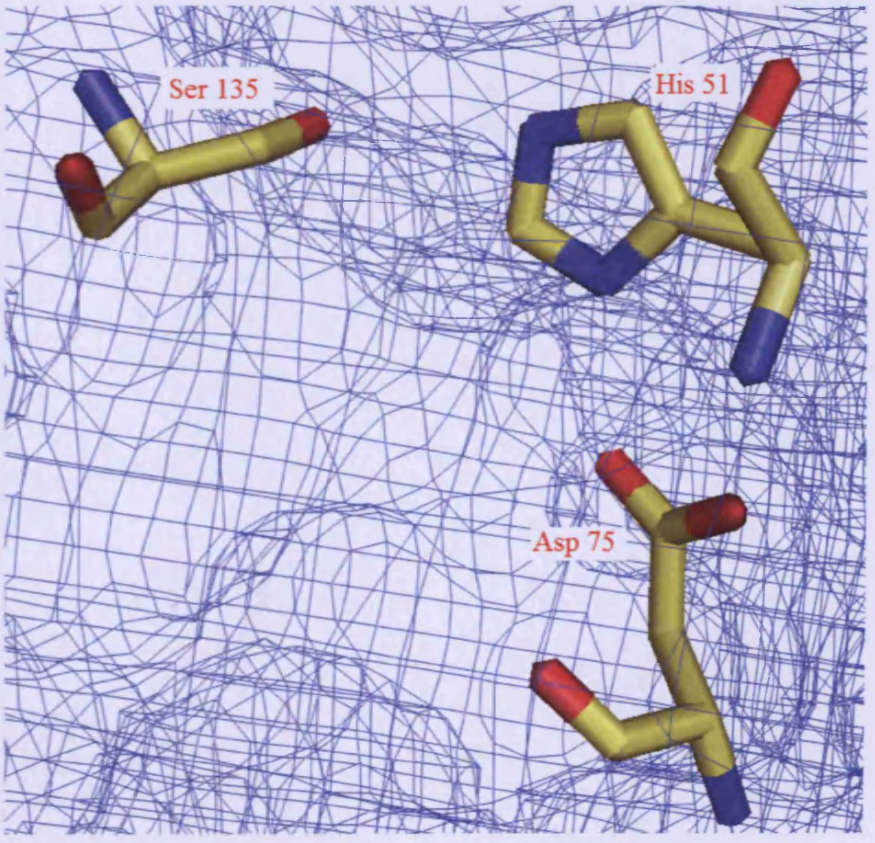
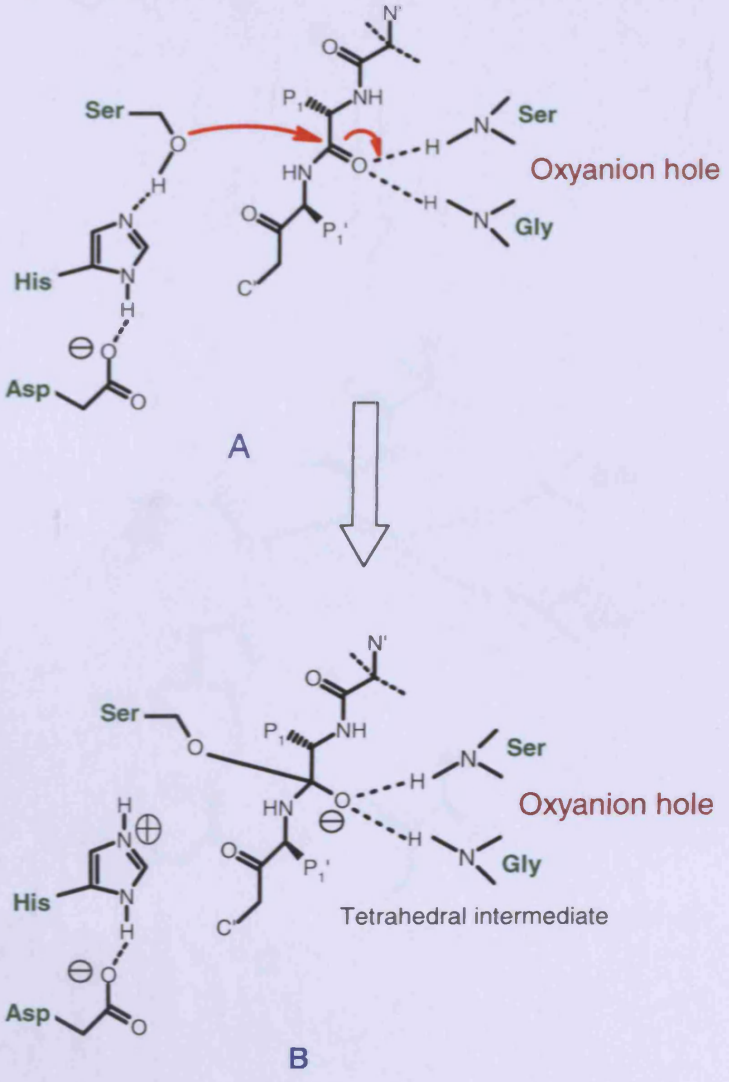


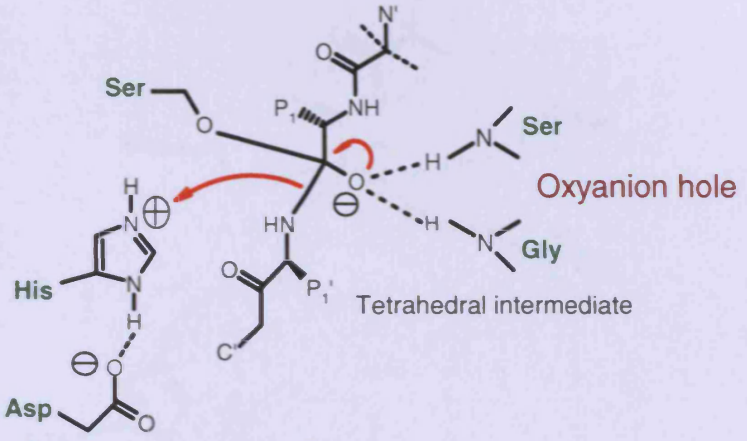
Figure 1.9 The active site of a serine protease consists of a catalytic triad that is composed of three residues Ser, His and Asp which are close to each other. According to DV NS3 protease numbering it will be Ser 135, His 51, and Asp 129

1.5.1.3- Catalytic Mechanism of Serine Protease

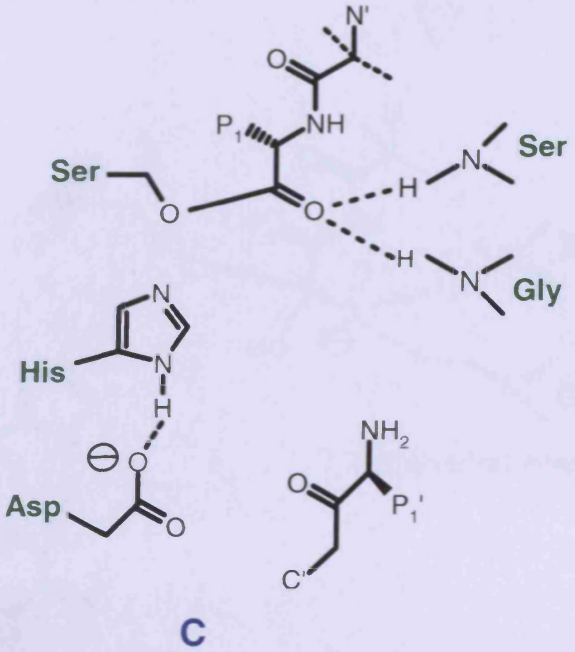
When the substrate binds to the active site "The Michaelis complex" is formed in which the –NH of His 51 will interact through a H-Bond formation with the carboxylic group of aspartic acid (Asp 75). This will catalyze the hydroxyl group of Ser residue 135 to act as a nucleophile to attack the carbonyl group of scissile amide bond in the substrate (**A**). In this case a tetrahedral intermediate is formed and stabilized by H-Bond formation with –NH of both Gly 136 and Ser 135 residues which will form what is called *Oxyanion hole* (**B**).[45]



A process which involves the transfer of proton from His of the catalytic triad to the amine of the tetrahedral intermediate takes place in order to liberate the free amine fragment (**C**).

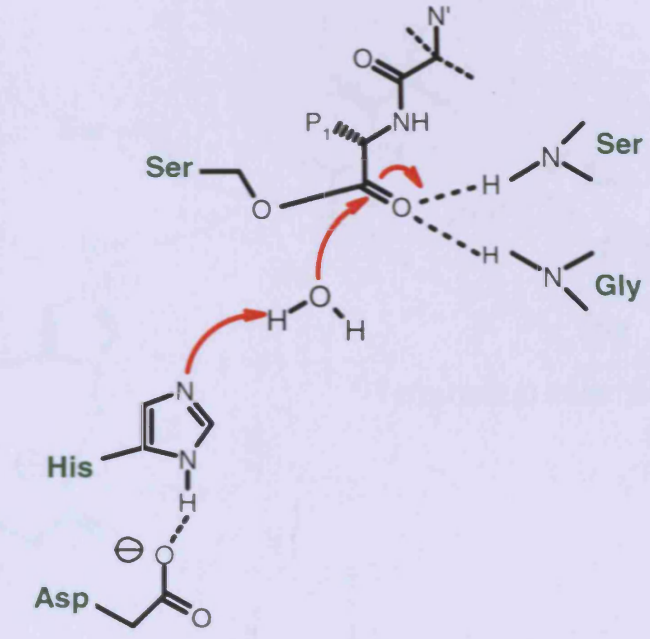


B

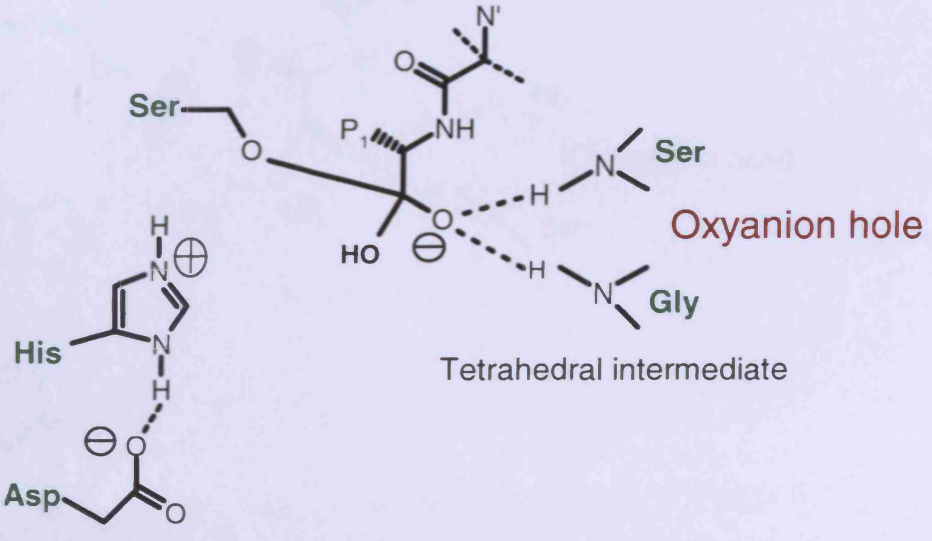


C

In the presence of a molecule of water the acyl-enzyme complex will interact with the water to form a new tetrahedral intermediate (**D**).

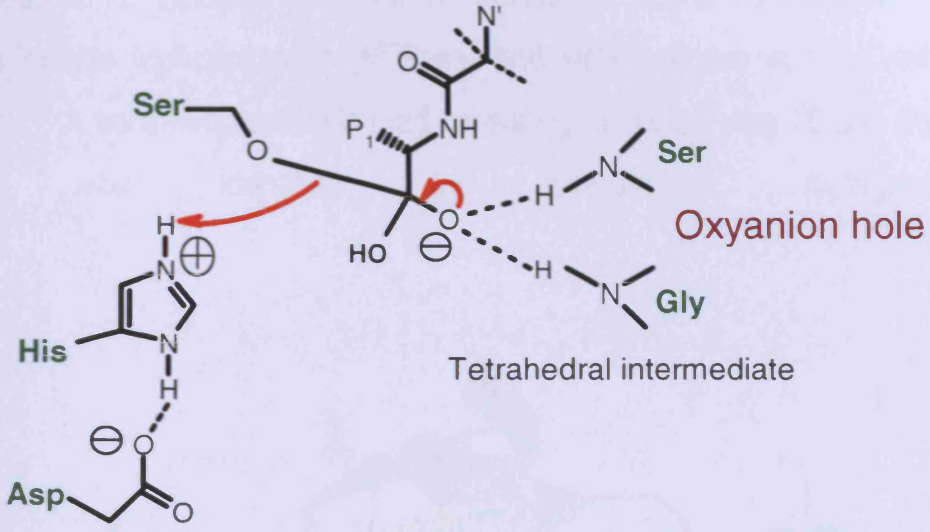


C



D

The newly formed tetrahedral intermediate will break in an acid medium to form carboxyl terminal from the cleaved substrate and Ser will become free.[46]



Tetrahedral intermediate

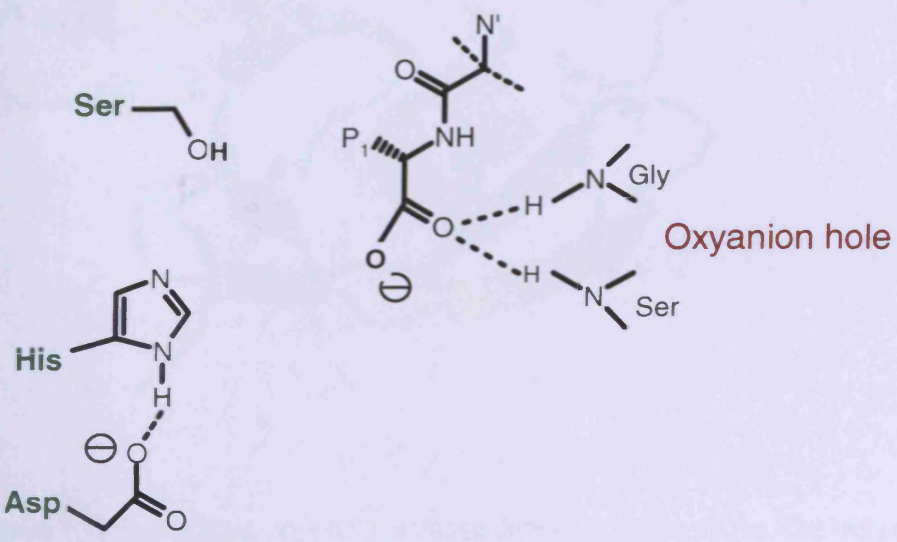


Figure 1.10 the mechanism of Serine protease.

1.5.1.4- NS3 protease Domain of DV.

DV NS3 protein is a multi-functional protein with a trypsin-like serine protease domain located within the N-terminal 180 amino acid residues [46] and A nucleoside triphosphatase (NTPase) and RNA helicase in the C-terminal moiety.[47] A conserved catalytic triad consisting of His 51, Asp 75 and Ser 135 and was identified by site-directed mutagenesis.[48]

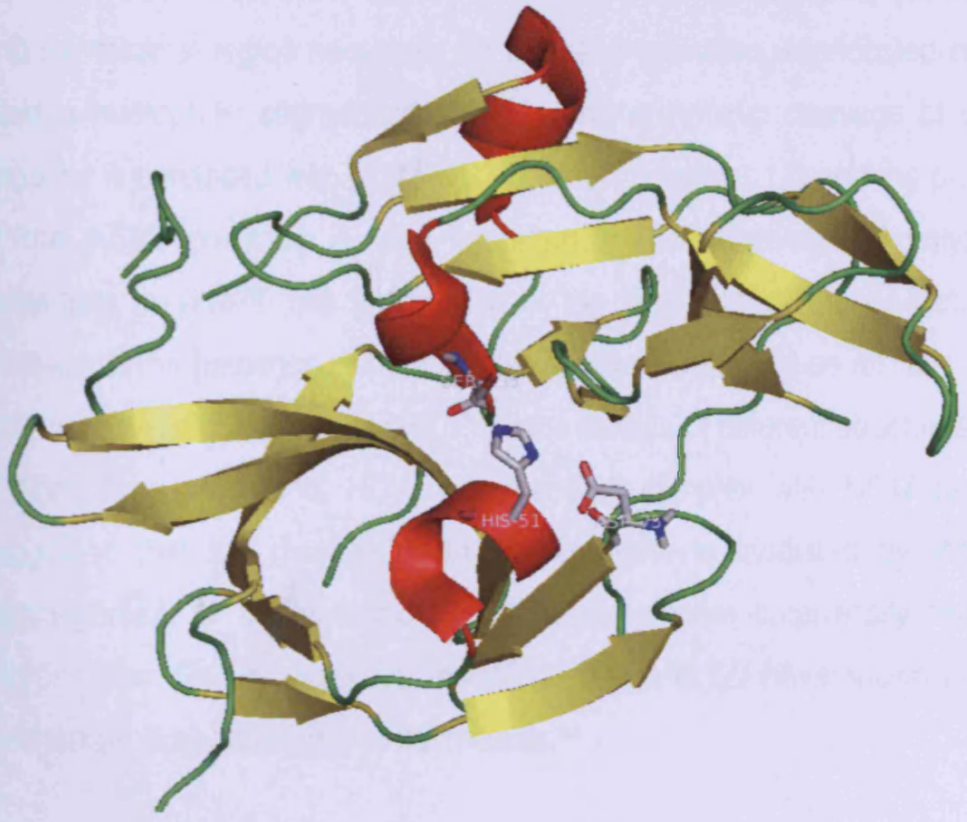


Figure 1.11 The dengue virus NS3 protease domain is shown above. The helix is colored red, sheets are in yellow, and loops are in green. The catalytic triad is shown in sticks.

The three dimensional structure of the dengue virus NS3 protease in complex with a co-factor NS2B was resolved at a resolution of 2.1 A°.[49] It has been observed that the structure of the dengue NS3/NS2B protease is highly similar to that of the WNV NS3/NS2B protease and also to the hepatitis C virus NS3/NS4A complex. The difference might be in the structure of the cofactor dependent activation mechanism of the two proteases.[50] The presence of the co-factor is essential for the activity of the NS3 protease which is a small activating protein important for optimal activity of the flaviviral NS3 proteases. In the case of DV NS2B is considered to be the co-factor required for this purpose and the minimal region necessary for protease activation was located in a 40-residue hydrophilic segment of NS2B.[51] The enzymatic cleavage of dibasic peptides is enhanced with NS2B-NS3 complex (Figure 1.12) and the presence of the NS2B co-factor is very important for the cleavage of polyprotein substrates in vitro.[52] The importance of the presence of the co-factor with protease is the presence of hydrophobic residues which act as an "anchor" for the protease-cofactor interactions. Previous studies of different structures of the NS3pro such as that of HCV, alone and in complex with NS4A-peptide[53] suggested that the mechanism of the co-factor is mediated by the local rearrangement of the catalytic triad towards a more catalytically favorable conformation. Recent molecular modeling studies in DV have shown that this mechanism is also available in flaviviruses.[54]

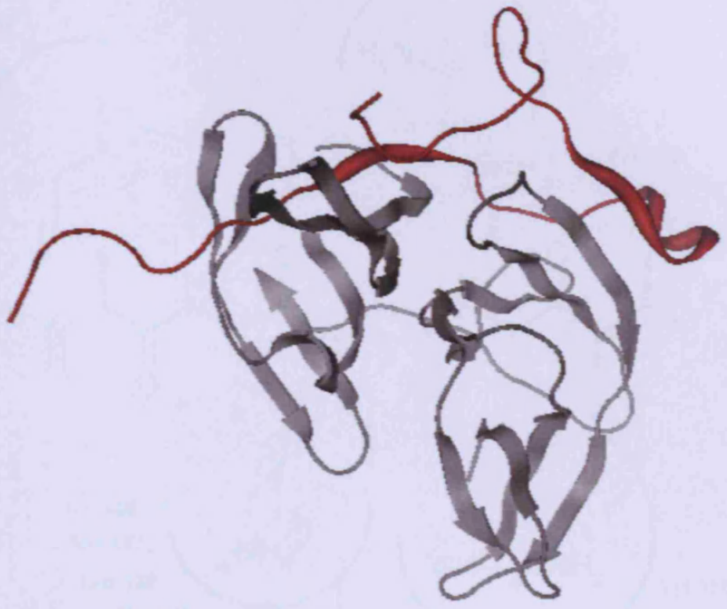


Figure 1.12 NS3/NS2B complex showing the NS2B chain in red color.

1.5.1.5- Description of the crystal structure of the Dengue virus NS3 serine protease.

The NS3/NS2B complex catalyses the cleavage of the polyprotein at specific sites; NS2A-NS2B, NS2B-NS3, NS3-NS4A, and NS4B-NS5. These sites of the polyprotein have (Lys-Arg, Arg-Arg, Arg-Lys) in the **P1** side chain and (Gln-Arg) in the **P2** side chain then short chain amino acids (Gly, Ala, Ser) in the **P1"** side chain. The selectivity of proteases for particular substances is due to the existence of specific binding sites present on the enzyme surface for amino acid side chains of the substrate(s). Here the substrate is oriented by binding of the amino acid side chain of the **P** substrate residue in the corresponding **S** pocket of the receptor (figure 1.13).[55]

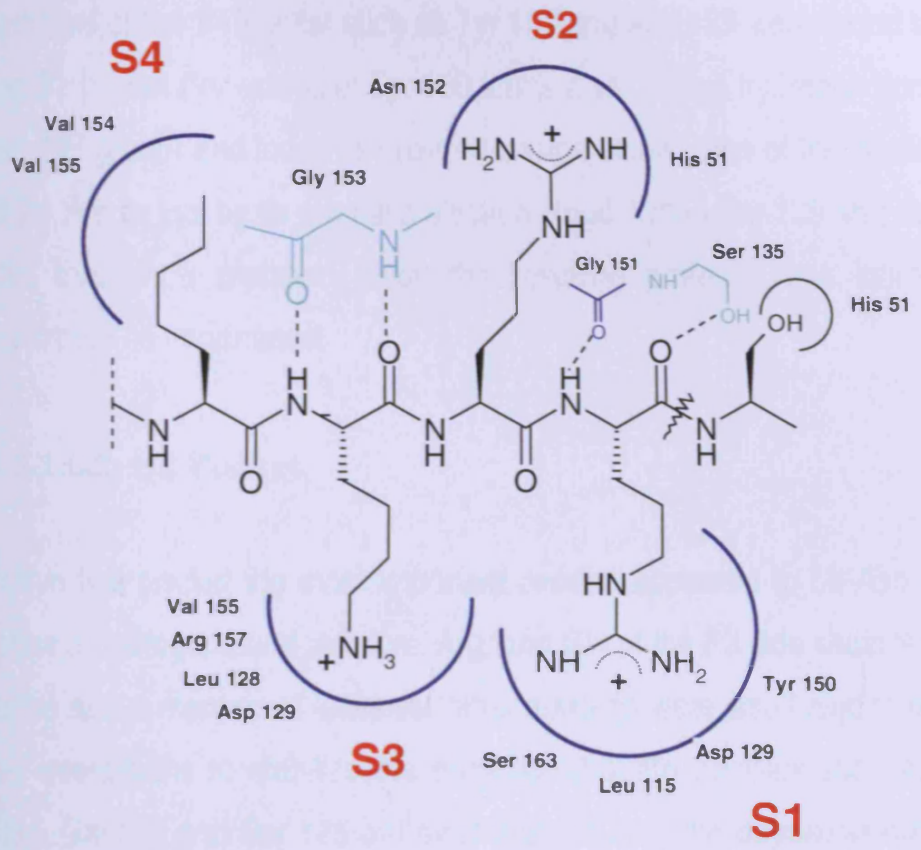


Figure 1.13 Substrate-receptor pockets complex.

1.5.1.6- Important Receptor Pockets.

1.5.1.6.1- S1 Pocket.

S1 pocket contains Ser 131, Tyr 150, Ser 163, Asp 129, Phe 130, Ser 135 and leu 115. These residues are capable of making interactions with guanidine N atoms of Arg residue at **P1** and they are important for the determination of substrate specificity. The following interactions were found; Leu 115 forms Van der Waals interactions with Arg from the substrate. Asp 129 lies in the bottom of **S1** and provides a salt bridge with Arg of the substrate side chain. Important

residues of the **S1** pocket such as Tyr 150 and Asp 129 were found to stabilize the **S1** pocket (for example; Tyr 150 forms a salt bridge hydrogen bond through its –OH group) and found to provide charge stabilization of the positive charge of **P1** Arg or Lys by its aromatic electron cloud. When Ser 135 was replaced by Cys through a mutation study the resulting protease was inactive which confirmed its importance.

1.5.1.6.2- S2 Pocket.

Within this pocket the most important residue appeared to be Asn 152 which forms a hydrogen bond with Lys, Arg, and Gln of the **P2** side chain.[56] In addition to the above mentioned residues, other residues were also found to be useful in the interactions to stabilize the enzyme-substrate complex such as Ala 160. Also, Gly 133 and Ser 135 are most likely to form the ***oxyanion hole*** which is involved in the formation of the tetrahedral transition state with an acyl-enzyme intermediate. Site-directed mutagenesis studies were done to some of these residues to prove its role in protease activity. When His 51 residue was replaced with Ala residue the activity of protease was abolished.[57] Also, upon replacement of Ser with Thr the activity was removed.[58] Mutations of the catalytic triad residues completely abolished the activity.[59] In DV-2 NS3/NS2B complex a site-directed mutagenesis of Asp 129, Tyr 150, and Gly153 were included [60] Mutation of Asp 129 with Glu, Ala, and Ser did not affect the activity of protease which was retained and upon replacement of Asp 129 with Arg, Lys, and Leu protease activity was decreased but not abolished.[60] Tyr 150 residue was also replaced by Ala, Val, and His where the activity was eliminated.

1.5.1.7- Dengue NS3 protease inhibitors

The success of some drugs such as ACE inhibitors of the metalloprotease class and HIV protease inhibitors of the aspartyl protease class, have served to increase the efforts to identify the inhibitors of other proteases like serine protease as drugs. The first inhibitor for the dengue NS3-protease was the **Bowman-Birk** inhibitor that was extracted from mung beans (MbBBI). The crystal structure of the inhibitor complexed with dengue NS3-pro was reported and can be found in the protein data bank with PDB code = 1df9.

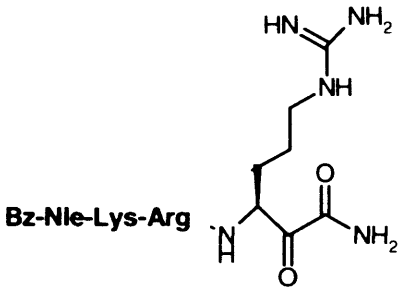
Figure 1.14 MbBBI cartoon representation (yellow) -Dengue NS3 protease (green). Arg and Lys residues in red color.

MbBBI is extracted from a leguminous plant [61] and contains an anticancer activity.[62] It is double-headed (composed of 2 β-sheets connected together with 7 disulfide bridges and a hydrophobic core) so it can inhibit two molecules of protease at the same time by the standard serine protease inhibition mechanism.[62] NS3 domain alone and the NS3/NS2B complex are inhibited by this inhibitor. While it was assumed that the NS2B cofactor is required for the proteolytic activity of the NS3, the NS3 alone can be inhibited by small molecule inhibitors.

Peptide inhibitors containing electrophilic functional groups are generally good inhibitors for serine protease as they will compete with the substrate for binding to Ser residue of the catalytic triad. The effect of different electrophilic functional groups on dengue protease was studied (Figure 1.15). [63]

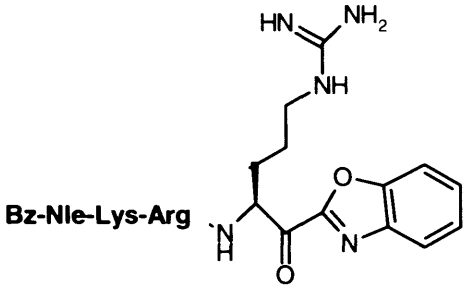
Compound	Ki (μM)
1	127.5
2	82.9
3	0.85
4	42.8
5	0.043

Table **1.1.** Showing the activity of tetrapeptide inhibitors with different electrophilic functional groups.



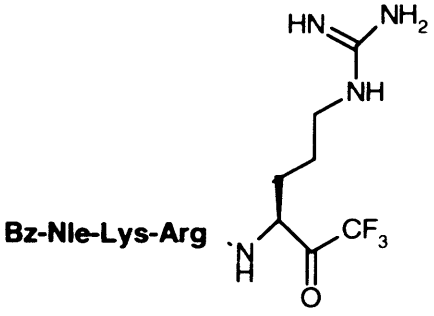
Tetrapeptide α-ketoamide

1



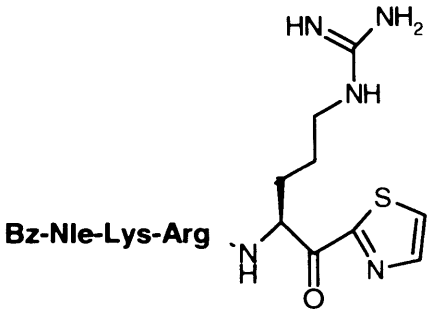
Tetrapeptide α-ketobenzoxazole

2



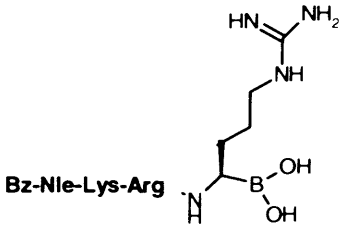
Tetrapeptide trifluoromethyl ketone

3



Tetrapeptide α-ketothiazole

4



Tetrapeptide boronic acid

5

Figure 1.15 tetrapeptide dengue NS3-protease inhibitors. Where Bz- Nle is benzoylated L-norleucine

Different functional groups were used such as –COOH, -OH, $-NH_2$, $-NHSO_2CF_3$, Benzoxazole, thiazole, $-CF_3$, and $-B(OH)_2$ and the activity was evaluated. In contrast to the HCV NS3-4A protease which was inhibited by a peptide acid dengue NS3 protease could not be inhibited by carboxylic acids. Moreover, peptide analog containing trifluoroacetylsulfonamide (which are isostere for carboxylic acids) was not active as well.[65] Tetrapeptide amides showed inhibitory activity but, at a high micro molar concentration. For inhibitors with α-keto heterocycle moiety it was proved to be less active than aldehydic moiety. In addition, the incorporation of trifluoromethyl ketone and boronic acid has yielded potent inhibitors specially as the tetrapeptide boronic acid was evaluated to be the most potent peptide inhibitor of dengue virus NS3 pro with Ki = 43 nM.[64] Generally, peptides without electrophilic functional groups did not show potent inhibition of dengue NS3 pro.

In contrast to the above mentioned points for the tetrapeptide inhibitors, peptides suffer from some limitations when used as drugs that can be summarized in the following points [66]

1. Peptides will have a very short duration of action because they are liable to many proteases and peptidases in our body.
2. Many peptides are too large to pass through the digestive tract into the blood so, it can not be taken orally.
3. Peptide inhibitors lack the desired specificity because they can interact with a number of enzymes in the blood stream and may cause unwanted side effects.

In a trial to find a non-peptidic inhibitor for DF/DHF some compounds were extracted from finger root, ***Boesenbergia rodunta*** (BR) and tested for their inhibition of the dengue NS3 protease. BR is a common spice belonging to a number of the ginger family (***Zingiberaceae***). The chalcone, **Cardamonin** which was isolated from BR and was recently reported to be active against HIV-1 protease.[67] In addition to Cardamonin some compounds were isolated from BR and were tested for their inhibitory activity towards DV-2 virus NS3/NS2B protease.[71] A Fluorogenic peptide substrate contains Arg-Arg residue was used to test the cleavage of Arg-Arg at the **S1** pocket. DV-2 NS3/NS2B protease with and without BR extracts compounds of different concentrations were tested. The cleavage activity of dengue 2 NS2B/NS3 protease at **S1** was inhibited by Panduratin A **11** and 4-hydroxypanduratin A **12** with an activity of 21 μM and 25 μM respectively

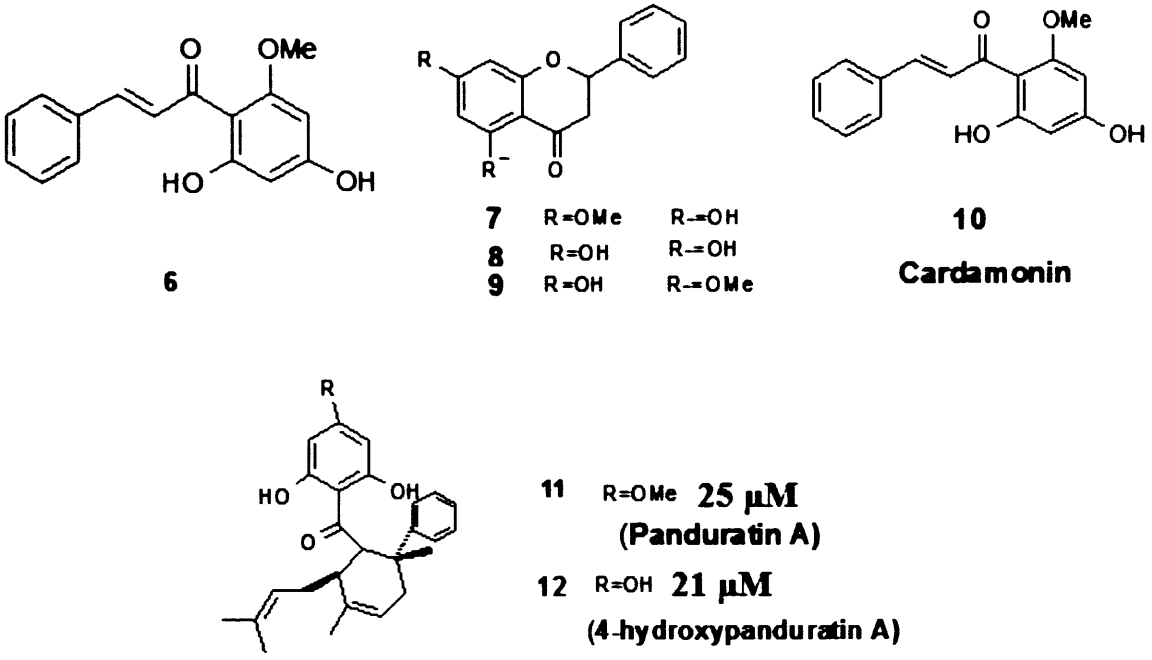
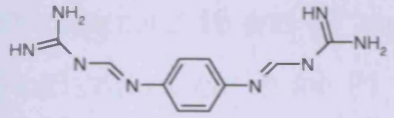


Figure 1.16 *Boesenbergia rodunta* active compounds

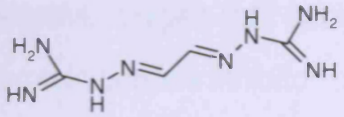
In another study, the crystal structure of dengue NS3 serine protease complexed with the Bowman Birk inhibitor was useful in understanding the important residues and interactions that may help in designing a potent inhibitor. Also, it was observed that NS2B enhances the hydrolysis of tripeptide substrates and has no effect on substrates that have only a **P1** side chain.[69] Furthermore it is similar in this feature to HCV NS3 protease in which the cofactor has a very low effect on **P1.**[70] Depending on the structure of the side chain present in **P1** (Arg, or Lys) [68], and by observing the interactions formed by MbBBI side chains in which Arg side chain from the inhibitor was involved by its guanidine moiety in the interactions with the residues found at the **S1** pocket of the DV NS3/NS2B protease complex. Searching for commercially available compounds that have guanidine scaffold which mimic this mode and can give the same interactions that are done by Arg. 17 compounds were identified from which only 6 compounds were commercially available and were tested against:

- DV-2 NS3/NS2B protease complex.
- WNV NS3/NS2B protease complex.

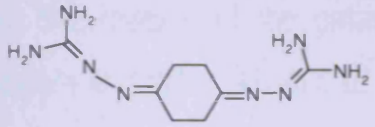
Only 5 compounds show activity from which compounds number **16** and **17** showed the best activity as shown in the table 1.1.



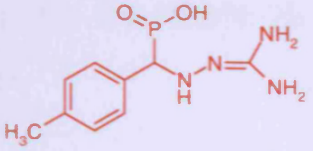
13



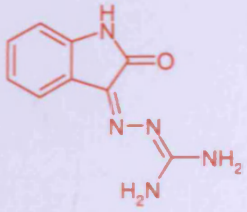
14



15



17



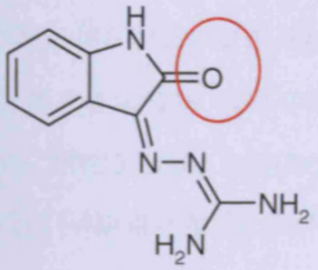
16

Figure 1.17 structures of the DV NS3/NS2B protease inhibitors

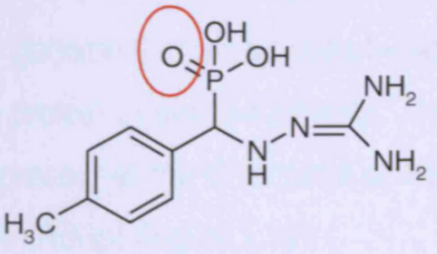
Compound	Ki μM	
	DV-2 NS3/NS2B protease	WNV NS3/NS2B protease
13	44±5	33±5
14	423±50	337±56
15	1783±113	1088±162
16	23±2	16±2
17	14±2	13±1

Table 1.2 The DV NS3/NS2B and WNV NS3/NS2B serine protease inhibitors

Both compound **16** and **17** have an electronegative oxygen that could mimic the carbonyl oxygen in the **P1** side chain of protein/peptide inhibitor complex. This oxygen is of the carbonyl group of indoline moiety in compound **16** and of phosphonate group in compound **17** and in both cases it makes hydrogen bond with Ser residue of the catalytic active site in the DV NS3/NS2B protease complex and WNV NS3/NS2B protease complex.



16



17

Figure 1.18 DV NS3/NS2B and WNV NS3/NS2B serine protease inhibitors, most active compound.

1.5.2- DV RNA dependent RNA polymerase (RdRp)

Viral polymerases represent an attractive drug targets for the development of specific antiviral compounds. Indeed, selective inhibitors against HIV-1 reverse transcriptase have been approved as drugs such as Didanosine, Stavudine.[72] The genomic RNA of the dengue virus is translated into a single polyprotein.[73] This is cleaved into three structural C-prM-E and seven nonstructural NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 proteins by both the viral and host proteases as mentioned before.[74] The viral serine protease is located within the N-region of the **NS3** [75], while the C-region has the RNA Helicase that is working to separate the double-stranded RNA template into individual strands to facilitate the process of replication of the genome that will be done by **NS5** [76-77]. **NS5** is the largest and most conserved protein in dengue proteins. There are some amino acids sequences (motifs) present in the C-terminal of **NS5** form the DV RNA-dependant-RNA-polymerase (RdRp) (Figure 1.19).

Figure 1.19 The Conserved motifs (A, B, C, D, E, and F) of dengue RdRp.

Polymerase is responsible for synthesizing a transient double-stranded replicative RNA intermediate. This RNA is composed of viral plus- and minus-strand RNA strands. The newly synthesized minus strand serves in turn as a template, allowing the RdRp to synthesize additional plus-strand genomic RNA.[78-81]

1.5.2.1- Crystal structure of DV RdRp

Most of the polymerases such as that of the ***Flaviviridae*** family, and other viral families share an overall architecture composed of palm, fingers, and thumb domains [82] and a common catalytic mechanism for nucleotide incorporation by using of two metal ions that are coordinated by two structurally conserved aspartic acid residues.[83] Viral RdRp are characterized by the presence of a connection between the thumb and the finger domains, making the active site appears to be encircled, as seen in HCV RdRp. There is an additional motif, called motif F that is not present in host cells, which has an important role in the stabilization of the recently formed base pair. The three-dimensional structure of the DV RdRp was defined at 1.8 °A by X-ray crystallography (figure1.20). [84]

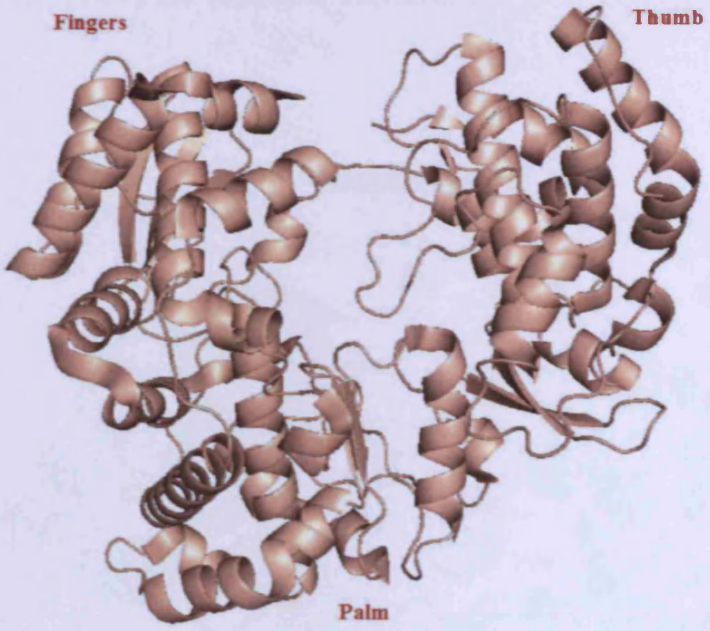


Figure 1.20 Crystal structure of Dengue virus RdRp (pdb code = 2J7U)

1.5.2.3- Catalytic active site

The catalytic active site of the DV RdRp is found in the palm domain. It consists of residues 497 to 542 and 601 to 705; it appears to be the most structurally conserved. The catalytic domain of the DV RdRp shows good superimposition with those of other RdRps (e.g., HCV), with RMS deviations of about 1.7 Å.

1.5.2.4- Different motifs found in DV RdRp

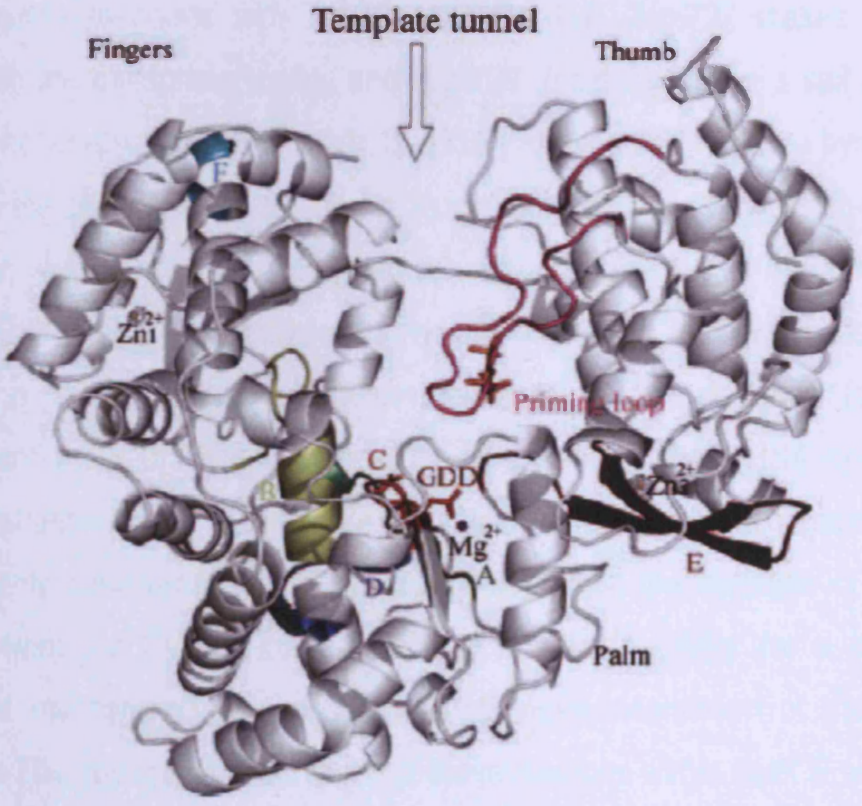


Figure 1.21 showing the different motifs of dengue RdRp. Motif A is in violet, motif B in white, motif C is in cyan, motif D is in pink, motif E is in blue, and motif F is in red color.[84]

DV RdRp has five important motifs that are distributed in its structure; Motif A has one hydrated Mg^{2+} ion close to the expected catalytic position. Motif C contains catalytic triad GDD (Gly 662, Asp 663 and Asp 664).

Motifs B and D are shown in the figure 1.21. Motif E is the most important conserved motif in the thumb domain. Additional interactions between the polymerase and the 3'-deoxy guanosine triphosphate (3'-dGTP) phosphate include hydrogen bonds with Thr-794 and Ser-796. Arg-737 makes a salt bridge with the phosphate moiety and Arg-729 (motif E) makes a salt bridge with the phosphate as well. However, Ser-710 (from motif E) makes a hydrogen bond with the phosphate. In a separate study, residues Ser-498 and Arg-517 in the Bovine viral diarrhea virus polymerase (corresponding to Ser-710 and Arg-729 in DENV RdRp) were mutated individually to alanine in order to test their contribution to an elongating (primer-dependent) versus de novo (primer-independent) mode of RNA synthesis. The de novo mode of the RNA synthesis was almost abolished by these single mutations, while the RNA elongation was reduced only approximately two- to nine folds. Thus, the mutation of these residues from the BVDV RdRp is able to confer specificity for a primer-dependent mechanism, as opposed to a de novo mechanism of the RNA synthesis. The structural conservation of these residues within motif E with the BVDV RdRp suggests a similar phenotype when Ser-710 and Arg-729 are mutated in the DENV RdRp enzyme. A comparable study of the HCV RdRp demonstrated a similar role played by Arg-386 and Arg-394 (also leading to a severe decrease in de novo initiation). These two residues are structurally equivalent to Arg- 729 and Arg-737 in the DENV 3 RdRp. Taken together; these data suggest an essential role for this GTP-binding site for de novo initiation of RNA synthesis by the DENV RdRp. Motif F plays an important role in the stabilization of the nascent base pair.[84]

1.5.2.5- Mechanism of Action of DV RdRp.

The process starts by the capped genomic (+) RNA that is found circularized (1) due to interactions between complementary sequence stretches at the 5' and 3' ends named UAR (Upstream of AUG Region) and CS (complementary sequences). The NS5 RdRp domain specifically binds to the promoter stem loop A (SLA) at the 5' end of the genome and initiates de novo RNA synthesis using the 3' end as a template (2). The **NS5** RdRp synthesizes the complete (–) strand (3). The double-stranded RNA Replicative form (RF), consisting of the genomic strand annealed to the neosynthesized (–) strand (4), serves as a template for the synthesis of a new genomic (+) RNA, in the Replicative intermediate (RI) (5).[85-86] (Figure 1.22).

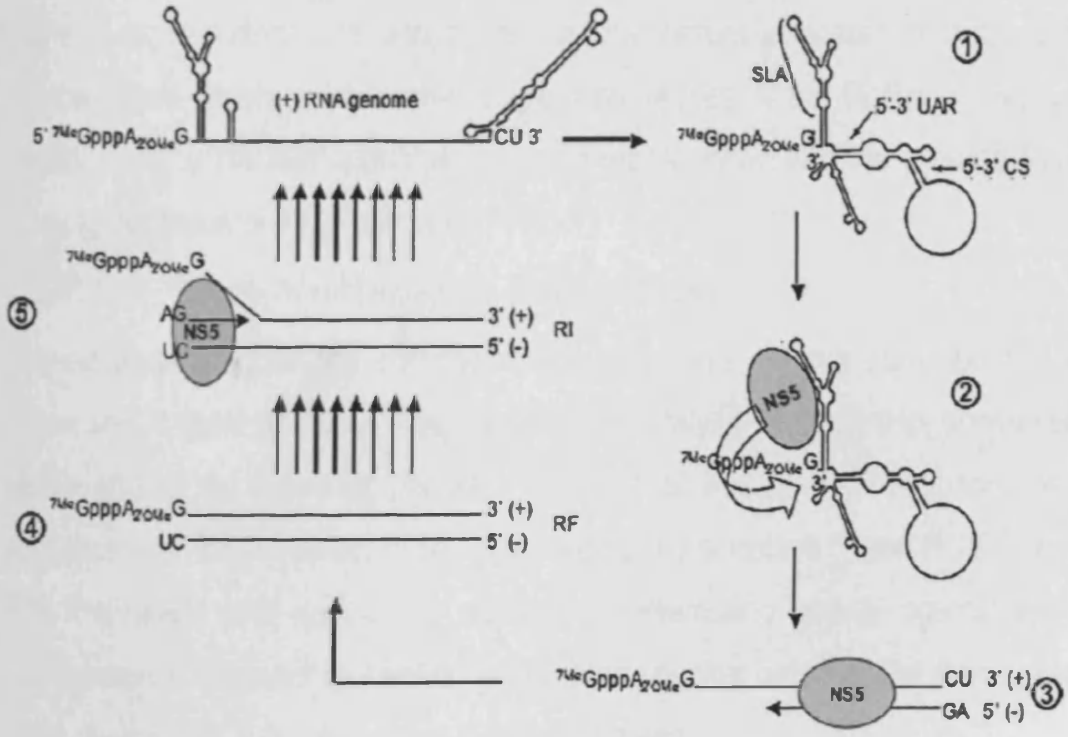


Figure.1.22 RdRp in Dengue virus replication.[85]

1.5.2.6- DV RdRp Inhibitors.

As mentioned before, the dengue virus RdRp is considered to be an interesting enzyme that could provide specific inhibition of the viral replication because polymerases are not found in the host cells.[87] Current RdRp inhibitors work either by directly chelating the catalytic active site in the palm domain, binding to the conserved active site or by binding to allosteric pockets.[88] According to the chemical structure of RdRp inhibitors they may be classified into:

- Nucleoside inhibitors.
- Non- Nucleoside inhibitors.
- Miscellaneous inhibitors.

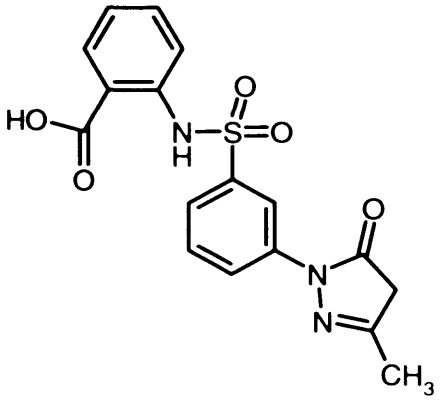
1.5.2.6.1- Nucleoside Inhibitors.

Nucleoside inhibitors are analogues of the naturally occurring nucleotide triphosphates which are the physiological substrates of the RdRp. These are metabolized to the triphosphates and then incorporated into the growing RNA causing inhibition of RNA synthesis.[89-90]

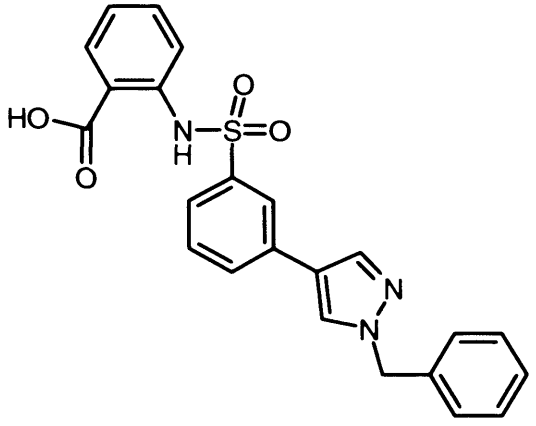
1.5.2.6.2- Non-Nucleoside Inhibitors.

Non-nucleoside inhibitors are chemical agents that do not have purine or pyrimidine in their structure. They bind to the catalytic site of RdRp, conserved active site or to allosteric pockets. The use of this kind of inhibitors was successful for the inhibition of the HIV-reverse transcriptase [91] and HCV RdRp. [92-95] Anthranilic acid derivatives showed an interesting activity against HCV polymerase.[96] They act by binding an allosteric pocket between the thumb and palm domain at 7.5 A° away from NTP binding site.

Recently N-sulfonylanthranilic acid derivatives were also discovered with some inhibitory activity against Dengue RdRp via an allosteric inhibition,[97] the anthranilic acid derivatives have confirmed its great affinity towards the RdRp enzyme. First they synthesized compound **18** that was tested and found to have IC50 = 7.2 μM inhibitory concentrations by which it was considered to be a good hit to be modified to give compound **19** with IC50 = 0.7 μM and here they reached an interesting lead compound.



18



19

Figure.1.23 DV RdRp anthranilic acid derivatives inhibitors.

1.5.3- DV NS3 Helicase as a target for drug design.

In the process of viral replication a negative stranded RNA is synthesized from a positive stranded RNA. The –RNA strand is then used as a template for the synthesis of complementary + RNA strand. As mentioned in the mechanism of the RdRp enzyme a double-stranded replicative form is synthesized. The helicase enzyme mediates the unwinding and separation of these two strands. Here the helicase works by disrupting the hydrogen bonds between the two strands. The energy required for this process is derived from the NTP hydrolysis. The helicase enzyme is found in the NS3 protein of the dengue virus, where the protease domain is located in the N-terminal region and the Helicase together with NTPase are in the C-terminal region.[98] Flavivirus helicases are classified into three super families (SF); Super family 1 (SF1), Super family 2 (SF2), and Super family 3 (SF3) beside two other small families. Flaviviridae helicases belong to **SF2.**[99] They contain seven conserved motifs (I, Ia, II, III, IV, V, and VI) as shown in (figure 1.24).

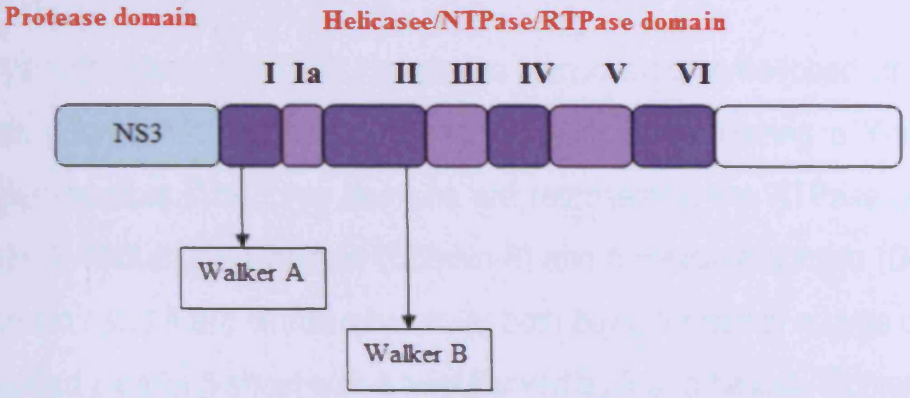


Figure 1.24 Schematic representation of the helicase motifs.

From the seven conserved motifs there are two important motifs which are Motif I and Motif II:

- **Motif I (Walker A):** also called phosphate-binding loop motif (P-loop) that binds the terminal phosphate group of NTP. The presence of a Gly-Lys-(Thr/or Ser) sequence is essential for this motif as the amino group of lysine residue will interact with the phosphate of MgATP/MgADP, and the hydroxyl group of threonine or serine will ligate the Mg^{+2} ion.[99]
- **Motif II (Walker B):** has a general form of DEAH that is for Asp-Glu-Ala-His respectively. The carboxylic group of aspartic acid coordinates the Mg^{+2} ion of MgATP/MgADP. Glutamic acid acts as a catalytic base in ATP hydrolysis.

1.5.3.1- The crystal structure of Dengue Helicase.

The crystal structure of DV NS3 Helicase is composed of three-lobed structural domains (Figure 1.25) of about 150 amino acids each forming a Y-shaped triangular structure. The three domains are representing the NTPase domain (Domain I), RNA-binding domain (Domain II) and a Helicase domain (Domain III). Domain I and II are structurally similar both being formed of a large central six-stranded parallel β-sheet with a twist flanked by four α-helices. Domain III is composed of a bundle of four parallel α-helices surrounded by three α-helices and augmented by two antiparallel β-strands. Domain I and II form a core that contains most of the conserved sequence motifs at the interface between the two domains. The **Walker motif A** is composed of a P-loop that is located on the surface of domain I and binds to the terminal phosphate group of the NTP. **Walker motif B** is also found in domain I and is responsible for the chelating of Mg^{2+}.[100]

1.5.3.2- Importance of Helicase Enzyme.

Helicase and polymerase form a complex with other host and viral proteins to form the viral replicase multiprotein complex. The +ss RNA virus needs a negative strand RNA to be synthesized by the replicase complex using the + RNA strand as a template. Then the newly synthesized – ve RNA is used as a template for further synthesis of +ve RNA strands. Because +ve and –ve strands are complementary, Helicase enzyme is needed for the separation of the strands.

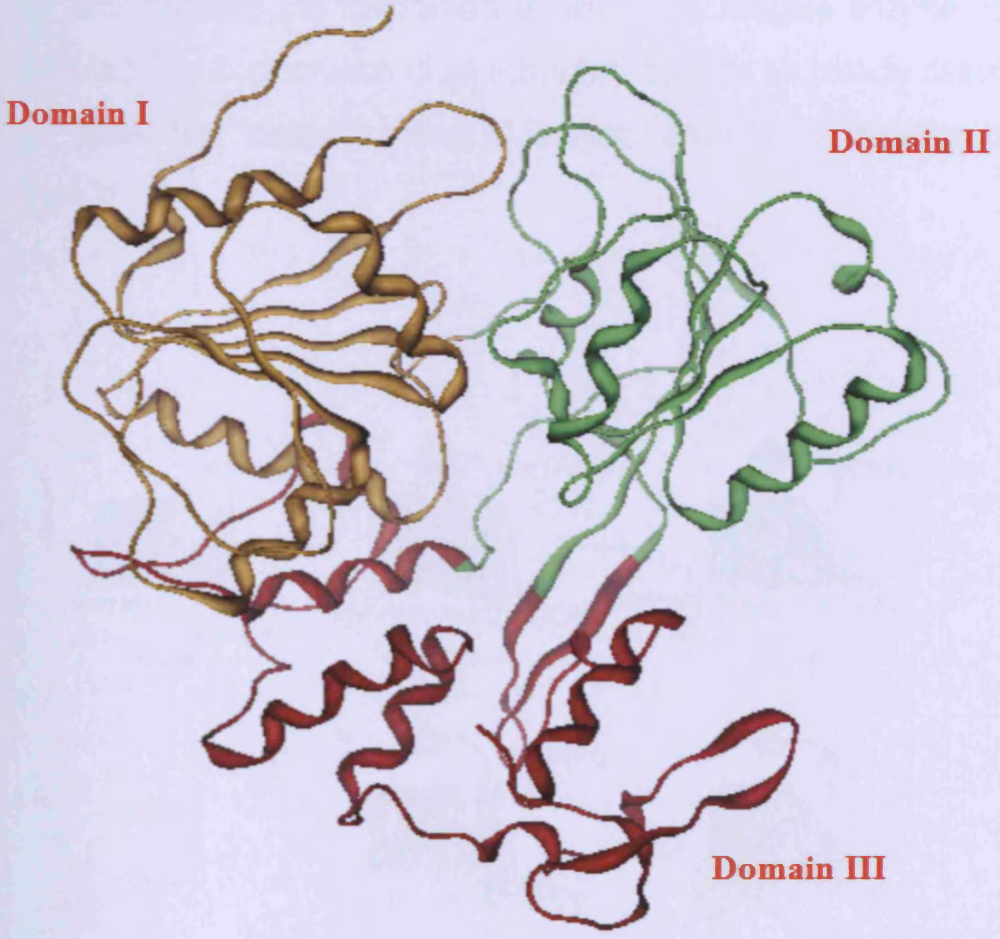


Figure 1.25 Crystal structure of Dengue virus Helicase enzyme showing Domain I (yellow), Domain II (green) and Domain III (red).

1.5.3.3- Mechanism of Flavivirus Helicase inhibition.

The conformation changes which occur in the helicase are an important issue for understanding the mechanism by which the helicase enzyme could be inhibited. The conformation of an active helicase can be broadly described as an "open" and "closed" complex of Domain I and II with a transition between them.[99]

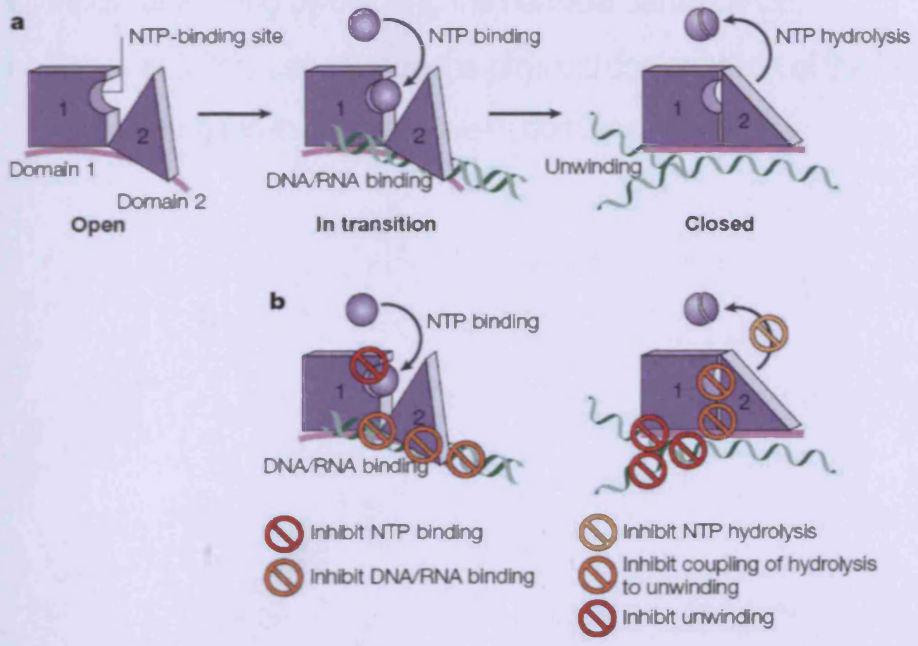


Figure 1.26 Potential mechanism of helicase inhibition.[91]

Inhibitors can have one of the following mechanisms to inhibit the helicase function:

- Inhibit the NTPase activity by direct competition with the NTP binding site.
- Inhibit nucleic acid binding.
- Inhibit NTP hydrolysis or NDP release by blocking the movement of domain II.
- Inhibit the process that couples the NTP hydrolysis to translocation and unwinding.
- Inhibit unwinding by blocking the helicase translocation.
- Other inhibitors can change the physical conformation of the helicase and altering the interface between domains I and II.

1.6- References

1. Knipe, D. M.; Howley, P. M. *Fields Virology;* 5th Edition. Eds. Lippincott-Raven Publishers, Philadelphia, **2007**.
2. Thiel, H. J.; Collett, M. S.; Gould, E. A. Family Flaviviridae. International committee on taxonomy of viruses. San Diego: *Academic press.* **2005**, 979-996.
3. Gubler, D. J. Dengue and Dengue Haemorrhagic fever: its history and resurgence as a global public health problem. In: Gubler, D. J.; Kuno, G. eds. Dengue and Dengue Haemorrhagic Fever. Walling Ford, UK: CAB International. **1997,** 1-22.
4. Gorbalenya, A. E.; Koonin, E. V. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic acid Res.* **1989,** *17*, 4713-4730.
5. Valle, R. P. C.; Falgout, B. Mutagenesis of the NS3 protease of dengue virus type 2. *J. Virol.* **1998**, *72*, 624-632.
6. Krishna, M. H. M.; Clum, S.; Padmanabhan, R. Dengue virus NS3 protease, crystal structure and insights into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effects. *J .Biol. Chem.* **1999**, *274*, 5573-5580.
7. Kim, J.; Morgenstern, K. A. Crystal structure of the hepatitis c virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell.* **1996**, *87*, 343-355.
8. Rushika, P.; Richard, J. K. Structural proteomics of dengue virus. *Curr. Opin. Med. Chem.* **2008**, *11*: 369-377.
9. Henchal, E. A.; Putnak, J. R. The Dengue Viruses. *Clin. Microbiol. Rev.* **1990,** *3*, 376–396.

10. Deubel, V.; Kinney, R. M.; Trent, D. W. Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of dengue type 2 virus, Jamaica genotype: comparative analysis of the full-length genome. *Virology.* **1988**, *16,* 234–244.
11. Fu, J.; Tan, B. H.; Yap, E. H. Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90). *Virology.* **1992**, *188,* 953–958.
12. Osatomi, K.; Sumiyoshi, H. Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology.* **1990**, *176*, 643–647.
13. Zhao, B.; Mackow, E.; Buckler-White A. Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding for structural proteins. *Virology.* **1986**, *155*, 77–88.
14. Morens, D. M.; Fauci, A. S. Dengue and hemorrhagic fever: a potential threat to public health in the United States. *JAMA.* **2008**, *299*, 214-216.
15. Gould, E. A.; Solomon, T. Pathogenic flaviviruses. *Lancet.* **2008**, *371*, 500-509.
16. Gubler, D. J. Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev.* **1998**, *11,* 480–496.
17. Siler, J. F.; Hall, M. W.; Hitchens, A. P. Dengue: its history, epidemiology, mechanism of transmission, etiology, clinical manifestations, immunity, and prevention. Philippine J. *Sci.* **1926**, *29,* 1–304
18. Simmons, J. S.; St John, H. F.; Reynolds, H. K. Experimental studies of dengue. Philippine J. *Sci.* **1931**, *44*, 1–251.
19. WHO. Dengue Hemorrhagic Fever: Diagnosis, Treatment, Prevention, and Control. Geneva: *WHO.* **1997.**
20. Nimmannitya, S. Clinical spectrum and management of dengue hemorrhagic fever. Southeast Asian J. *Trop. Med. Public Health.* **1987**, *18*, 392–387.

21. Sumarmo, W. H.; Jahja, E. Clinical observations on virologically confirmed fatal dengue infections in Jakarta, Indonesia. *Bull. World. Health. Organ.* **1983**, *61*, 691–701.

22. Thomas, S. J.; Strickman, D. Vaughn, D. W. Dengue epidemiology: virus epidemiology, ecology, and emergence. *Adv. Virus. Res.* **2003**, *61*, 235–289.

23. http://library.thinkquest.org/05aug/00235/whatisit.html.

24. Aruna, S.; Padmanabhan, R. Molecular targets for flavivirus drug discovery. *Antiviral Res.* **2009**, *81*, 6-15.

25. Mukhopadhyay, S.; Kuhn, R. J.; Rossmann, M. C. A structural perspective of the flavivirus like cycle. *Nat. Rev. Microbiol.* **2005**, *3*, 13-22.

26. Stiasny, K.; Heinz, F. X. Flavivirus membrane fusion. *J. Gen. Virol.* **2006**, *87*, 2755-2766.

27. Yu, M. I.; Zhang, W.; Holdaway, H. A.; Yi, L.; Kostyuchenko, V. A.; Chipman, P. R.; Kuhn, R. J.; Rossmann, M. G.; Chen, J. Structure of the Immature Dengue Virus at Low pH Primes Proteolytic Maturation. *Science.* **2008**, *319*, 1834–1837

28. Premkumar, A.; Horan, C. R.; Gage, P. W. Dengue Virus M Protein C-Terminal Peptide (DVM-C) Forms Ion Channels. *J. Membr. Biol.* **2005**, *204*, 33–38

29. Klasse, P. J.; Bron, R.; Marsh, M. Mechanisms of Enveloped Virus Entry into Animal Cells. *Adv. Drug Delivery Rev.* **1998**, *34*, 65–91

30. (a) Kuhn, R. J.; Zhang, W.; Rossmann, M. G.; Pletnev, S. V.; Corver, J.; Lenches, E.; Jones, C. T.; Mukhopadhyay, S.; Chipman, P. R.; Strauss, E. G.; Baker, T. S.; Strauss, J. H. Structure of Dengue Virus: Implications for Flavivirus Organisation, Maturation, and Fusion. *Cell.* **2002**, *108*, 717–725. (B) I-Mei, Y.; Jue, C. Structure of the Immature Dengue Virus at low pH Primes Proteolytic Maturation. *Science.* **2008**, *319*, 1834-1837.

31. Chambers, T. J.; Hahn, C. S.; Galler, R. and Rice, C. M. Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* **1990**, *44*, 649–688.

32. Pieter, L.; Erik, D. C.; Johan, N. Perspectives for the treatment of infections with Flaviviridae. *Clin. Microbiol. rev.* **2000**, *13* (1), 67-82.

33. Behrens, S. E. L.; Tomei, R. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **1996**, *15*, 12–22.

34. Chambers, T. J.; Ronald, C. W.; Arash, G.; David, W.; McCourt, J.; Fernaldo, B.; Robert, J. F.; Charles, M. R. Evidence that the N-terminal domain of non-structural protein NS3 from yellow fever virus is a serine protease responsible for site- specific cleavages in the viral polyprotein. Proc. *Nat. Acad. Sci.* USA. **1990**, *87*, 8898-8902.

35. Chambers, T. J.; Nestorowicz, A.; Amberg, S. M. Rice, C. M. *J. Virol.* **1993**, *67*, 6797-6807.

36. Helene, M.; Nicolas, M.; Barbara, S.; Jean, L. R.; Karine, A.; Jean, C. G.; Hughes, T.; Thai, L.; Subash, V.; Julien, L.; Bruno, C. The flavivirus polymerase as a target for drug discovery. *Antiviral Res.* **2008**, *80*, 23-35.

37. Dahai, L.; Ting, X.; Randall, P.; Watson, D.; Scherer-Becker, A. S.; Wolfgang, J.; Sui, S.; Yeong, C. H.; Wang, S. P;. Lim, A.; Strongin, S. G.; Julien, L Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J.* **2008**, *27*, 3209–3219.

38. Sahar, K.; Sonia, B.; Dimitrios, V; Anna-Claire, C.; Antonio, C.; Colin, B. Pieter, L.; Johan, N.; Andrea, B. Discovery of a novel HCV helicase inhibitor by a de novo drug design approach. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2935–2937.

39. Rich, D. H. *Comprehensive Medicinal Chemistry.* **1990**, *2* (Chapter 8), 391.

40. Leung, D.; Abbenante, G.; Fairlie, D. P. Protease inhibitors: current status and future prospects. *J. Med. Chem.* **2000**, *43* (3), 305-341.
41. Gibson, W.; Hall, M. R. T. An essential herpes virus proteinase. *Am. Chem. Soc.* **1996**, *211*, 240.
42. Bernstein, W. B.; Dennis, P. A. Repositioning HIV protease inhibitors as cancer therapeutics. *Cur. Opin. HIV AIDS.* **2008**, *3* (6), 666-75.
43. Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Res. Commun.* **1997**, *27* (2), 157-162.
44. Babine, R. E.; Bender, S. L. Molecular recognition of protein-ligand complexes. *Application to drug design, Chem. Rev.* **1997**, *97* (5), 1359-1472.
45. Donmienne, L.; Giovanni, A.; David, P. F. Protease inhibitors: current status and future prospects. *J. Med. Chem.* **2000**, *43* (3), 305- 341.
46. Li, H.; Clum, S.; You, S.; Ebner, K. E.; Padmanabhan, R. The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 converge within a region of 20 amino acids. *J virol.* **1999**, *73*, 3108-3116.
47. Gorbalenya, A. E.; Koonin, E. V. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic acid Res.* **1989**, *17*, 4713-4730.
48. Valle, R. P. C. Falgout, B. Mutagenesis of the NS3 protease of dengue virus type 2. *J. Virol.* **1998**, *72*, 624-632.
49. Krishna Murthy, H. M.; Clum, S.; Padmanabhan, R. Dengue virus NS3 protease, crystal structure and insights into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effects. *J. Biol. Chem.* **1999**, *274*, 5573-5580.

50. Kim, J.; Morgenstern, K. A. Crystal structure of the hepatitis c virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell.* **1996**, *87*, 343-355.

51. Falgout, B.; Miller, R. H.; Lai, C. J. Deletion analysis of dengue virus type 4 nonstructural protein NS2B: Identification of a domain required for NS2B-NS3 protease activity. *J virol*, **1993**, *67*, 2034-2042.

52. Yusof, R.; Clum, S.; Krishna Murthy, H. M.; Padmanabhan, R. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids. *J. Biol.* Chem. **2000**, *275*, 9963-9969.

53. Yan, Y. Li, Y.; Munshi, S. Sardana, V. Complex of NS3 protease and NS4A peptide of BK strain hepatitis c virus: a 2.2 A° resolution structure in a hexagonal crystal form. *Protein Sci.* **1998**, *7*, 837-847.

54. Chanprapaph, S.; Saparpakorn, P.; Sangma, C. competitive inhibition of the dengue virus NS3 serine protease by synthetic peptides representing polyprotein cleavage sites. *Biochem. Biophys. Res. Commun.* **2005**, *330*, 1237-1246.

55. Krishna Murthy, H. M.; Clum, S. Padmanabhan.. Dengue virus NS3 serine protease. *Biol. Chem.* **1999**, *274* (4), 5593-5580.

56. Wengler, G. G.; Czaya, P. M.; Hegemann, J. H. In vitro synthesis of West Nile virus proteins indicates that the amino-terminal segment of the NS3 protein contains the active centre of the protease which cleaves the viral polyprotein after multiple basic amino acids. *J. Gen.***1991**, *72*, 851–858.

57. Lobigs, M. Proteolytic processing of a Murray Valley encephalitis virus non-structural polyprotein segment containing the viral proteinase: accumulation of a NS3-4A precursor which requires mature NS3 for efficient processing. *J. Gen. Virol.* **1992**, *73*, 2305–2312.

58. Chambers, T. J.; Weir, A.; Grakoui, D. W.; McCourt, J. F.; Bazan, R. J.; Rice, C. Evidence that the N-terminal domain of non-structural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proc. Natl. Acad. Sci.* USA. **1990**, *87*, 8898–8902.

59. Rosaura, P. C. V.; Barry, F. Mutagenesis of the NS3 Protease of Dengue Virus Type 2. *J. Virol.* **1998**, *72* (2), 624–632.

60. Ikenaka, T.; Norioka, S. Bowman-Birk family serine proteinase inhibitors. In the proteases (Barrett, A. J. S & Salvesan, R., ed). *Academic press*, New York. **1986**

61. Kennedy, A. R. The Bowman-Birk inhibitor from soyabeans as an anticarciniogenic agent. *Am. J. Clin. Nutr.* **1998**, *68*, 1406-1412.

62. Bode, W.; Huber, R. Natural proteinase inhibitors and their interactions with proteinases. *Eur. J. Biochem.* **1992**, *91*, 433-451.

63. Perona, J.; Craik, C. S. Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **1995**, *4*, 337-360.

64. Bazan, J. F.; Fletterick, R. Detection of a trypsin- like serine protease domain in flaviviruses and pestiviruses. *Virology.* **1989**, *171*, 637-639.

65. Valle, R. P.; Falgout, B. Mutagenesis of the NS3 protease of dengue virus type 2. *J. Virol.* **1998**, *72*, 624-632.

66. Burley, S. K.; Petsko, G. A. Weakly polar interactions. *Advan. Protein Chem.* **1988**, *39*, 125-189.

67. Ryan, M. D.; Monaghan, S.; Flint, M. Virusencoded proteinases of Flaviviridae. *J. Gen. Virol.* **1998**, *79*, 947-959.

68. Zheng, Y.; Sejal, J.; Patel, W. W. Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. *Bioorg. Med. Chem. lett.* **2006**, *16*, 36-39.

69. Malancona, S.; Colarusso, S.; Ontoria, J. M. SAR and Pharmacokinetic studies on phenethylamide inhibitors of hepatitis C virus NS3/NS4A serine protease. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4575.

70. Zheng, Y.; Sejal, J. Peptide inhibitors of dengue virus NS3 protease. Part 1 : Warhead. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 36.

71. Tan, S. K.; Richard, P.; Rohana, Y.; Halijah, I.; Norzulaani, K.; Noorsaadah, A. Inhibitory activity of cyclohexenyl chalcone derivatives and flavanoids of fingerroot, Boesenbergia rotunda (L.) towards dengue-2 virus NS3 protease. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3337-3340.

72. Wu, J. Z., N. H.; Yao, M.; Hong, Z. Recent advances indiscovery and development of promising therapeutics against hepatitis C virus NS5B RNA-dependent RNA polymerase. *Mini Rev. Med. Chem.* **2005**, *5*, 1103–1112.

73. Wengler, G. Terminal sequences of the genome and replicative-form RNA of the flavivirus West Nile virus: absence of poly (A) and possible role in RNA replication. *Virology.* **1981**, *113*, 544–555.

74. Halstead, S. B. Dengue. *Curr. Opin. Infect. Dis.* **2002**, *15*, 471–476.

75. Erbel, P., N.; Schiering, A.; D'Arcy, M.; Renatus, M.; Kroemer, S. P.; Lim, Z.; Yin, T. H.; Keller, S. G; Hommel, U. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **2006**, *13*, 372–373.

76. Sampath, A., T.; Xu, A.; Chao, D. H.; Luo, J.; Lescar, S.; Vasudevan, G. Structure-based mutational analysis of the NS3 helicase from dengue virus. *J. Virol.* **2006**, *80*, 6686–6690.

77. Xu, T. A.; Sampath, A.; Chao, D. Y.; Wen, M.; Nanao, P.; Chene, S. G.; Vasudevan, J. Structure of the dengue virus helicase/ nucleoside triphosphatase catalytic domain at a resolution of 2.4 angstrom. *J. Virol.* **2005**, *79*, 10278–10288.

78. Bartholomeusz, A.; Thompson, P. Flaviviridae polymerase and RNA replication. *J. Viral. Hepat.* **1999**, *6*, 261–270.

79. Kapoor, M. L. W.; Zhang, M. Ramachandra, J.; Kusukawa, K. E.; Padmanabhan, R. Association between NS3 and NS5 proteins of dengue virus type-2 in the putative RNA replicase is linked to differential phosphorylation of NS5. *J. Biol. Chem.* **1995**, *270*, 19100–19106.

80. Nomaguchi, M. M. Ackermann, C.; Yon, S.; Padmanabhan, R. De novo synthesis of negative-strand RNA by dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. *J. Virol.* **2003**, *77*, 8831–8842.

81. You, S. B.; Falgout, L.; Padmanabhan, R. In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5_- and 3_-terminal regions that influence RNA structure. *J. Biol. Chem.* **2001**, *276*, 15581–15591.

82. Kohlstaedt, L. A. J.; Wang, J. M.; Friedman, P. A.; Rice, C.; Steitz, T. A. Crystal-structure at 3.5 angstrom resolution of HIV-1 reverse-transcriptase complexed with an inhibitor. *Science.* **1992**, *256*, 1783–1790.

83. Steitz, T. A. Structural biology—a mechanism for all polymerases. *Nature* **1998**, *391*, 231–232.

84. Thai, L. Y. Crystal structure of dengue virus RNA dependant RNA polymerase. *J. Virol.* **2007**, *4753*-4765.

85. Celiam, M.; Henry, C. CN Washington structure-based drug design. *Science & Technology pharmaceuticals.* **2001**, *79*, 23, 69-74.

86. Douglas, B.; Helen, D.; John, R. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discovery.* **2004**, *3*, 935-949.

87. Hélène, M.; Nicolas, M.; Barbara, S.; Jean-Louis, R.; Karine, A. The flavivirus polymerase as a target for drug discovery. *Antiviral Res.* **2008**, *80*, 23-35.

88. Kajohn, B.; Gabriele, K. Inhibitors of positive-sense (ss) RNA viruses RNA-dependant RNA polymerases. *Curr. Enzyme Inhib.* **2009**, *5*, 234-244.

89. Harki, D, A.; Graci, J, D. *Chembiochem.* **2007**, *8*, 1359-1362.

90. Crotty, S.; Magg, D.; Arnold, J. J. The broad spectrum antiviral ribo nucleoside Ribavirin is an RNA virus mutagen. *Nat. Med.* **2000**, *6*, 1375-1379.

91. Tsai, C.; Lee, P, Y.; Stollar, V. Antiviral therapy targeting viral polymerase. *Current. Pharm. Des.* **2006,** 12, 1339-1355.

92. Hashimoto, H.; Mizutani, K. A preparation of heterocyclic compounds as remedies for Hepatitis C. Int. Pat. Appl. WO01/47883. *Japan Tobacco Inc.* **2001**.

93. Dhanak, D.; Kaura, A.; Shaw, A. Novel anti-infectives. Int Pat. Appl. WO 01/85172. SmithKline Beecham Corporation, Int. **2001**.

94. Beaulieu, P. L.; Fazal, G.; Gillard, J. Viral polymerases inhibitors. Int Pat. Appl. WO 02/04425. Boehringer Ingelheim, (Canada) Ltd. **2002**.

95. Chan., I.; Reddy, T. J. Identification of N, N-disubstituted phenylalanines as a novel class of inhibitors of hepatitis C NS5B polymerase. *J. Med. Chem.* **2003,** *46*, 1283-1285.

96. Nittoli, T.; Curran, K.; Insaf, S. Identification of anthranilic acid derivatives as a novel class of allosteric inhibitors of hepatitis C NS5B polymerase. *J. Med. Chem.* **2007**, 3, 50 (9), 2108-2116.

97. Zheng, Y. N-sulfonylanthranilic acid derivatives as allosteric inhibitors of dengue viral RdRp. *J. Med. Chem.* **2009**, *52* (24), 7934-7937.

98. Peter, B. NTPase/Helicase of Flaviviridae: inhibitors and inhibition of the enzyme. *Acta. Biochim. Pol.* **2002**, *49* (3), 597-614

99. Ann, D.; Kwong, B.; Govinda, R.; Kuan-The, J. Viral and cellular RNA helicases as antiviral targets. *Nat. Rev. Drug Discovery.* **2005,** *4*, 845-853.

100. Dahai, L. Insights into RNA unwinding and ATP hydrolysis of the flavivirus NS3 protein. *EMBO J.* **2008**, *27*, 3209-3219.

Chapter 2

Molecular Modeling in Drug Design.

2- Molecular Modeling in Drug Design.

2.1- Introduction

The process of drug discovery is a complicated process that starts with an idea and could end with a new drug. It aims to develop a new molecule which could be potent, selective, and have good bioavailability. Furthermore, this process is very expensive consuming both time and money. Modern techniques such as computational chemistry and molecular modeling are widely used nowadays in order to save time and money. The main features of such tools are that they can be used for the understanding, studying and designing or modifying the compounds which could be potential inhibitors against the target protein. This is carried out by getting the main interactions between the compounds and the protein, ranking the compounds and taking the top scoring compounds into consideration and the excluding of compounds that do not show any interactions. Many successful drugs have been designed by such methods and have been proven to be very active against their target. Molecular modeling by definition is the science of representing molecular structures and simulating their behaviour.[1] Programs used for molecular modeling allow us to perform a number of operations which can be used in our research for drug design such as: the study of the target protein, energy minimization, docking studies, virtual screening, optimizing lead compounds, the design of de novo compounds, the study of the main interactions between compounds and protein, predicting the activities of unknown compounds, and studying structure-activity relationships (QSAR).

2.2- Computer-Aided Drug Design.

Computer-aided drug design methods are mainly dependent on the presence or absence of the crystal structure of a target protein. If the structure is available structure-based design will be the ideal method, and if it is not available and there is no lead compound homology modeling will be the alternative method. In a situation where there is no crystal structure but, a lead compound is available a ligand-based approach is the method of choice. Structure-based drug design is when the crystal structure of the protein is available and is obtained either by X-ray crystallography or by nuclear magnetic resonance (NMR). Here we can visualize the protein and its active site and observe/predict how the small compounds can interact with the active site residues. Different methods such as docking and virtual screening are common in structure-based drug design methods to obtain new compounds for specific proteins.

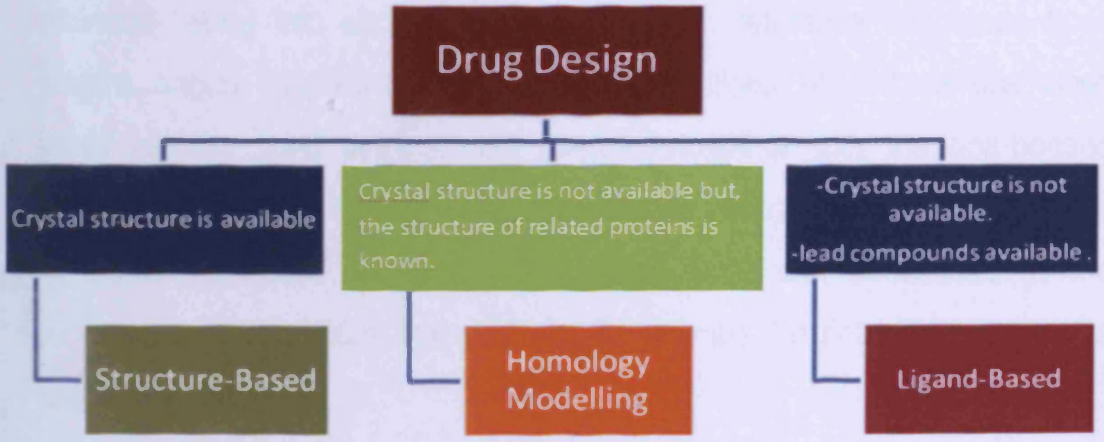


Figure 2.1 Schematic representation showing the different approaches in drug design.

The active site here is considered to be a space which needs to be filled with suitable molecule(s) or fragments that could be linked together to give the final molecule. The most important factors to be taken into account are the shape, size, charge, number of electron donating groups, number of electron accepting groups, and essential pharmacophores.[2] The following sections will summarize all the computational methods that were used in the present work showing the application of molecular modeling techniques in the process of drug design.

2.3-Molecular Mechanics (MM).

Molecular mechanics (MM) is the process used in order to calculate molecular geometries and energies. It considers the atoms as if they were rubber balls of different sizes connected together with springs (bonds). MM calculates the energy as a function of the nuclear position of atoms and this could be explained by the Born-Oppenheimer approximation which states that electrons move around the nuclei in a fast movement due to the difference between the mass of both nuclei and electrons.[3] MM can also calculate the total energy of a molecule taking into account its deviation from reference unstrained bond lengths, angles, torsions and non-bonded interactions. All of these unstrained items (angles, bond lengths, and torsions) together with the non-bonded interactions (Van der Waals and electrostatic interactions) and any force constants are called the force fields. Force fields can be defined as the functions and parameters that are used to describe the potential energy of a defined system.[4]

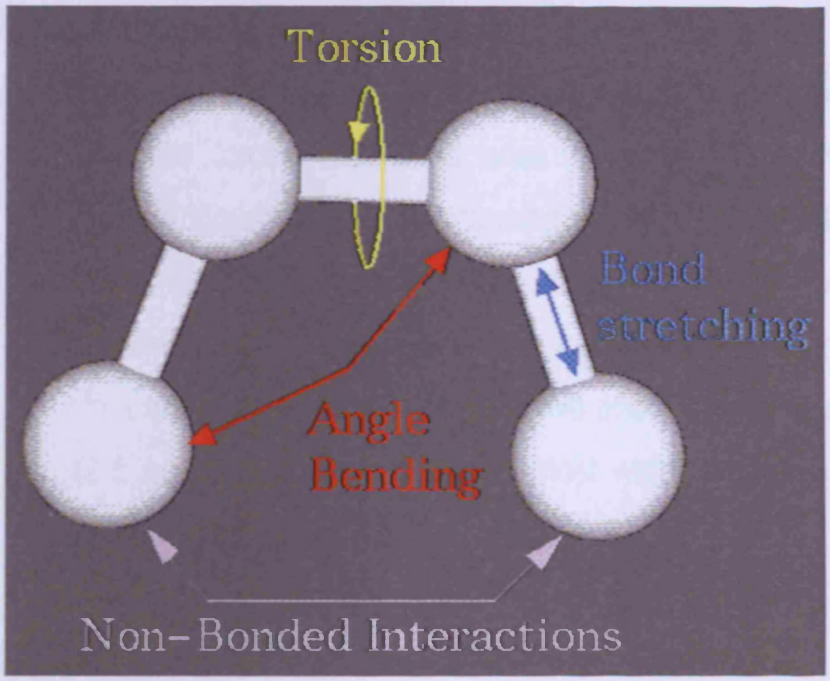


Figure 2.2 Type of bonds.

The potential energy can be calculated by the following equation:

$$E_{tot\ (MM)} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{elec} + ...$$

Where:-

- E_{tot} : is the total energy of a molecule.
- E_{str} : is the bond-stretching energy where the bond is treated as a spring according to Hooke's law in which the stretching of a spring is directly proportional to the load added to it and the potential energy will be proportional to the deviation from the equilibrium point.
- E_{bend} : is the angle-bending energy.
- E_{tors} : torsional energy produced by rotation of a bond.

- **Non-bonded interactions:**
  1. E_{vdw}: van der Waals energy which could be defined as the attractive forces that could arise from the distortion of the electron clouds and hence causing dipoles that by its role will induce dipoles in the neighboring atoms and finally result in an attractive effect. If the atoms become very close their electron clouds will overlap causing a repulsive effect. The Lennard-Jones 12-6 equation is considered the most widely used to calculate Van der Waals potential.[5-6]

$$E = 4\varepsilon [\sigma/r]^{12} - [\sigma/r]^{6}$$

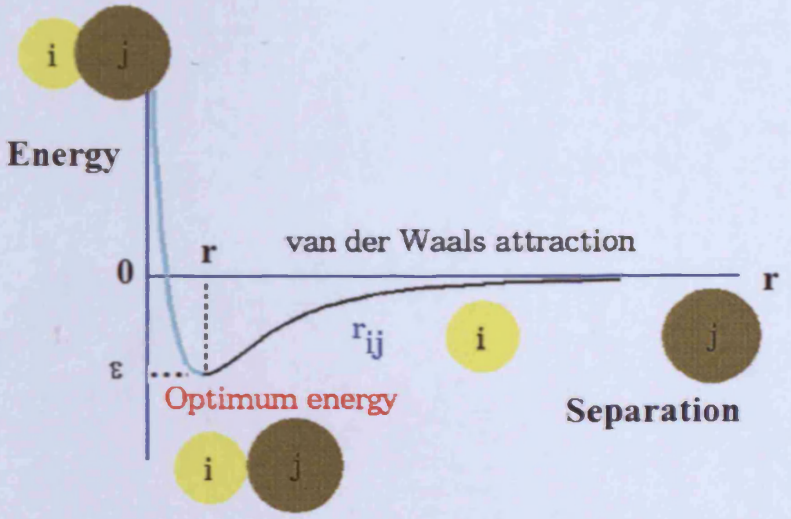


Figure 2.3 Lennard-Jones potential (12-6) function.[7]

Where:

$-r^{-12}$: is the repulsive component.

$-r^{-6}$: is the attractive component.

r: the distance between the interacting atoms.

σ: the distance from which the interaction energy is zero.

ε: minimum energy when the distance between the nuclei is the sum of their Van der Waals radii.

2. E_{elec}: electrostatic energy that is resulted from electrostatic interactions and could be expressed by Coulomb's law. It could be calculated after assignment of partial charges to all atoms because of the unequal distribution of charge in the molecules.

$$E_{elec} = \frac{1}{4\pi\varepsilon} - \frac{q1q2}{r^2}$$

Where:

r: The distance between two atoms of partial charges q1 and q2.

ε: Permittivity of the medium.

2.4- Force Fields.

Force fields (FF) are mathematical equations that provide some parameters which are important in solving certain problems. By definition it is the functions and parameters that are used to describe the potential energy of a defined system. There are several FF that have been developed to deal with different types of atomic coordinates. FF can compute bonded and non-bonded interaction energies.[8] We can find FF that are more suitable for some specific systems rather than other systems. So, the choice of the proper FF will mainly depend on the system to be used. For example, the **AMBER** FF is suitable for proteins and nucleic acids and not used for small organic molecules.[9] While **MMFF** (Merck Molecular force field) is widely used with small organic molecules.[10]

2.5- Energy Minimisation (EM).

In the process of MM the total energy is minimized with respect to the atomic coordinates.[4] Molecules can have different conformations due to rotations about single bonds. The conformations of lower energy are considered the most stable. So, the process of EM is mainly used to search for the optimum conformations with low energy using EM algorithms.[3] The process of EM involves adjustments in the geometry of the molecules which are followed by energy calculations and the discarding of any geometry with high energy. Once the minimum energy is reached the process is stopped. The stable state of any molecule is known as local minima which is the minimum point in the energy. Flexible molecules will have more local minima than rigid molecules.[5-6] EM algorithms are classified into two groups: non-derivative methods (e.g. simplex method) that do not require calculations of potential energy and derivative methods (e.g. steepest descent, conjugate gradient, and Newton-Raphson).[4-6]

EM is widely used in molecular modeling applications. It can be used with MM, molecular dynamics (MD), molecular docking, homology modeling (HM) and also, with conformational search methods such as systematic and stochastic methods to generate the starting point for the minimisation run.

2.6- Conformational Analysis

Molecular conformations are commonly defined as structures that can be interconverted by rotation about single bonds. The conformations available to a molecule can have a great effect on its biological activity and in some cases one of these conformations may be responsible for an observed behaviour. So, when we are going to analyze the effects of a 3D structure on molecular properties, we should consider all possible conformations. Some aspects should be taken into account such as, the conformational space which can be represented schematically by the following plot.

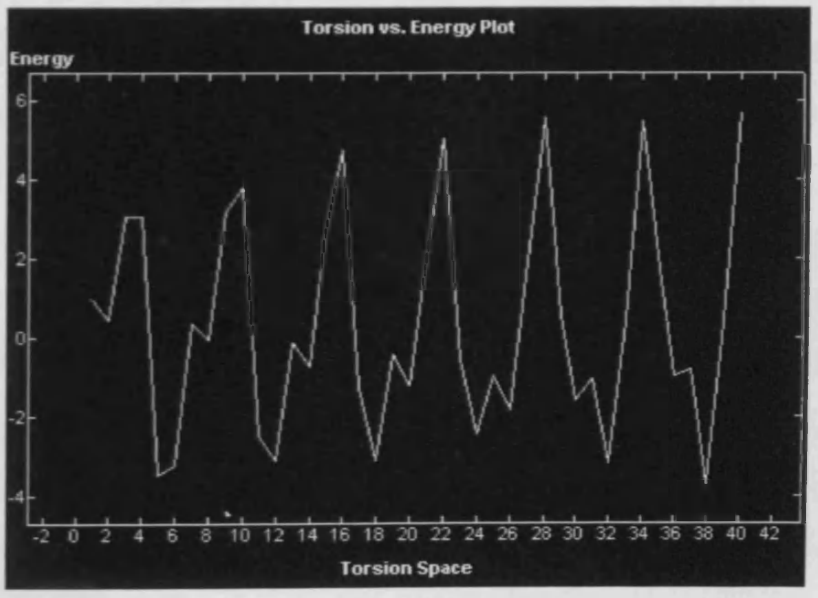


Figure 2.6 energy versus torsional-angle plot where the x-axis represents the torsion space and the y-axis represents energy. [11]

The minima in the above conformational plot correspond to low energy conformations, while the maxima represent transition states between low energy conformations. The lowest energy conformer is referred to as the global minimum and lies at the bottom of the deepest potential energy well. Conformers that lie at the bottom of potential wells with energies higher than the global minimum energy are referred to as local minima. A complete conformational search via incremental rotations about all rotatable bonds often leads to a combinatorial explosion of structures that becomes impossible to enumerate. For example, a systematic 0-360 degree search of each rotatable bond in a structure containing 6 rotatable bonds would produce more than 2 billion structures presuming that each angle is searched in 10-degree increments. Many more structures would result if the increment was lowered. [11]

2.6.1- Systematic Conformational Search

The main aim of the Systematic Conformational Search is to generate a collection of reasonable molecular conformations which may or may not be at local minima. This is by systematically rotating bonds in a molecule by discrete increments. A generated conformation is rejected if it contains two atoms whose mutual van der Waals energy exceeds a threshold (by default, 10 kcal/mol). This ensures that the output conformations contain no conformations with heavily overlapped atoms. The main advantage of the systematic search is that its output will contain most of the local minima. The systematic search also produces reasonable molecular geometries that do not lie at potential energy minima. Such conformers can be important in situations where the minimum energy conformation of a bound structure (*e.g.*, a ligand docked to a protein) does not correspond to a minimum energy conformation of the unbound structure (*i.e.*, gas phase or free in solution). In any molecule, all bonds except

bonds to terminal atoms are available for rotation. Such bonds are called rotation bonds. For each rotation bond, a list of possible relative dihedral increments, or steps, is determined. For example, the CH3-OH rotation bond might be assigned to rotate three times (*e.g.*, 0, 120 and 240 degrees). Once the list of increments has been determined, the application generates all combinations of conformations according to the rules laid out in the increment list. Conformations with severe non-bonded contacts are eliminated from the set of generated conformers. The rotation increments are added to the dihedral angle of the starting conformation. For this reason, it is important to perform an energy minimization prior to running the Systematic Conformational Search.[12]

2.7- Molecular Alignment.

There are many different procedures proposed for aligning molecules. Different methods work better in different situations and there is no best method for all cases. There are three categories; substructure overlap, pharmacophore overlap and docking. [5] Substructural overlap is the simplest. It assumes that the molecules share a common core of atoms. This core is overlapped in each of the molecules in the dataset. On the other hand, pharmacophore overlap does not assume any particular common core for the active molecules. Instead, it assumes that the pharmacophore features involved in binding are identified (or can be identified) and attempts to maximize the overlap of these features between the molecules in the dataset. Docking is a third strategy that is based on the idea that if you know how the molecules can bind to the active site, you can use those coordinates for the alignment. The bound conformation is not necessarily to be the best conformation.

2.8- Site Finder.

The aim of Site Finder is to calculate possible active sites in a receptor from the 3D atomic coordinates of the receptor. This is to determine potential sites for ligand binding in docking calculations or Multifragment Search. Active sites are usually hydrophobic pockets that involve side chain atoms. There are some methods that use interaction energies between the receptor and different probes in an attempt to locate energetically favorable sites. However, this procedure requires the assignment of proton locations and partial charges to the receptor atoms and this is not always easy. While van der Waals energies can indicate sterically available regions, the long-range nature of electrostatic potentials makes the interpretation of energy levels difficult (e.g., a carboxylate in an active site will emphasize positively charged probes even though negatively charged probes like carbonyl oxygen may be part of the bound ligand). Alternatively, purely geometric methods seek to locate "pockets" without the use of energy models. This is advantageous since proton locations are not required. Another method uses a grid representation of the molecular volume and computes exterior site scores by projecting rays from the receptor exterior to the surface. The deeper and more surrounded a site is, the higher it scores. The Site Finder falls into the category of geometric methods since no energy models are used. Instead, the relative positions and accessibility of the receptor atoms are considered along with a rough classification of the chemical type.

The following methodology principles have been used:

a) Identify regions of tight atomic packing. This is not the same as locating pockets, since surface sites may still be regions of tight packing.
b) Filter out sites that are "too exposed" to solvent. In other words, sites that are on protrusions are unlikely to be good active sites.
c) Use hydrophobic/hydrophilic classifications. This coarse classification of chemical type is used to separate water sites from the more likely hydrophobic sites.
d) Use a definition of hydrophilic that is invariant to protonation state and tautomer state (this means no distinction between donor and acceptor atoms).
e) Avoid grid-based methods since grid methods are not invariant to the rotation of the atomic coordinates and can consume large amounts of memory.

The Site Finder methodology is based upon Alpha Shapes which are a collection of 3D points in triangulated form. For each collection of four points there is an associated sphere called an alpha sphere. These spheres have differing radii including infinite radii (corresponding to the planes of the convex hull of the point set). The collection of alpha spheres is pruned by eliminating those corresponding to inaccessible regions of the receptor as well as those that are too exposed to solvent. In addition, only the small alpha spheres are retained since these correspond to locations of tight atomic packing in the receptor. Also, each alpha sphere is classified as either "hydrophobic" or "hydrophilic" depending on whether the sphere is in a good hydrogen bonding spot in the receptor. Hydrophilic spheres not near a hydrophobic sphere are eliminated (since these generally correspond to water sites). Finally, the alpha

spheres are clustered using a single-linkage clustering algorithm to produce a collection of sites. Each site consists of one or more alpha spheres at least one of which is hydrophobic. [13-18]

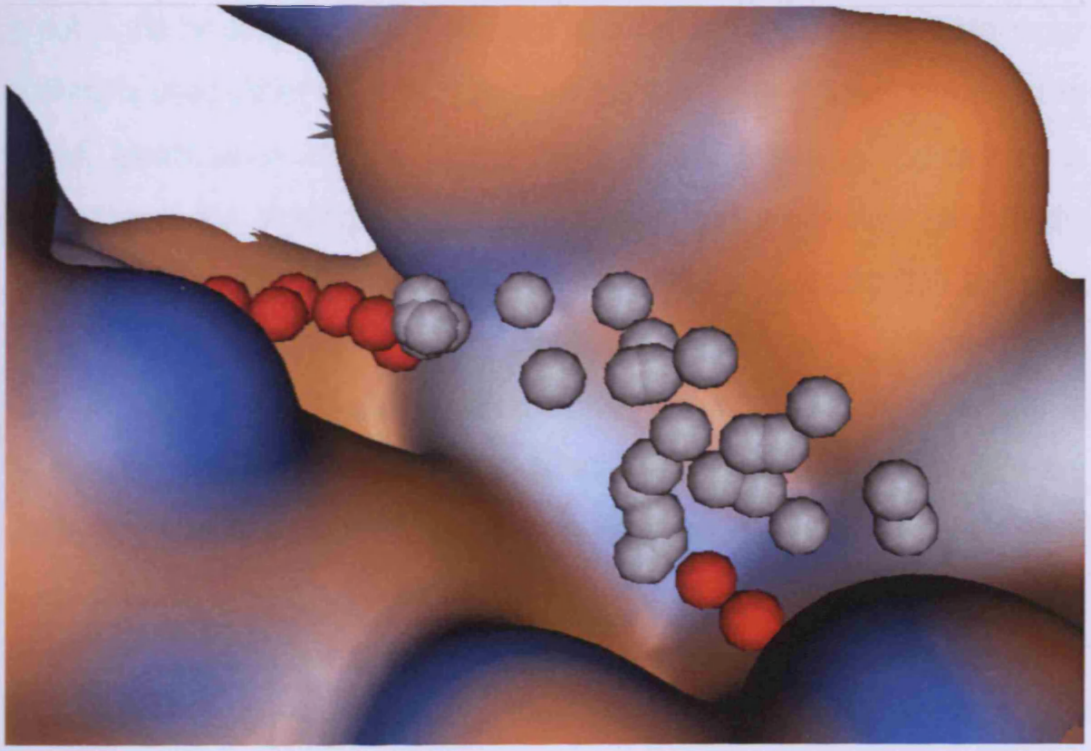


Figure 2.7 This represents a group of alpha spheres indicating an allosteric site for inhibitor design.

2.9- Molecular Docking.

Molecular docking simulations represent a widely employed computational tool in drug discovery. It attempts to predict the structure of intermolecular complex between the ligand and macromolecule (target protein). The first docking attempts were done manually using interactive computer modelling. The ligand is put in the binding site and minimised to avoid bad steric clashes. Molecular docking is used either for prediction of the binding mode of a well known active ligand, identification of new ligands using virtual screening approaches or predicting of the binding affinities of a related series of active compounds. Docking procedures are classified depending on the approximation level into three categories; rigid body docking in which both protein and ligand are treated as rigid bodies, semi-flexible docking where the ligand only is considered a flexible and fully flexible docking in which both the ligand and protein are treated as flexible molecules. A standard docking protocol consists of a step-wise process. First, a proper search algorithm predicts the various configurations of the ligand (poses) within the target binding site. In the second step, each docked pose is evaluated and ranked assessing the intermolecular interaction and estimation of the binding free-energy. The ability of a standard docking protocol to achieve its ultimate goal provides a reliable binding mode. Prediction strongly depends on the accuracy of the scoring function used. Molecular docking involves many degrees of freedom. Of the well known six degrees of freedom of translational and rotational freedom are of one molecule relative to the other. It also involves, the conformational degrees of freedom of the ligand and the protein. In order to perform docking for a large number of compounds the degrees of freedom should be fast.[7]

2.9.1- Docking Algorithms.

Docking algorithms are used by the computational programs to deal with the flexibility of the ligands. The accuracy of predicting the ligand orientation in molecular docking is mainly due to a good docking algorithm. There are three methods of docking algorithms; systematic methods, stochastic methods, and simulation methods.[19]

2.9.1.1- Systematic Methods.

(Incremental construction and multiconformer database.)

These methods explore all degrees of freedom of conformations due to the systematic variation of all torsional angles in the molecule and this can lead to a large number of conformations. To overcome this an incremental search in which the molecule is split into different rigid fragments then linked them together in an incremental construction manner or by splitting the molecule into two parts; rigid part (core) and flexible part (side chains). Here the core is docked first then an incremental construction of the flexible parts takes place.[20-22]

2.9.1.2- Stochastic Methods.

(Monte Carlo, Genetic algorithm and Tabu search).

In stochastic methods a random changes are made to the ligand in order to find a new one that can be evaluated and the process is repeated until, it resulted in a number of conformations.[23]

2.9.1.3- Simulations Methods.

(Molecular Dynamics MD and EM)

Molecular Dynamics MD simulations are used to simulate the different parts of a protein-ligand complex at different temperatures.[24-25] They can also bring the ligand in local minima.

2.10- Scoring Functions (SF).

SF are mathematical methods used by molecular modeling programs in order to predict the strength of non-bonded interactions and the binding affinities of potential ligands to their proposed site of action. Some SF is dealing with hydrogen-bond energies or entropic loss which occurs to ligands due to solvation effects or the binding of ligands to their site. Scoring functions are used to evaluate and rank the resulting docked poses. This creates a correlation between the atomic coordinates and energy values. Poses of low energy are considered to have a higher chance for binding than those with high energy. Both FF and SF are similar in that they are mathematical functions but, SF are considered to be the speed determining point in a docking algorithm. In the process of molecular docking molecular algorithms are capable of generating a large number of conformations. Some of these conformations will be rejected due to their high energy clash with protein. The conformations of low energy should be assessed by using some SF which will be able to identify and rank the docked orientation for all docked ligands or conformations. There are four types of SF; Force Field-based SF, Empirical SF, Knowledge-based SF, and Consensus SF.

2.10.1- Force Field-Based SF.

Here the SF mainly depends on MMech FF which calculates the energies of both the receptor-ligand interactions energy and the ligand energy which could be due to steric strain resulting from binding. Because it deals with MMech FF it will use the Van der Waals (12-6 Lennard-Jones) and electrostatic (Coulomb's law) energy equation. FF-based SF have some limitations due to their consideration of a single protein conformation and because they were developed for gas-phase in addition to their high cost.[26]

2.10.2- Empirical SF.

Empirical SF is mainly used to produce the binding energies which are correlated with experimental affinity. Here a training set is used and the evaluated binding energies of the complex structures of this set are used in a regression analysis which will require the structures and the binding constants. The SF use hydrogen bonds, electrostatic and hydrophobic interactions. The main disadvantage of this type is that it depends mainly on the training set.[27-31]

2.10.3- Knowledge-Based SF.

This type of SF is used to generate experimental structures and does not deal with binding energy. The modeling of a receptor-ligand complex will depend on atomic interactions so; a number of atom-type interactions will be defined according to molecular environment. It is considered to be a simple method that could screen a large number of databases. The main disadvantage of this kind of SF is that it strongly depends on the data encoded in limited sites of protein-ligand complexes. Any interactions that are commonly used are considered to be attractive. However, the less frequently interactions are considered to be repulsive.[26]

2.10.4- Consensus SF.

It uses information from different scores to balance errors in single scores and to be able to improve and identify true ligands.[26, 31] An example of that is FlexX SF. If terms in different SF are significantly correlated the value of consensus scoring might be limited and it could amplify calculation errors and will not be able to balance them.

2.11- Pharmacophore Search.

The main purpose of a pharmacophore search is to perform 3D searches of conformation databases using the annotations that are related to ligand receptor interactions. Here the pharmacophore can be defined as the group of structural features that are related to the ligand's activity and recognition at the receptor site. In the MOE program the pharmacophoric structural features are represented by labeled points in space. There is a set of points for each conformation that could be considered as a set of structural features that may contribute to the pharmacophore of that ligand. By this it is possible to search a database of conformations with a pharmacophore query and the purpose of the use of that query is to select a limited subset from a large number of conformations. The output will have only the conformations that are matched with the pharmacophoric features of the query.

2.12- Structure-Based Virtual Screening.

Drug discovery has traditionally been successful by a combination of random screening and rational design. Nowadays, and by the huge progress in drug discovery approaches, new methods have been developed. One of these approaches which are of great importance in drug design is virtual screening. It can be named as in silico screening, attracting increasing levels of interests in the field of drug design. By definition, virtual screening is a computational approach used for the evaluation of binding properties using large ligand databases that may be fragment-like, lead-like, drug-like and can be used in rational drug design. The main principle of this process involves the computational analysis of chemical databases to identify compounds appropriate for a given biological receptor. This strategy implies that some information is available regarding either the nature of the receptor binding site

or the type of the ligand that is expected to bind effectively to the receptor. An important class of this approach is based on virtual screening using a pharmacophore model that is derived from a known active Ligand. In this case a pharmacophore query is generated depending on the structure features of the reported active compound, taking in account the whole essential pharmacophoric features of this compound such as hydrogen bond acceptors, and hydrogen bond donors. A second class is docking and scoring techniques that predict the positions of bound Ligands and related binding affinities in case we have a 3D crystal structure of the target protein. Virtual screening can provide us with a potential hits that may be considered as a good starting point in drug discovery. In order to perform a successful virtual screening we should have a crystal structure of the target protein complexed with an inhibitor or to have data about the most active inhibitors for this target beside all the data about the active site/s of this protein, databases that can be used for the screening. The pharmacophore query that will be built depends on the active inhibitor by the aid of a molecular modeling program that will be also used also for docking and filtering of hits depending on their scoring. Finally, synthesis and measurement of the inhibitory activity will be the last step after the selection of the final hits. In comparison to in-vitro high-throughput screening, virtual screening technique is a faster and less expensive method for the selection of compounds that could be tested as it is if commercially available and can be synthesized or after the modification and removal of the undesired side chains that do not contribute to the built query and can affect the docking results.

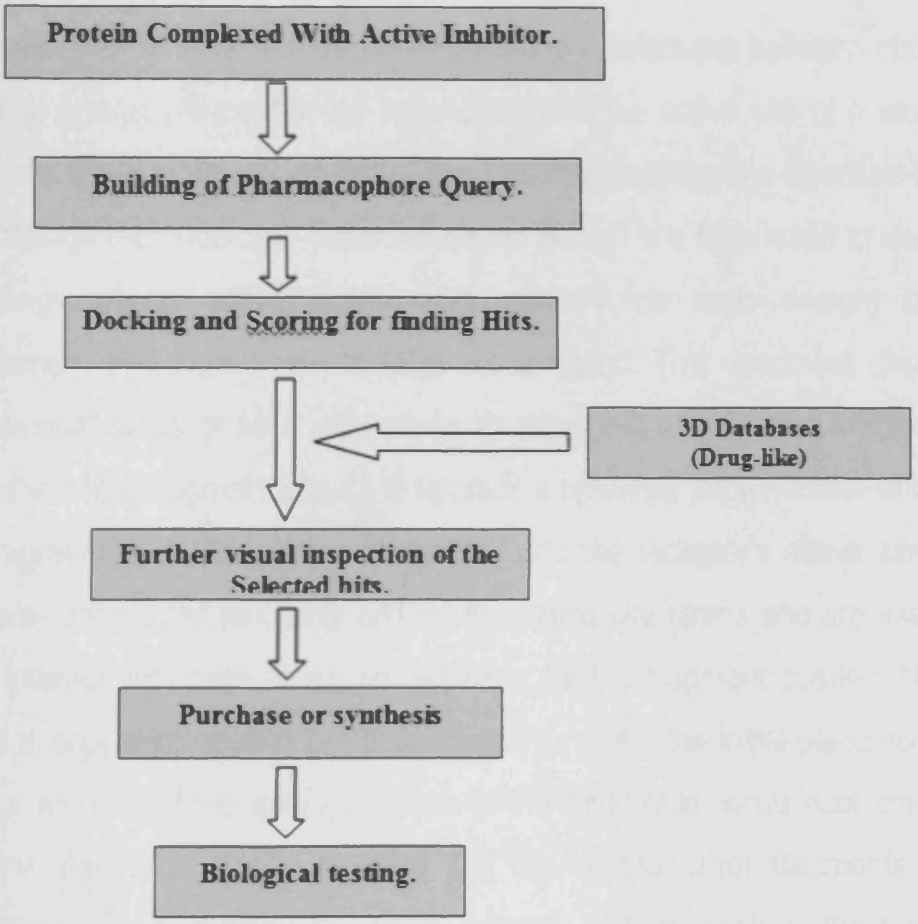


Figure 2.4 Schematic representation of the key steps of The virtual screening process.

2.13- Multifragment search (MFS).

The purpose of MFS is to help understand the interactions between chemical functional groups (referred to as fragments) with the active site of a receptor. MFS uses a 3D structural model of the receptor, making it a structure-based ligand design methodology. Structure-based design is a term used to describe the design of an active ligand from scratch (*de novo* design) or the improvement of a known (structurally) active ligand. This document describes the individual steps of MFS as well as its input and output. The fundamental idea behind Multifragment Search is to place a relatively large number of copies of a fragment, say 200 copies of ethane, into the receptor's active site. The fragments are placed randomly around the active site atoms and are assumed not to interact with each other; no regard is paid to fragment overlap. Next, a special energy minimization protocol is used to refine the initial placement: the receptor atoms feel the *average forces* of the fragments, while each fragment feels the *full force* of the receptor but not of the other fragments. After preparation, the calculation begins in earnest. In turn, each of the fragment classes (*i.e.*, all chemically identical fragments) are energy minimized; the other fragments and the receptor are held fixed. During this energy minimization, fragments interact only with the receptor and not with any other fragments. This is followed by the energy minimization of the receptor atoms with the fragment atoms held fixed. The receptor is subjected to the mean force of the fragments. This energy minimization protocol is repeated until convergence within a user-defined RMS gradient.[32]

2.14- Homology Modelling (HM).

The ultimate goal of protein modeling is to predict a structure from its primary sequence with an accuracy that is comparable to the best results achieved experimentally.[33] This would be a good chance to use rapidly generated *in silico* protein models in the case that the crystal structure of the target protein is not available. Many proteins are simply too large for NMR analysis and cannot be crystallized for X-ray diffraction. In practice, homology modeling is a multi-step process that can be summarized in seven steps; the first step is to identify the best template depending on the percent of identity between it and the query sequence. Programs such as BLAST or FASTA are used for such purpose.[34-35] Alignment will be performed in the next step to determine the regions of low identity where the alignment will be difficult and may need correction. Then the coordinates of the template residues will be copied in order to build the backbone for the new model. The alignment between the model and template sequences can have gaps, which is due to gaps in the model sequence deletions or in the template sequence insertions. Both cases imply a conformational change of the backbone. These conformational changes cannot happen within regular secondary structure elements. It is therefore safe to shift all insertions or deletions in the alignment out of helices and strands, placing them in loops and turns which is done in the loop modeling step. When we have conserved residues in both the template and query it will be easier to copy these conserved residues entirely from the template to the model and achieve a higher accuracy than by copying just the backbone and repredicting from the side chains takes place. The next step will be the model optimisation to improve the models. The most straightforward approach to model optimization is simply to run a molecular dynamics simulation of the model. Finally, model validation takes place to make sure that the final model can be used.

2.15- Molecular Dynamics (MD).

MD is a powerful and widely used tool in chemistry, physics, and materials science. This technique is a scheme for the simulation of the natural motion of molecules and allows the prediction of the static and dynamic properties of substances directly from the underlying interactions between the molecules. It helps us to study the dynamic and thermodynamic properties by numerically solving an equation of motion, which is the formulation of the rules that govern the motion executed by the molecule. It applies the Newton's second law of classical mechanics which states that atoms in a molecule interact with each other according to the rules of the employed force field at regular time intervals.

$$Fi\ (t) = mi\ ai\ (t)$$

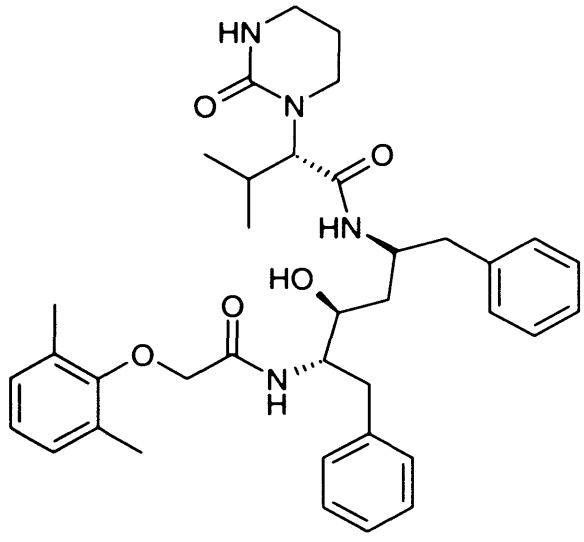
Where:

- Fi (t) = The force (F) of atom (i) at time (t).
- Mi = The mass of atom i.
- ai = The acceleration of atom i at time (t).

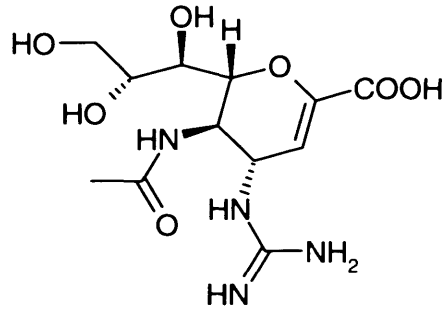
It is mainly concerned with the atomic interactions, geometries, and energies in a given moment. MD predicts the temporal behaviour of a molecular system over time. MD is a good method for the generation of configurations for proteins and large molecules that can not be simulated by quantum mechanics due to their large size. Also, it provides a suitable refinement for models that were built by HM and alignment studies in order to adjust their coordinates.[4]

2.16- Aim of work and objectives.

Structure-based drug design methods utilize the knowledge of a three dimensional structure of an enzyme/receptor to develop small molecules able to bind to the desired target, generating a specific biological response. These computer-based methodologies are now becoming an integral part of the drug discovery process and, although the principles of molecular recognition are far from being completely understood, some marketed compounds (i.e. Zanamivir, Lopinavir) have been developed with the help of the successful application of structure-based design techniques.



Lopinavir
HIV protease inhibitor



Zanamivir
Antiinfluenza

Figure 2.5 Structure of real drugs developed by structure-based approach.

- The first aim of this work is to modify the structure of Panduratin A (derived from BR extract) that has an activity = 25 μM depending on molecular modeling and docking results into the **S1** pocket of Dengue virus NS3 protease. That is done by removal of undesired substituents such as 2-(3-methyl-3-butenyl) and 6-phenyl moieties.
- The second aim of our project is to design and synthesize small non-peptidic molecules that could be potential inhibitors for **DV-2** NS3 protease using the virtual screening approach.
- Design and synthesis of novel Dengue RdRp allosteric inhibitors by fragment-based drug discovery methods and depending upon the presence of a 3' GTP inhibitor complexed with Dengue virus RdRp. The aim is to find a suitable fragments that could be linked together to reach to a potential inhibitor.

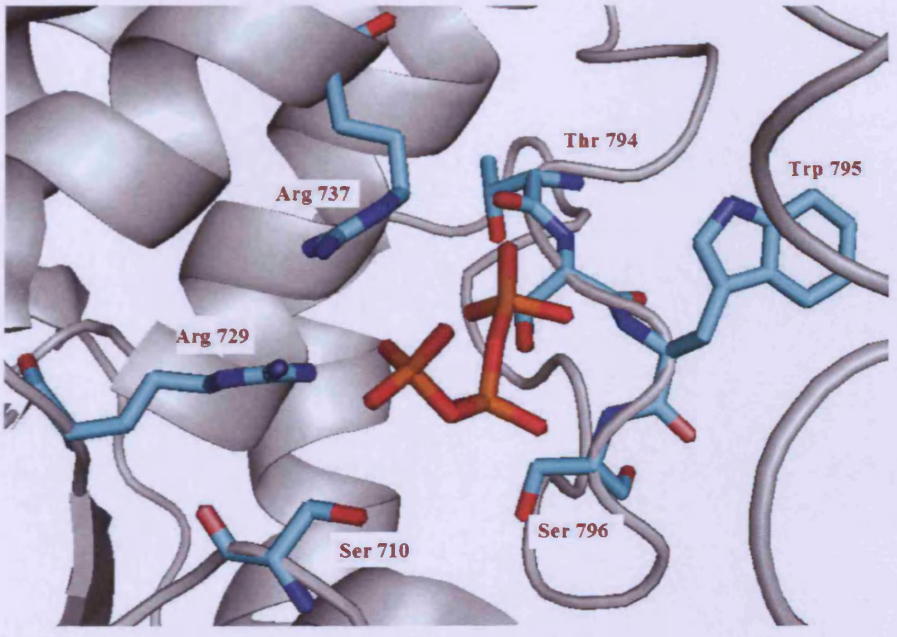


Figure 2.6 Potential site for allosteric RdRp inhibitor design.

- Docking studies of well known Flaviviral Helicase inhibitors against HCV, Dengue and WNV helicase which required the building of a homology model for WNV helicase in order to do a comparative docking study that could result in some aspects important in the design of novel dengue virus helicase inhibitors.

2.17- References.

1. http://www.netsci.org/science/compchem/feature01.html
2. Celiam, M.; Henry, C. CN Washington structure-based drug design. Science & Technology pharmaceuticals, June 4, **2001**, 79, 23, 69-74.
3. Jao, P. A. N. PhD thesis, Faculty of medical and human sciences, University of Manchester. **2006**.
4. Hans-Dieter, H.; Gerd, F. Molecular modeling: basic principles and applications. WILEY VCH.
5. Leach, A. R. Molecular Modeling: Principles and Applications; Prentice Hall: Harlow, **2001**.
6. Goodman, J. M. Chemical applications of molecular modeling; The Royal society of chemistry: Cambridge **1998**.
7. Krumarine, J.; Raubacher, F.; Brooijmans, N.; Kuntz, I. Principles and methods of docking and ligand design. Structural bioinformatics. Bourne P. and Weissig H. Eds.; John Wiley & Sons Inc.: Hoboken, NJ. **2003**, 443-476.
8. Kandil, S. Computer-aided drug design, PhD thesis, Welsh school of pharmacy **2009.**
9. Cornell, W.; Cieplak, P.; Bayly, C.; Gould, I.; Merz, K.; Ferguson, D.; Spellmeyer, D.; Fox, T.; Caldwell, J.; Kollman, P. A second generation force-field for the simulation of proteins, nucleic acids and organic molecules. *J. Amer. Chem. Soc.* **1995**, *117*, 5179-5197.
10. Halgren, T. Merck molecular force field. 1. Force field form, scope, parameterization and performance of MMFF94. *J. Comp. Chem.* **1996**, *17*, 490-519.
11. Molecular operating Environment (MOE). Chemical computing group. Inc. Montreal, Quebec, Canada.

12. Ferguson, D. M.; Raber, D. J. A New Approach to Probing Conformational Space with Molecular Mechanics: Random Incremental Pulse Search. J. Am. Chem. Soc. **1989,** *111*, 4371-4378.
13. Edelsbrunner, H. Weighted Alpha Shapes; *Technical Paper of the Department of Computer Science of the University of Illinois at Urbana-Champaign*; Urbana, Illinois 61810.
14. Del Carpio, C. A.; Takahashi, Y.; Sasaki, S. A New Approach to the Automatic Identification of Candidates for Ligand Receptor Sites in Proteins: (I) Search for Pocket Regions. *J. Mol. Graph.* **1993,** *11*, 23–42.
15. Edelsbrunner, H.; Facello, M.; Fu, R.; Liang, J.; Measuring Proteins and Voids in Proteins; *Proceedings of the* 28th Hawaii International Conference on Systems Science. **1995**, 256–264.
16. Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985,** *28,* 849–856.
17. Hendlich, M.; Rippman, F.; Barnickel, G. LIGSITE: Automatic and Efficient Detection of Potential Small Molecule-binding Sites in Proteins. *J. Mol. Graph.* **1997,** *15,* 359–363.
18. Miranker, A.; Karplus, M. Functionality Maps of Binding Sites: A Multiple Copy Simultaneous Search Method. *Proteins.* **1991,** *11*, 29-34.
19. Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput. –Aided Mol. Des.* **2002,** *16*, 151-166
20. DesJariais, R. Docking flexible ligands to macromolecular receptors by shape. *J. Med. Chem.* **1986,** *29*, 2149-2153.
21. Klebe, G.; Rarey M. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996,** *261*, 470-489.

22. Kuntz, I.; Leach, A. conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comput. Chem.* **1992,** *13*, 730-748.

23. Olson, A.; Goodsell, D. Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins.* **1993**, *17*, 1-10.

24. Brooijmans, N.; Kuntz, I. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biolmol. Struct.* **2003,** *32*, 335-373.

25. Di Nola, A.; Berendsen, H.; Roccatano, D. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins.* **1994,** *19*, 174-182.

26. Kitchen, D.; Decornez, H.; Furr, J.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature.* **2004,** *3*, 935-949.

27. Clark, D. et al. PRO Ligand: an approach to de novo molecular design. 1. Application to the design of organic molecules. *J. Comput. Aided. Mol. Des.* **1995,** *9*, 13-32.

28. Murray, C. et al. PRO-SELECT: combining structure based drug design and combinatorial chemistry for rapid lead discovery. Technology. *J. Comp. Aided. Mol. Des.* **1997,** *11*, 193-207.

29. Bohacek, R.; McMartin, C. Multiple high diverse structures complementary to enzyme binding sites: results of extensive application of de novo design method incorporating combinatorial growth. *J. Am. Chem. Soc.* **1994,** *116*, 5560-5571.

30. Pearlman, D.; Murcko, M. CONCEPTS: new dynamic algorithm for de novo design suggestion. *J. Comput. Chem.* **1993,** *14,* 1184-1193.

31. Eldridge, M.; Murray, C.; Auton, T.; Paolini, G.; Mee, R. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997,** *11*, 425-445.

32. Miranker, A.; Karplus, M. Functionality Maps of Binding Sites: a Multiple Copy Simultaneous Search Method; *Proteins: Struct., Funct., and Genet.* **1991,** *11*, 29–34.

33. Elmar, K.; Sander, B.; Nabuurs.; Gert, V. Structural Bioinformatics: Homology modelling. Edited by Philip E. Bourne and Helge Weissig. Wiley- Liss Inc.

34. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990,** *215*, 403–10

35. Pearson, W. R. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. **1990,** 183, 63–98.

Chapter 3

Docking and modification of Panduratin A and 4-Hydroxypanduratin A toward the inhibition of Dengue NS3 protease.

Chapter 3

3.1- Importance of S1 pocket.

According to the literature, [1] the **S1** pocket could be used for the development of small molecule that could inhibit the DV NS3/NS2B protease. The **S1** pocket could be used alone instead of depending on all the pockets that may require a large peptide inhibitor rather than a small molecule inhibitor.[2]

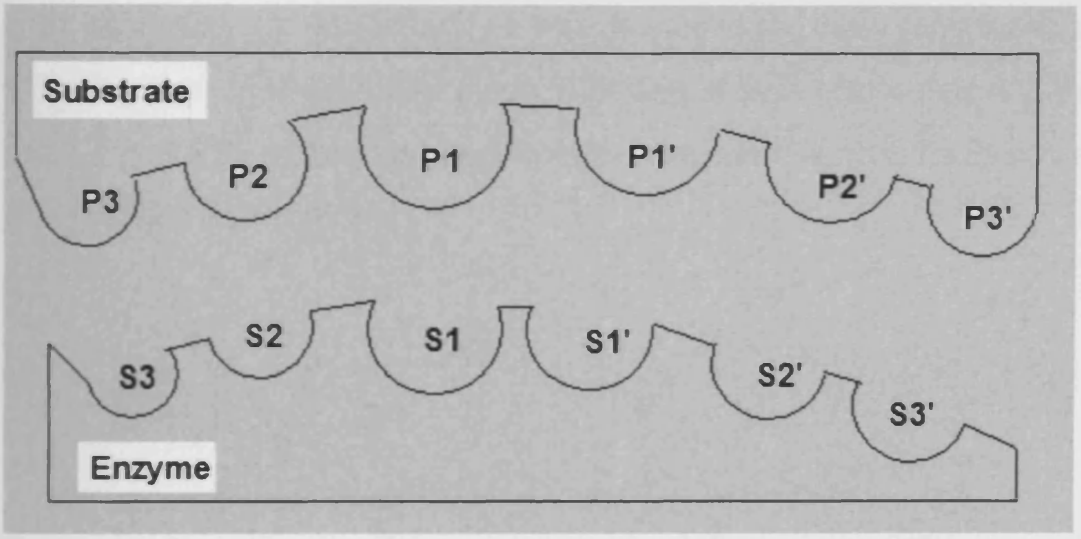
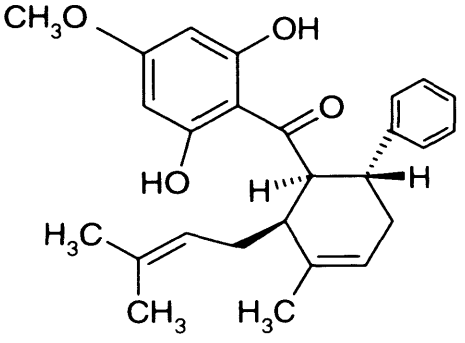


Figure 3.1 Schematic representation of substrate side chains and their fitting in the NS3 protease pockets

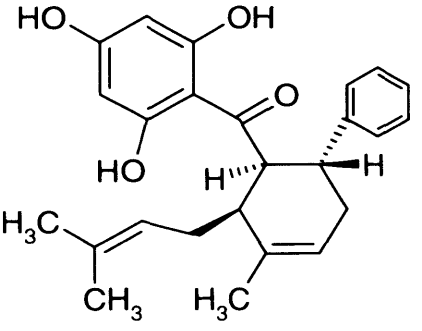
3.2- Docking of Panduratin A and 4-Hydroxypanduratin A.

A Fluorogenic peptide substrate containg Arg-Arg residue was used in order to test the cleavage of Arg-Arg at the **S1** pocket. Dengue 2 protease with or without BR extracts compounds of different concentrations were buffered at pH 8.5 and tested. The cleavage activity of dengue 2 NS2B/NS3 protease at **S1** was inhibited by Panduratin A **11** and 4-hydroxypanduratin A **12** with an activity of 21 μM and 25 μM respectively. [3] It was decided to use these compounds as a starting point to find the best mode of binding of such compounds with the dengue virus NS3 protease by docking these compounds against the **S1** pocket of the dengue NS3 protease.



Panduratin A

11 (25 μM)



4-Hydroxypanduratin A

12 (21 μM)

Figure 3.2 Chemical Structure of Panduratin A and 4-hydroxypanduratin A.

3.3- Preparation of the crystal structure.

The crystal structure of the dengue virus NS3/NS2B (pdb code = 2FOM) is considered to be the active form of the protease enzyme due to its combination with the cofactor as reported in the literature.[4] The crystal structure was downloaded from pdb, all hydrogens were added and minimized. The residues of the **S1** pocket were identified to be used as the docking site (Figure 3.3). The structures of **11** and **12** were built by MOE builder and saved as moe and mol2 for the docking that was done by both MOE [5] and FlexX 3.3.1.[6]

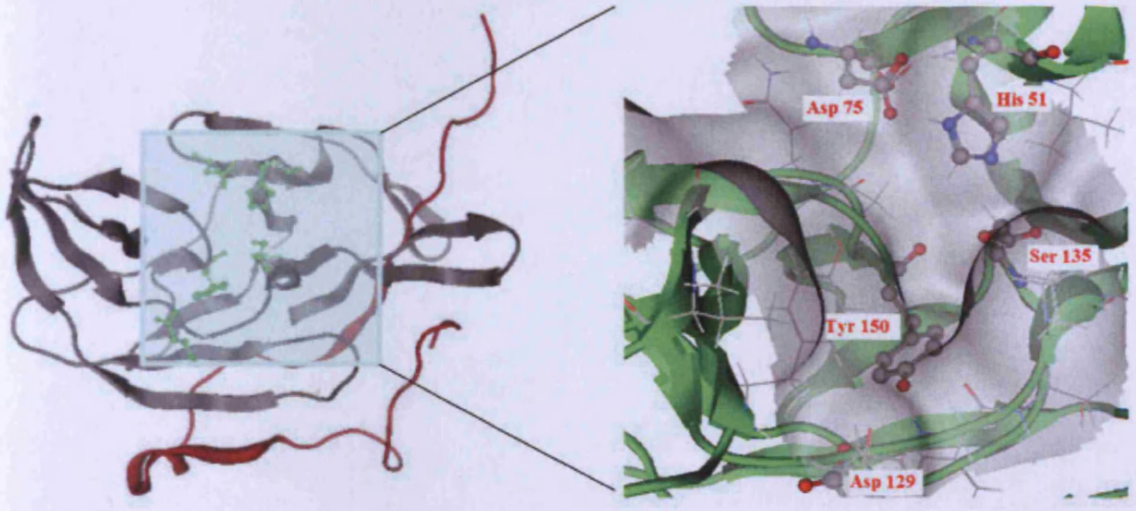


Figure 3.3 Identification of DV NS3 **S1** pocket and its main residues.

3.4- Docking results.

Poses of best scoring (Table 3.1) from both programs were visualized. It was found that the 2-methyl-1-propenyl side chain present in position number 2 of the cyclohexene ring did not show any kind of interactions and had some clashes with different residues within the **S1** pocket.

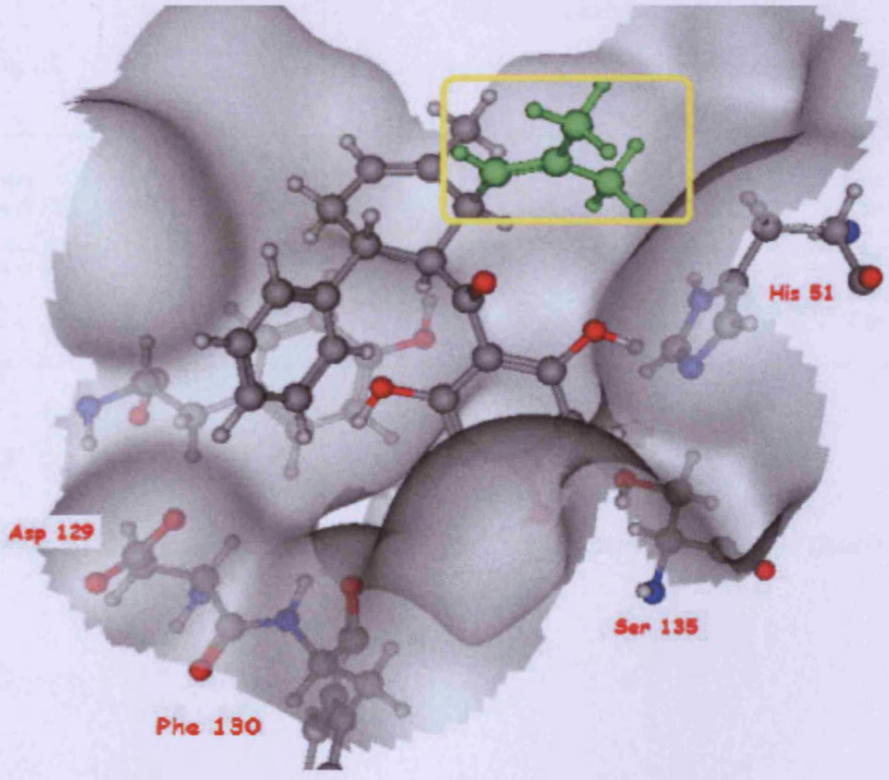


Figure 3.4 Docking of 4-hydroxypandurantin A (**12**) within the S1 pocket

Moreover, it was found that the 6-phenyl ring could be retained or replaced by a basic moiety or by an electron donating group that could increase the chance of interact ion with the Asp 129 and/or Phe 130 residues present in the bottom of the **S1** pocket. The 2, 4, 6-trihydroxyl groups could be replaced by 2, 4, 6-trimethoxy groups that have shown some interactions with Ser 135 –OH group. The presence of the carbonyl functional group of 3-cyclohexenyl methanone was important in the docking against the specified site of action.

Compound	FlexX Score Kcal/mol	MOE docking results and scores.		
		Affinity Kcal/mol	Interacted moiety	Main Residue
4-hydroxy panduratin A **12**	**-6.893**	**-11.25**	**-C=O** **-OCH_3**	**Ser 135** **Tyr 150**
panduratin A **11**	**-6.660**	**-11.1**	**-C=O**	**Ser 135** **Tyr 150**

Table 3.1 Docking scores of MOE and FlexX for BR derivatives and the designed compounds.

A proposed mode of interaction for the new modified compounds was suggested. The oxygen atom at position 2 of the phenyl ring could interact with the hydroxyl group of Ser 135 forming a hydrogen bond. In addition, the carbonyl group is proposed to form a hydrogen bond with the hydroxyl group of Tyr 161. However, the presence of a basic moiety or electron donating group on the other side may have a kind of hydrogen bonding with the carboxylate moiety of Asp 129 or the –C=O group of both Asp 129, and Phe 130.

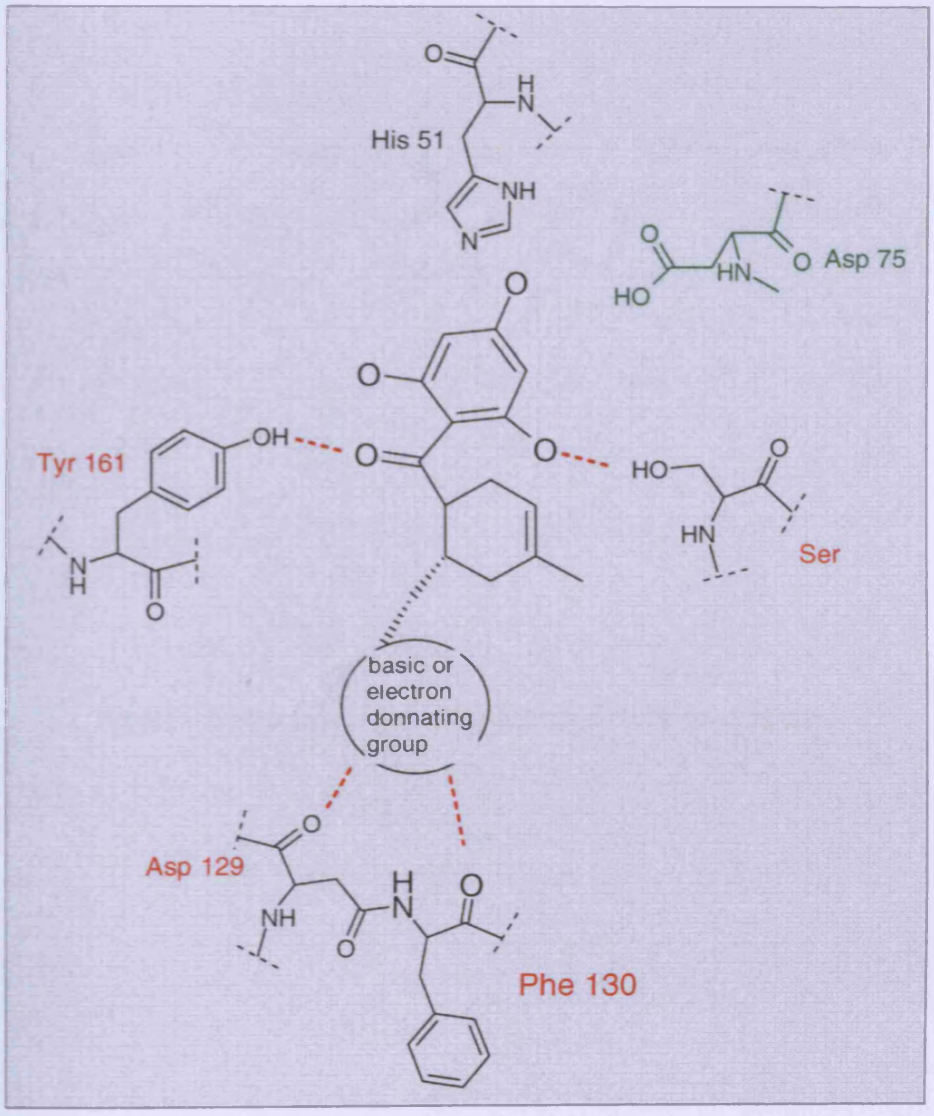
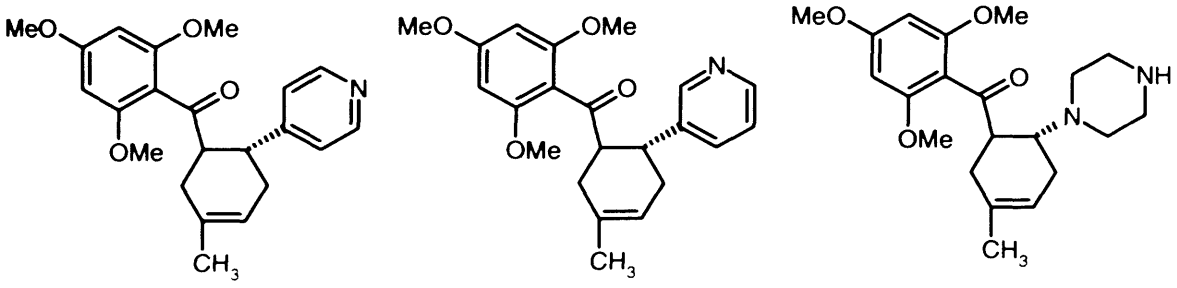


Figure 3.5 Predicted mode of binding of the modified compounds..

3.5- Design of new compounds.

According to the docking results three compounds were designed in which the phenyl ring was replaced by piperazine, 4-pyridine and 3-pyridine. 2, 4, 6-trihydroxy groups substituted with 2, 4, 6-trimethoxy groups were drawn by the MOE builder and were docked against the **S1** pocket of DV NS3 protease.



20 **21** **22**

Figure 3.6 Structure of the newly designed modified structures.

3.6- Docking of 20.

Docking of **20** has shown some interactions in which the 2-methoxy and 4-methoxy groups were involved in hydrogen bond formation with –OH groups of Tyr 161 and Ser 135 respectively. Also the –C=O group of methanone moiety was involved in another hydrogen bond with –OH of Tyr 161, while the 4-pyridine ring was not involved in any kind of interactions.

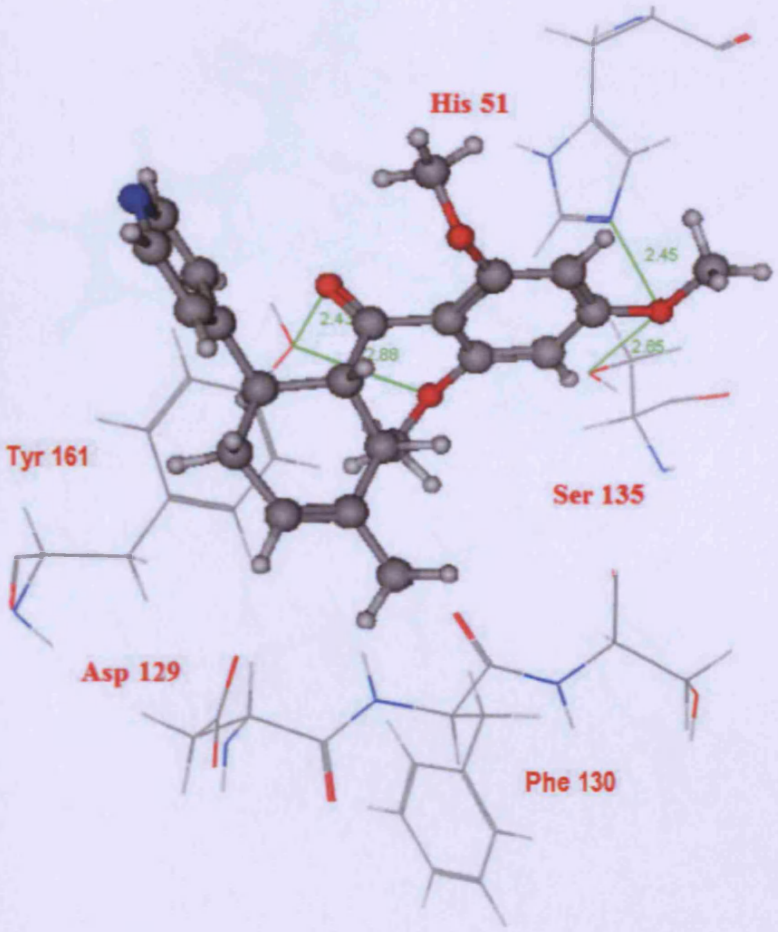


Figure 3.7 Docking results of compound **20** showing interactions with His 51, Ser 135, and Tyr 161 residues.

3.7- Docking of 21.

The same results of **20** were seen in the docking of **21** and the 3-pyridine ring was not involved in any interactions as was expected.

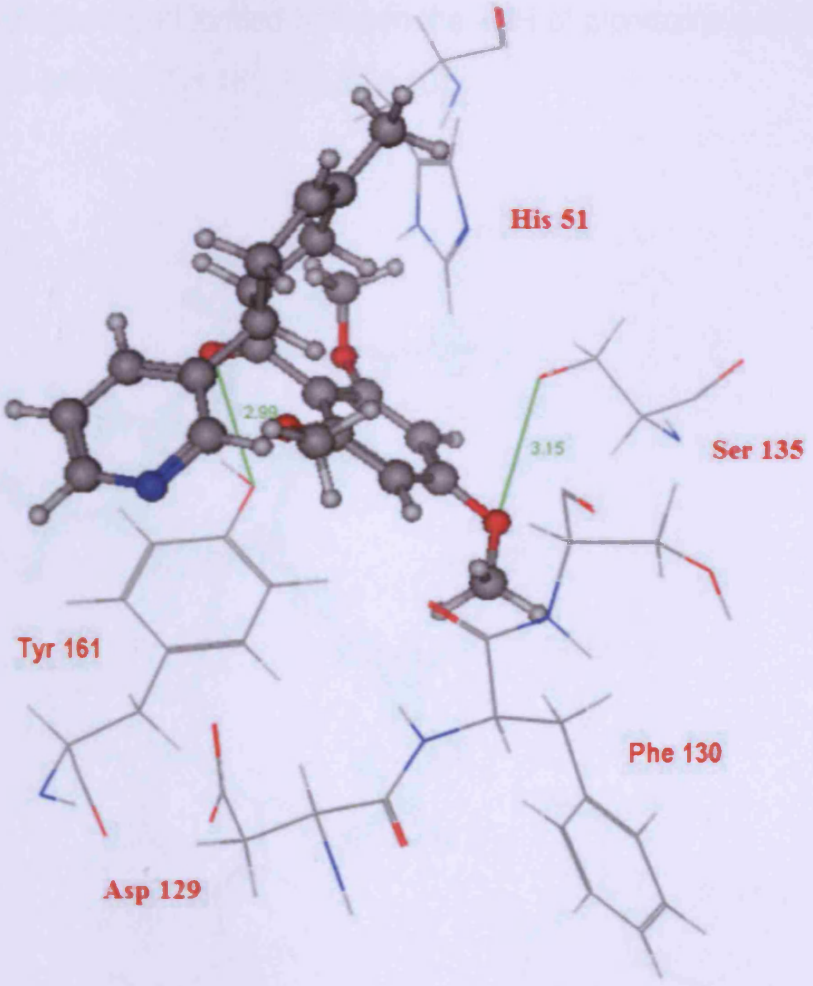


Figure 3.8 Docking of compound **21** showing interactions with Ser 135 and Tyr 161 residues.

3.8- Docking of 22.

Compound **22** has shown two modes of binding: the first mode in which the –C=O group of methanone moiety was involved in the hydrogen bond formation with –OH group of Tyr 161 (Figure 3.9), and the second mode has shown another hydrogen bond formed between the –NH of piperazine and an oxygen atom of –OH group of Tyr 161 (Figure 3.10).

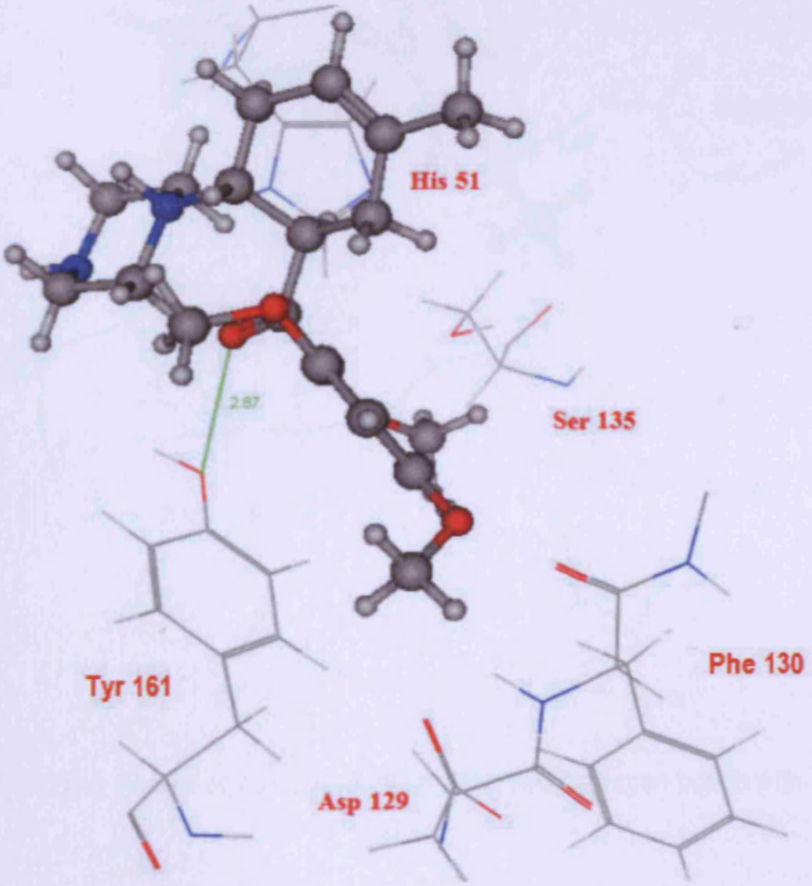


Figure 3.9 Binding mode 1 of compound **22** showing hydrogen bond with Tyr 161 residue.

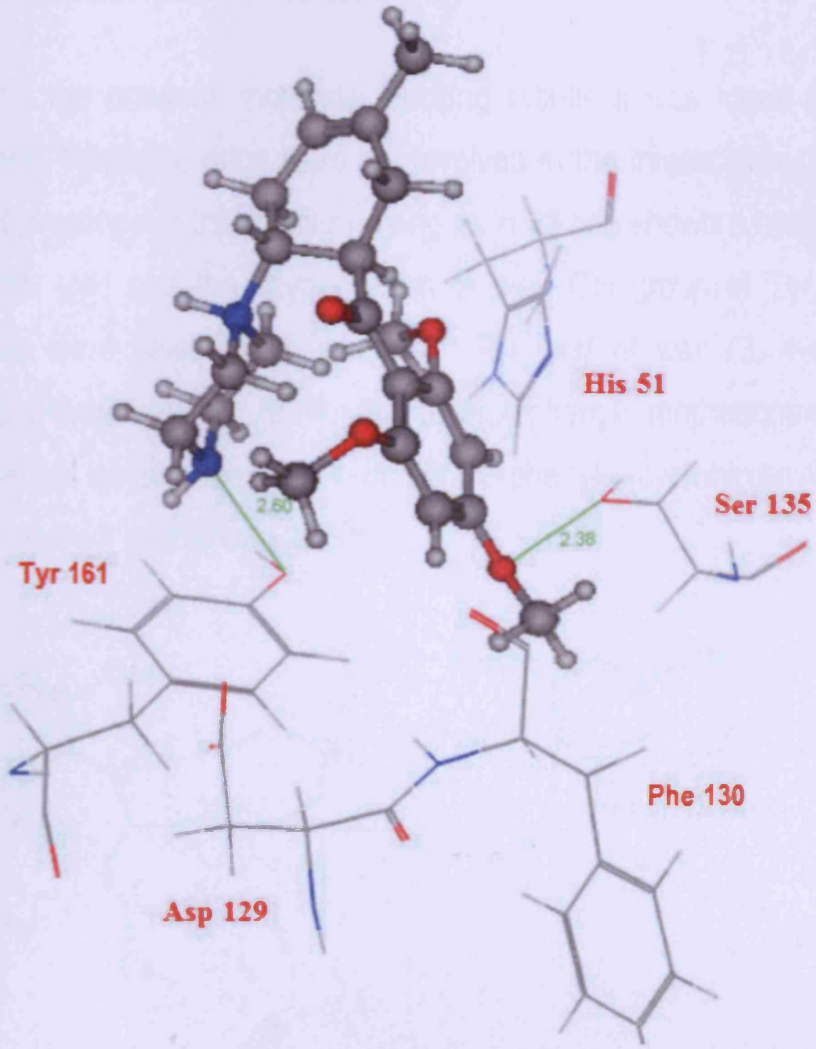
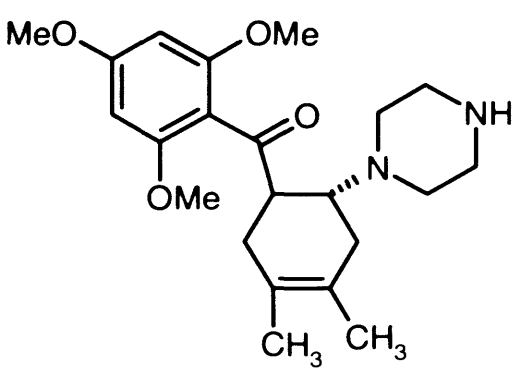


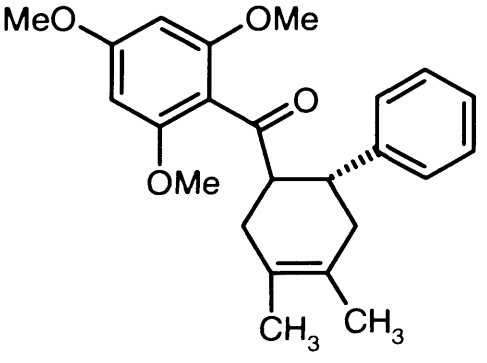
Figure 3.10 Binding mode 2 of compound **22** showing two hydrogen bonds with Tyr 161 and Ser 135.

3.9- Synthesis of (3-methyl-6-phenyl-3-cyclohexenyl)(2,4,6-trimethoxyphenyl) methanone.

Building on the previous molecular docking results it was found that the 3-pyridine and 4-pyridine rings were not involved in the interactions. It was also found that presence of the piperazine ring as in **22** has shown a hydrogen bond between its –NH and the oxygen atom of the –OH group of Tyr 161. Two compounds were selected for synthesis. The first of was (3, 4-dimethyl-6-piperazino-3-cyclohexenyl) (2, 4, 6-trimethoxyphenyl) methanone (**modified 22**). While the second was (3, 4-dimethyl-6-phenyl-3-cyclohexenyl) (2, 4, 6-trimethoxyphenyl) methanone **27.**



Modified 22



27

Figure 3.11 The chemical structure of modified **22** and **27** in which 3,4-dimethyl substitution of the cyclohexene ring is introduced for the ease of synthesis.

Docking of **27** has revealed some interactions including the hydrogen formed between the oxygen atom of the 4-methoxy group and the –OH group of Ser 135. Also, the hydrogen bond between the –C=O group of methanone moiety and the –OH group of Tyr 161.

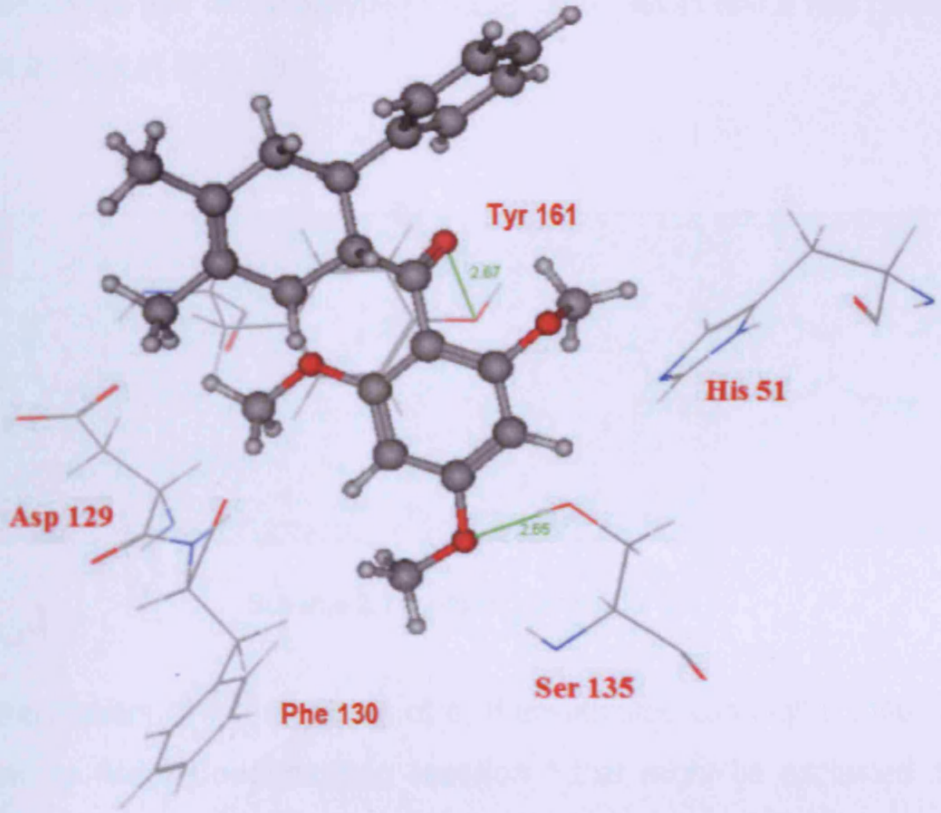
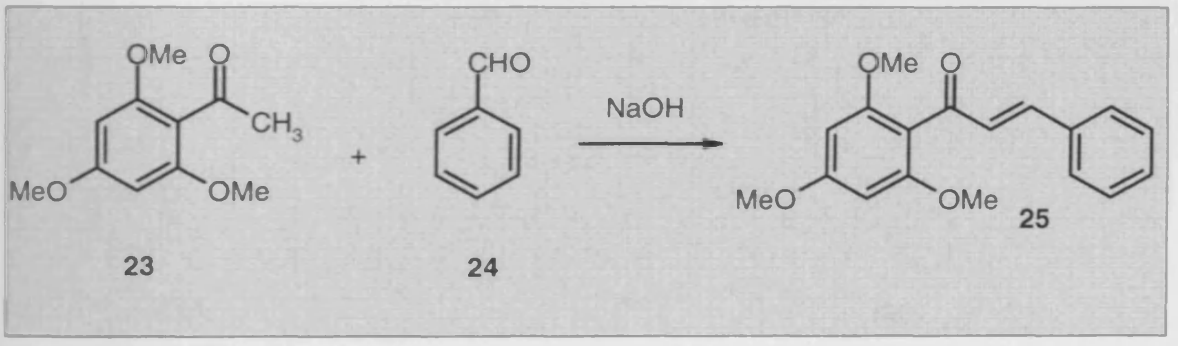


Figure 3.12 Docking results of compound **27** showing the main interactions with Tyr 161 and Ser 135.

Searching for the best synthetic method for both of them, it was found that substituted cyclohexenyl methanone could be synthesized by the Diels Alder reaction starting with an α, β-unsaturated carbonyl compound. For (3-methyl-6-phenyl-3-cyclohexenyl) (2, 4, 6-trimethoxyphenyl) methanone the starting compound was 3-phenyl-1-(2, 4, 6-trimethoxyphenyl) -2-propen-1-one that was synthesized according to the literature [7] by the reaction of 2, 4, 6-trimethoxy acetophenone with benzaldehyde in presence of NaOH and it was isolated as yellow crystals in 80 % yield.



Scheme 3.1 Synthetic pathway for **25.**

The mechanism of the synthesis of α, β-unsaturated carbonyl compounds is referred to **Aldol Condensation reaction** [8] that might be explained in five steps. In the first step the hydroxyl ion of NaOH will facilitate the removal of acidic hydrogen from 2, 4, 6-trimethoxy acetophenone leading to the formation of a strong nucleophilic enolate ion (Figure 3.13. **I**). The second step is the nucleophilic attack of the enolate ion of the electrophilic carbon of benzaldehyde resulting in an intermediate alkoxide (**II**).

The formed alkoxide will deprotonate a water molecule resulting in a hydroxide ion and β-hydroxyl carbonyl compound that is called Aldol product (**III**). Step 4 involves the dehydration of the previously formed Aldol product when a hydroxyl ion attacks the acidic hydrogen resulting in the formation of an active enolate (**IV**). In the last and final step the active enolate will be rearranged to form the 3-phenyl-1-(2, 4, 6-trimethoxyphenyl)-2-propen-1-one (**V**) **25.**

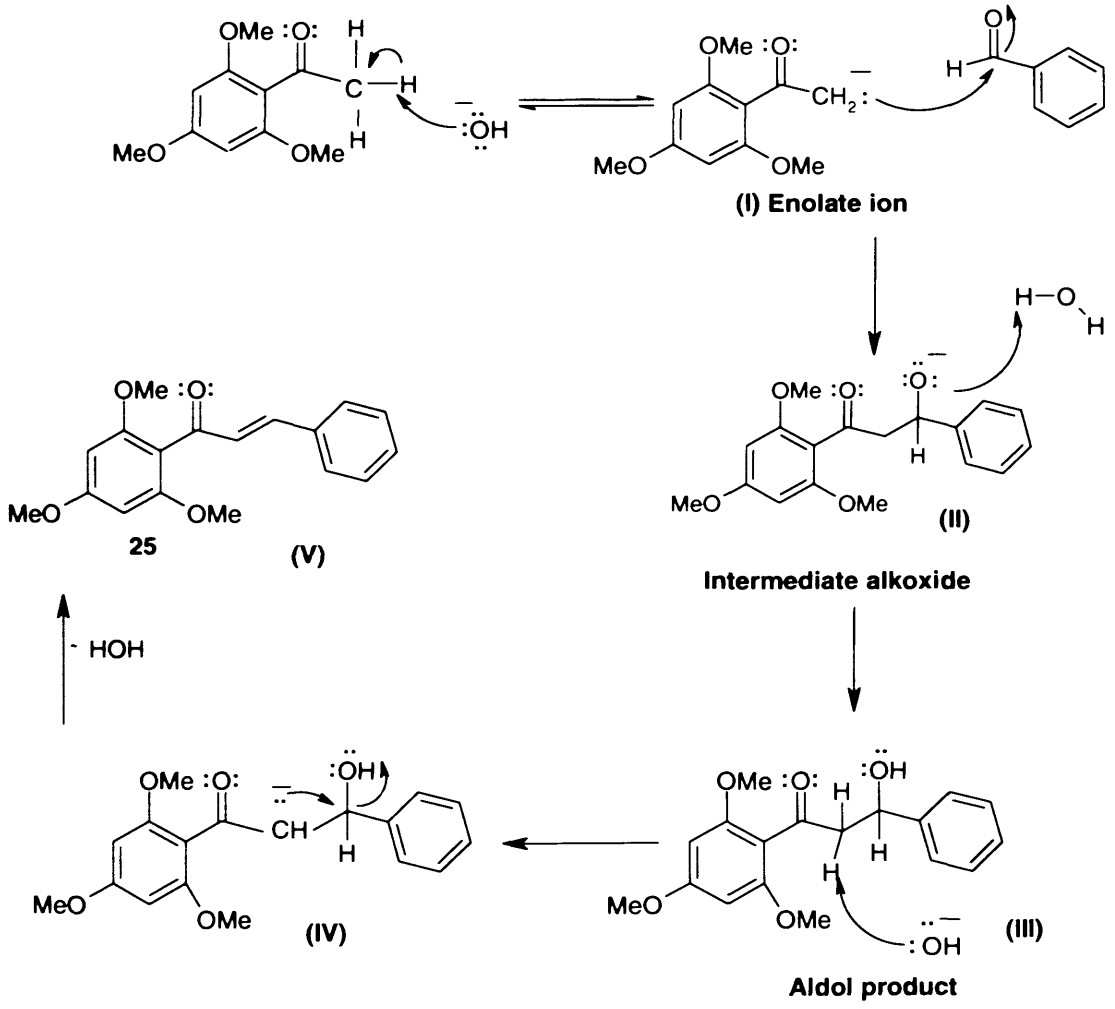
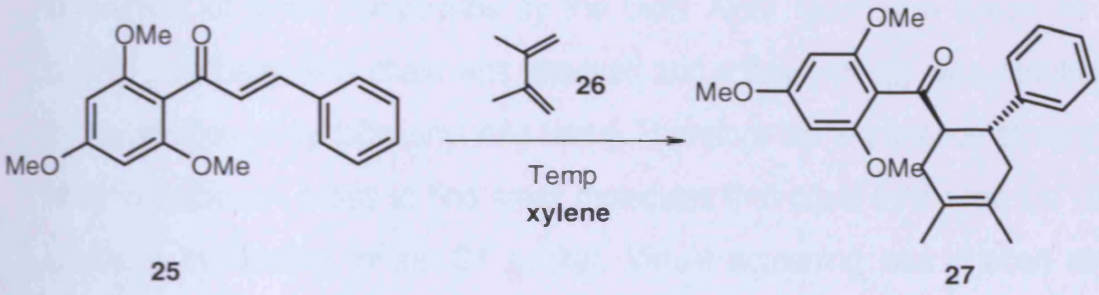


Figure 3.13 Aldol condensation mechanism for the synthesis of chalcones.

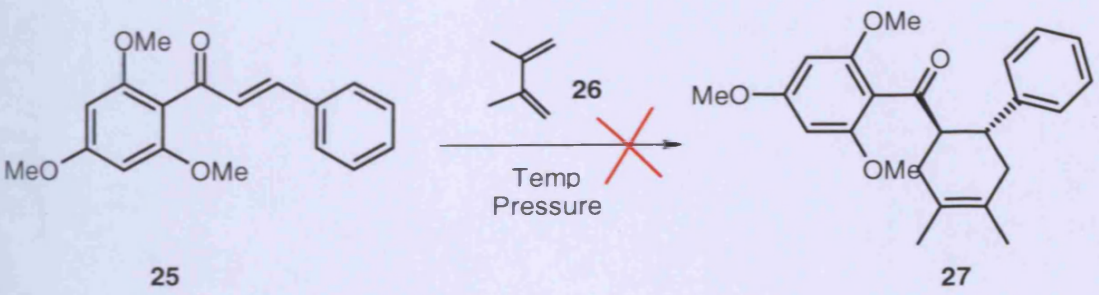
After the synthesis of **25** it was supposed that the reaction of 3-phenyl-1-(2, 4, 6-trimethoxyphenyl)-2-propen-1-one with diene such as dimethyl butadiene that was available commercially at high temperature would result in the formation of **27.**



25 27

Scheme 3.2 Synthetic pathway for **27.**

The reaction was first done under high temperature and it was repeated under both high temperature in an oil bath and a U-tube was used to apply the high pressure but, unfortunately, it did not give the required compound.



25 27

Scheme 3.3 Synthetic pathway for **27** on application of pressure.

3.10- Conclusion.

Docking of panduratin A **11** and 4-hydroxypanduratin A **12** against DV NS3 protease (pdb code = 2FOM) has revealed they have a binding mode in which the substituted benzoyl moiety was found interacting with some residues in the **S1** pocket such as Tyr 161, Ser 135, and His 51. The trial to synthesize some derivatives of these compounds by the Diels Alder reaction in which the 2-methyl-1-propenyl side chain was removed and a basic moiety was introduced in the position of cyclohexenyl ring failed. Therefore the importance of using a second approach arises to find small molecules that could inhibit the DV NS3 protease by binding to the **S1** pocket. Virtual screening was chosen as a second approach to search for new scaffolds that could be good hits to be used as a starting point.

3.11- References.

1. Krishna Murthy, H. M.; Judge, K.; Padmanabhan, R. Crystal structure of Dengue virus NS3 serine protease in complex with a Bowman-Birk inhibitor: Implications to Flaviviral polyprotein processing and drug design. *J. Mol. Biol* **2000,** 301, 759-767
2. Perona, J.; Craik, C. S. Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **1995,** *4*, 337-360.
3. Tan, S. K.; Richard, P.; Rohana, Y.; Halijah, I.; Norzulaani, K.; Noorsaadah, A. Inhibitory activity of cyclohexenyl chalcone derivatives and flavanoids of fingerroot, Boesenbergia rotunda (L.) towards dengue-2 virus NS3 protease. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 3337-3340.
4. Krishna Murthy, H. M.; Clum, S.; Padmanabhan, R. Dengue virus NS3 serine protease. *J. Biol. Chem.* **1999,** *274* (4), 5593-5580.
5. Molecular Operating Environment (MOE), Chemical computing group.Inc. Montreal, Quebec, Canada. **2005.**
6. Rarey, M.; Kramer B.; Lengauer T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm *J. Mol. Biol.* **1996,** *261*(3), 470-89.
7. Charles, K.; Bradsher; Lennard, J.; Wissow. Aromatic cyclodehydration. Use of intermediates obtained by the diene addition reaction. *J. Am. Chem. Soc.* **1946,** *68* (11), 2149-2151.
8. McMurry, J. Organic Chemistry. 5th ed. Brooks/Cole, **1999.**

Chapter 4

Design of Dengue Virus NS3 Serine Protease Inhibitors.

Chapter 4

4.1- Design of DV NS3 Protease Inhibitors by Structure-based Virtual Screening.

Virtual Screening (VS) has successfully been used in the process of drug discovery.[1] DV has three crystal structures available in pdb. The first one is complexed with MbBBI (pdb code = 1DF9) that was used previously for the development of non-peptidic inhibitors.[2-3] The second crystal structure is the dengue virus NS3/NS2B (pdb code = 2FOM) which is considered to be the active form of protease enzyme, [4] while the third one is the dengue virus NS3 domain alone (pdb code = 1BEF).[5] To date there is no crystal structure of dengue NS3 protease that is complexed with a small molecule inhibitor. The crystal structure of the dengue virus NS3 in complex with MbBBI was used for discovery of dengue inhibitors and resulted in some compounds with high micro molar inhibition activity. The structure of MbBBI might cause some conformational changes to the site of action of NS3 protease that might be the main cause of inhibition beside its two Arg and Lys side chains that were found interacting with the **S1** pocket residues. Also, if we depend only on the basic residue found in the **S1** pocket that would result in a basic compound to be used as an inhibitor. But, the interesting point is that the **S1** pocket alone could be used for discovery of novel NS3 protease inhibitors for dengue. The assumption of the use of the **S1** pocket for design of flaviviral NS3 protease inhibitors has been reported in the literature.[6-7] For this reason we have identified the main residues in the **S1** pocket together with the oxyanion hole.

A similarity search was done to find a similar NS3 structure from the *Flaviviridae* family that has a high percentage of identity and similarity especially for its **S1** pocket. WNV NS3 serine protease has been found to be the best choice with identity = 62.5 % and was inhibited by a potent peptidic inhibitor that was complexed with its crystal structure (pdb code = 2FP7) [8] and was found to be similar to that of the DV NS3/NS2B protease.

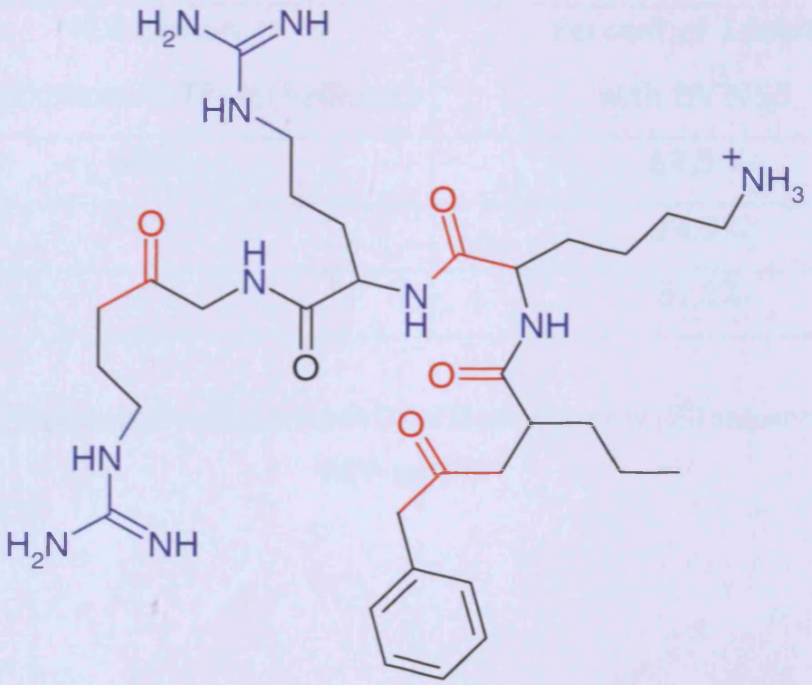


Figure 4.1 Structure of the peptide inhibitor of WNV NS3/NS2B protease.

The developed idea was to use the similarity between the DV NS3 serine protease and WNV NS3 serine protease by means of using the peptidic inhibitor in building a pharmacophore model that could be used for VS searching for potential hits. To begin building the model, the crystal structures of the two viruses NS3 proteases were downloaded for sequence alignment as the first step.

NS3 domain (Serine protease/NTPase/helicase)	Percent of Identity with DV NS3.
WNV	62.5 %
HCV	24.9%
YFV	51.2%

Table 4.1 Percentage of identity between DV NS3 sequence and NS3 sequence of WNV, HCV, and YFV

4.2- Superimposition of the crystal structures of DV and WNV NS3 proteases.

A superimposition of the two crystal structures of dengue and WNV NS3 proteases with pdb code = 2FOM and 2FP7 respectively was done by the MOE 2007.08. By downloading both crystal structures into the MOE [9] window. The residues of the dengue virus NS3 **S1** pocket slightly deviated from that of the WNV **S1** pocket; RMSD was found 1.37 A° between the two backbones.

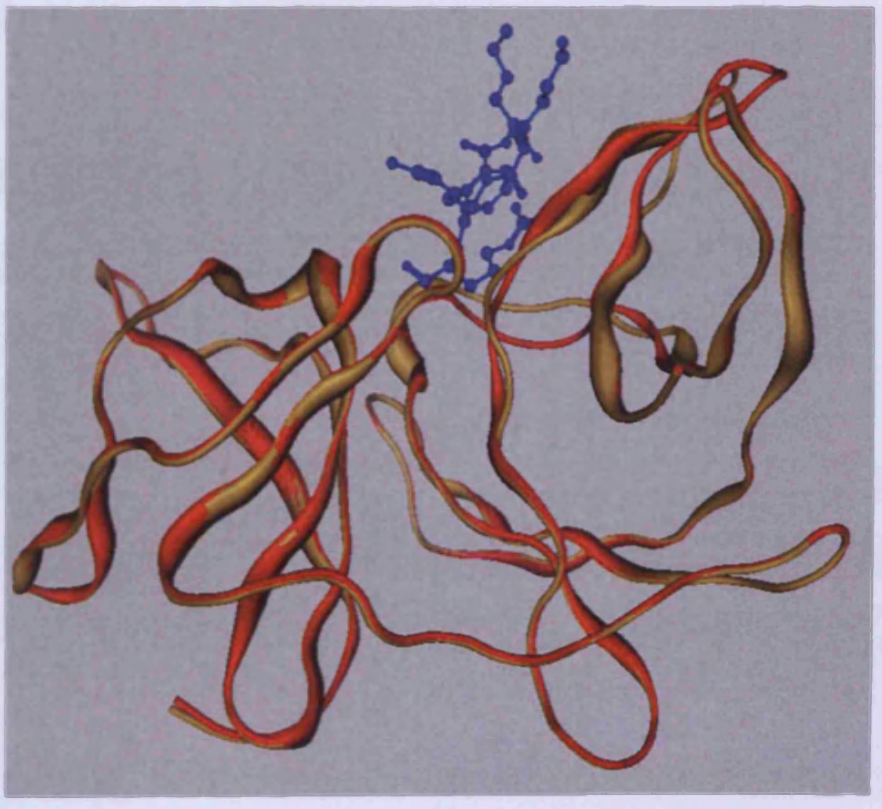


Figure 4.2 Superimposition of both 2FOM (in yellow) of dengue NS3 protease and 2FP7 (in red) for WNV NS3 protease with the peptidic inhibitor (in blue)

The alignment was also performed with the ClustalW program provided by the Uniprot website (http://services.uniprot.org/clustalw/), between the serine protease/NTPase/helicase NS3 sequence of the Dengue 2 virus that has the sequence from 1475-2093, and WNV that has the serine protease/NTPase/helicase NS3 sequence 1502-2120 (Figure 4.2). Identical residues are marked with a star and are shaded with grey as found in the following figure.

```
VKKQRAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRG  1530  DEN2
QYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKG  1556  WNV
  .:*.*****.***    *.:   *.***  :*:**  * ****  **.***:**.*:*
AVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGL  1590  DEN2
AALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGHDEVQMIVVEPGKNVKNVQTKPGV  1616  WNV
*.**    *::* *..**:* :.*** ***: :*:  :***::.:***** : *******:
FKTNAGTIGAVSLDFSPGTSGSPIIDKKGKVVGLYGNGVVTRSGAYVSAIAQTEKSIEDN  1650  DEN2
FKTPEGEIGAVTLDYPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMEEPA  1676  WNV
***   * ****:**:..*******:**:*.*:********:  .*:*:***.* *:  *
ME-IEDDIFRKRRLTIMDLHPGAGKTKRYLPAIVREAIKRGLRTLILAPTRVVAAEMEEA  1709  DEN2
PAGFEPEMLRKKQITVLDLHPGAGKTRKILPQIIKEAINKRLRTAVLAPTRVVAAEMSEA  1736  WNV
*  :* :::**:::*:**********:: ** *::***:: *** :***********.**
LRGLPIRYQTPAIRAEHTGREIVDLMCHATFTMRLLSPVRVPNYNLIIMDEAHFTDPASI  1769  DEN2
LRGLPIRYQTSAVHREHSGNEIVDVMCHATLTHRLMSPHRVPNYNLFIMDEAHFTDPASI  1796  WNV
**********.*:: **:*.****:*****:* **:** *******:**************
AARGYISTRVEMGEAAGIFMTATPPGSRDPFPQSNAPIIDEEREIPERSWNSGHEWVTDF  1829  DEN2
AARGYIATKVELGEAAAIFMTATPPGTSDPFPESNAPISDMQTEIPDRAWNTGYEWITEY  1856  WNV
******:*:**:****.*********: ****:***** * : ***:*:**:*:**:*::
KGKTVWFVPSIKAGNDIAACLSKNGKKVIQLSRKTFDSEYAKTRTNDWDFVVTTDISEMG  1889  DEN2
VGKTVWFVPSVKMGNEIALCLQRAGKKVIQLNRKSYETEYPKCKNDDWDFVITTDISEMG  1916  WNV
 *********:* **:** **.: *******.**::::**.* :.:*****:********
ANFKAERVIDPRRCMKPVILTDGEERVILAGPMPVTHSSAAQRRGRIGRNPKNENDQYIY  1949  DEN2
AMFKASRVIDSRKSVKPTIIEEGDGRVILGEPSAITAASAAQRRGRIGRNPSQVGDEYCY  1976  WNV
*****.****.*:.:**.*: :*: ****. * .:* :************.: .*:* *
MGEPLENDEDCAHWKEAKMLLDNINTPEGIIPSMFEPEREKVDAIDGEYRLRGEARTTFV  2009  DEN2
GGHTNEDDSNFAHWTEARIMLDNINMPNGLVAQLYQPEREKVYTMDGEYRLRGEERKNFL  2036  WNV
 *.. *:*.: ***.**:::***** *:*::..:::******  ::********** *..*:
DLMRRGDLPVWLAYRVAAEGINYADRRWCFDGVKNNQILEENVEVEIWTKEGERKKLKPR  2069  DEN2
EFLRTADLPVWLAYKVAAAGISYHDRKWCFDGPRTNTILEDNNEVEVITKLGERKILRPR  2096  WNV
:::* .*********:*** **.* **:***** :.* ***:* ***: ** **** *:**
WLDARIYSDPLALKEFKEFAAGRKSLTLNLITEMGRLPTFMTQKARDALDNLAVLHTAEA  2129  DEN2
WADARVYSDHQALKSFKDFASGKRSQ-IGLVEVLGRMPEHFMVKTWEALDTMYVVATAEK  2155  WNV
* ***:***  ***.**:**:*:::*  :.*:  :**:* .:  *: :***.: *: ***
```

Figure 4.3 ClustalW sequence alignment between DV NS3 and WNV NS3.

By visualization of the site at which the peptidic inhibitor fits in both DV NS3 protease and WNV NS3 protease before and after the superimposition, it was found that it had the same binding site. The most important here was the Arg side chain from the inhibitor that was found in the **S1** pocket of both DV and WNV NS3 protease that could give a clear overview of the possible interactions between the guanidine moiety of the bound ligand and the residues at the **S1** pocket of the DV NS3 protease.

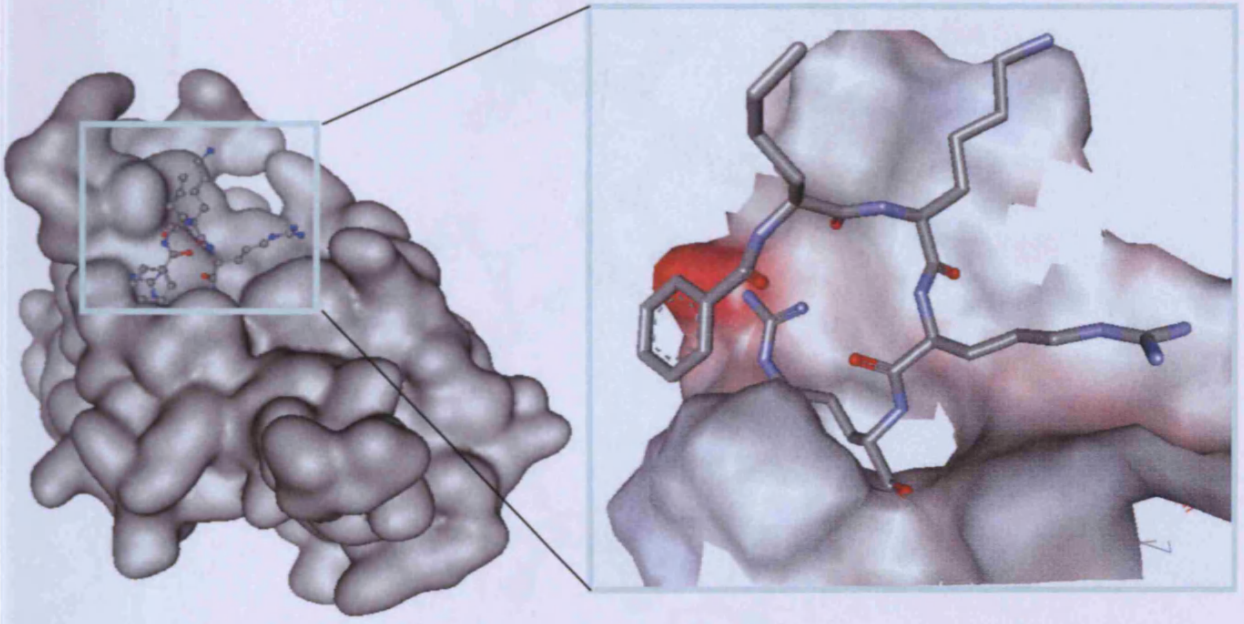


Figure 4.4 The site of the binding of peptidic inhibitor
That is mainly within the S1 pocket.

The observed difference between the two pockets of both DV and WNV NS3 serine proteases was the presence of Tyr 130 in WNV instead of Phe 130 in DV. The overall orientation of the residues at the **S1** pocket in both DV NS3 protease and WNV NS3 protease that were involved in the interactions was almost the same.

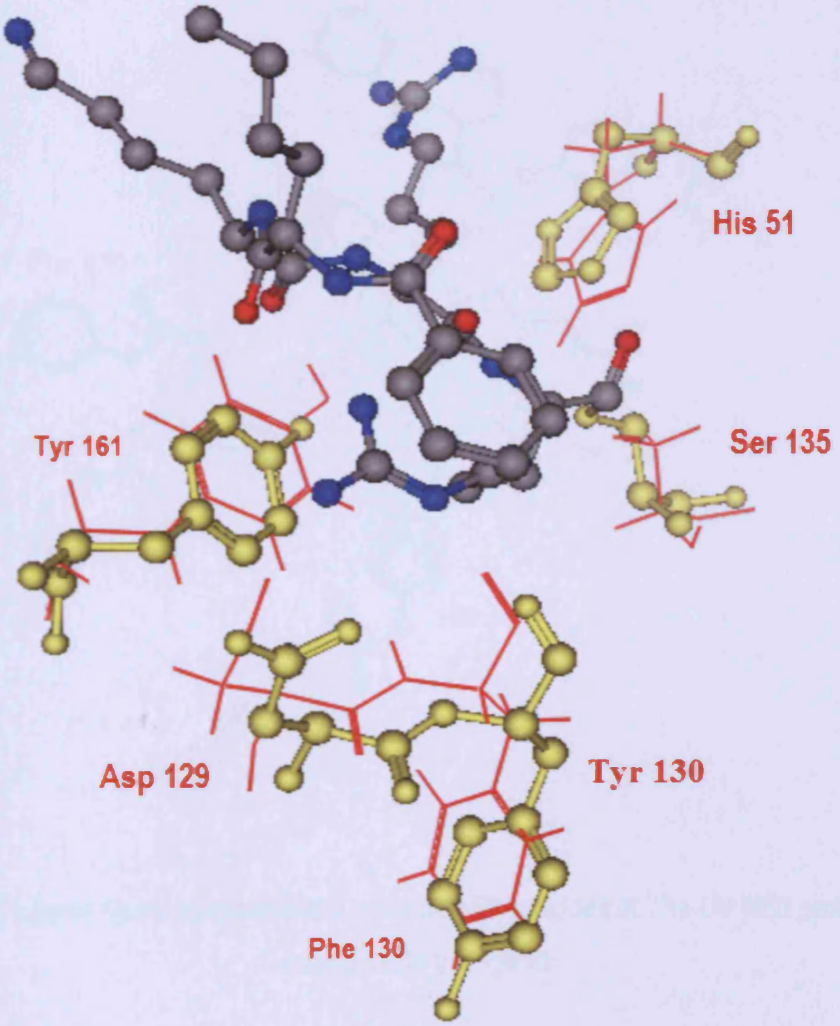


Figure 4.5 WNV residues in ball and stick (Yellow) and DV in the line representation (red) are showing the same orientation of their functional groups.

4.3- Building of the Pharmacophore model.

A variety of pharmacophoric features were introduced by the ligand such as the two carbonyl groups of amide bonds that were involved in the hydrogen bond formation with Tyr 161 and His 51 of WNV NS3 protease and Tyr 161 and His 51 of DV NS3 protease. Also, it was found that the Asp 129 residue was involved by its carboxylic –C=O group as an acceptor with the positively charged N atom of ligand present in its guanidinium moiety in both WNV and DV. Comparison of the formed interactions in both the WNV NS3 protease and the DV NS3 protease with residues of the **S1** pocket and with the measured distances is shown in figure 4.7 and figure 4.8. Table 4.2 shows the involved residues together with the measured distances for each enzyme.

WNV NS3 protease		DV NS3 protease	
Tyr 130	2.8 A°	Phe130	1.83 A°
Asp 129	2.87 A°	Asp 129	2.00 A°
Tyr 161	2.83 A°	Tyr 161	2.62 A°
His 51	2.77 A°	His 51	2.23 A°
Gly 153	2.86 A°		

Table 4.2 Comparison of the involved residues and measured distances in the interactions with the peptide inhibitor.

The DV NS3 protease-Ligand interactions have shown that the artificially fixed ligand was fitted in the same site as in the case of the WNV NS3 protease. Taking into account the similarity between the two enzymes to find a small potential compound that could be suitable for the DV NS3 protease, a virtual screening (VS) approach was chosen for that purpose. Building a pharmacophore model depending on the previous data was done by choosing the pharmacophoric features that were involved in the interactions including two cationic/donating features derived from the guanidine moiety $-NH_2$ and –C=N groups. It also includes two accepting groups derived from two separate carbonyl groups. Each feature was built with the radius = 1.4 A°. A pharmacophoric model was built and was ready to search databases to get hits. Two types of databases were used: the first was a zinc database of drug-like compounds subset (http://zinc.docking.org/). A number of drug-like subsets were downloaded in mol2 format that were converted to mdb format by the MOE. A conformation import was carried out in order to get the conformations for the databases that will be searched. The second type of databases was the MOE databases provided by the MOE package.

4.5- Selection of Hits.

Searching of hits is a long process that needs a long time as each database will produce a large number of conformations that need to be searched. Also, when the process of searching is finished a large number of hits may be produced as well. Conformational import of Zinc databases was done to generate all possible conformations. Then searching of both Zinc and Moe databases was performed starting with a total number of conformations = 8 million conformations. Filteration according to pharmacophore features resulted in a total number of 300 hits. All the produced hits were visualized in order to be filtered and restricted to a small number that could have a high priority. The hits were filtered according to Lipinski's rule of five [11] which states that a compound to be considered having drug-like properties should be of molecular weight less than 500, have cLogp less than 5, have a number of hydrogen bond donors less than 5, and the number of hydrogen bond acceptors should be less than 10. A second step of filtration was performed according to RMSD from the pharmacophore model. Hits with RMSD more than 1.2 were rejected because it was observed that RMSD values more than 1.2 did not fit well. As a result a number of 15 hits of different scaffolds with a variety of chemical scaffolds including substituted pyridine, triazine, piperazine, thiadiazole, pyrimidine, thiopyrimidine, thiadiazolane, and hydrazinecarboximidothioate may have the chance to provide at least one potential lead compound to be selective for our enzyme. Docking of all these hits against the DV NS3 protease in the **S1** pocket as a docking site was essential to find out the best binding mode by docking of each hit and testing weather there is any modification that needed to be performed.

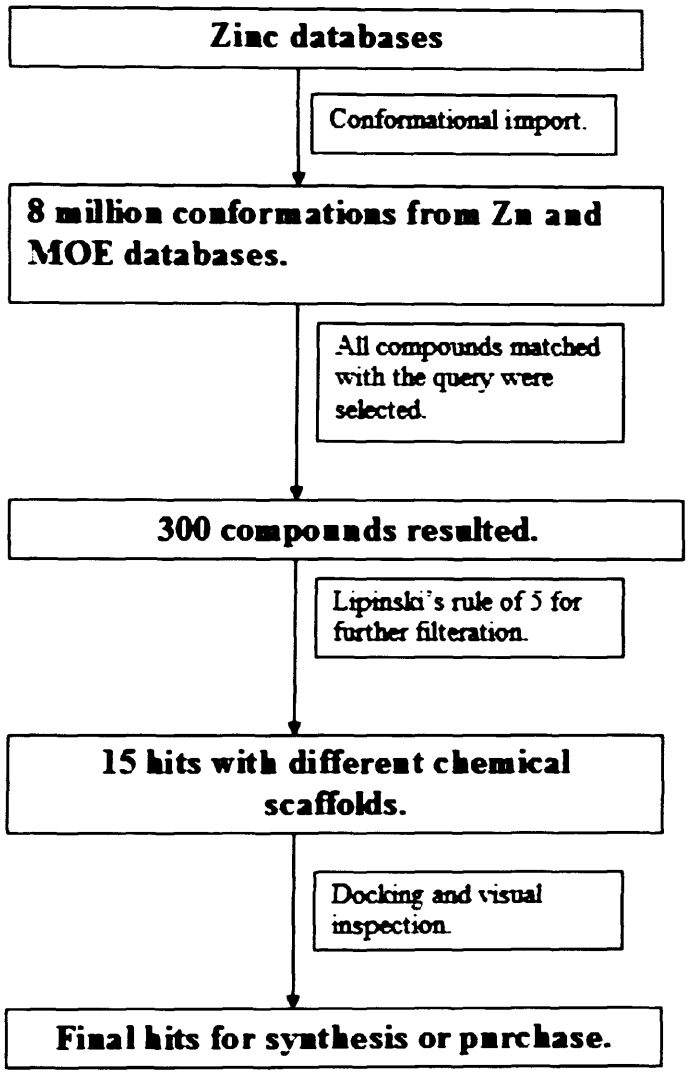
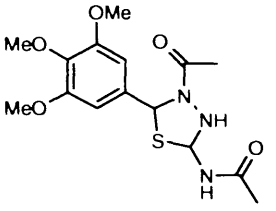
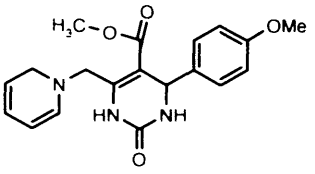
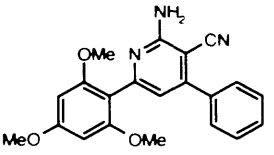
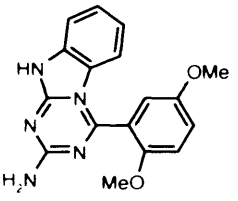
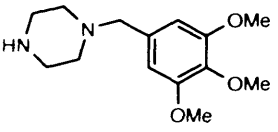
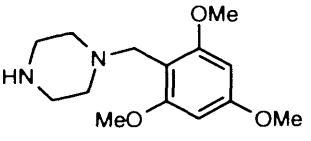
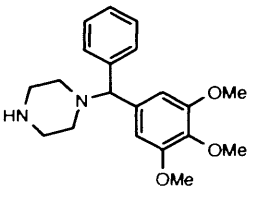
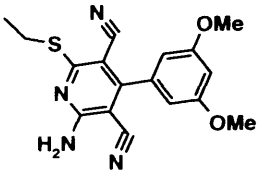


Figure 4.11 A diagrammatic chart describes the steps of virtual screening.

database reference No	Structure	RMSD
• ZINC05084823 • Molecules = 7273S		1.21 0.93
ZINC05007760		1.07
ZINC02204729		1.10
• ZINC04741639 • MOE = 1972394		0.959 1.0
• ZINC00172148 • MOE = 106864		1.02 1.0
ZINC04948349		1.21
ZINC03727553		1.12
• ZINC04480654 • Chemical block = A3847/0163399		0.932 0.80

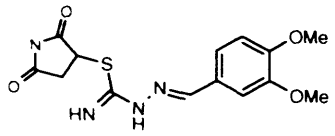
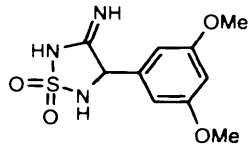
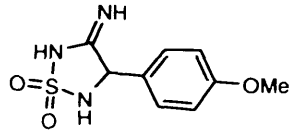
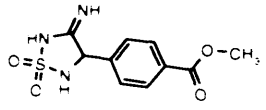
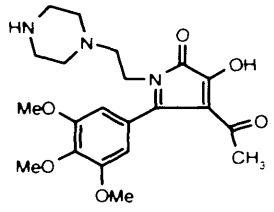
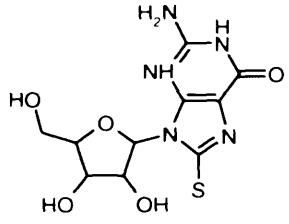
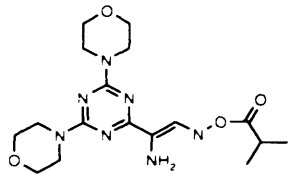
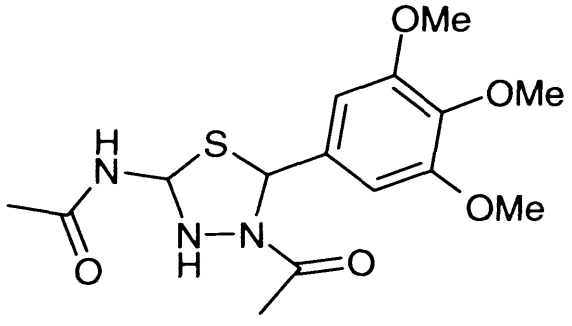
• ZINC03901312 • MOE = 13954		0.927
• ZINC05477623 • Molecules = 193035		0.93 1.08
ZINC05093402		0.92
ZINC05042684		0.92
• ZINC04904027 • MOE = 111636		0.527 0.92
Molecules 10304		0.72
Chemical block A3899/0165868		0.919

Table 4.3 Selected hits with their database numbers and RMSD values.

4.6-Docking and Synthesis of the Hits.

As mentioned before a number of different chemical scaffolds were found matching our pharmacophore model. Most of these hits were not available commercially. As a result a chemical synthesis of those was the proper method. Before the chemical synthesis docking was done against the DV NS3 protease **S1** pocket in order to observe the predicted binding mode using the London dG scoring provided in the MOE. A suitable synthetic pathway was selected to synthesize each hit as was found or after modification that may be needed for best fitting and/or synthesis. In the following section the docking and synthesis of each scaffold will be discussed in detail.

4.6.1- Docking of Thiadiazole based scaffold Hit.



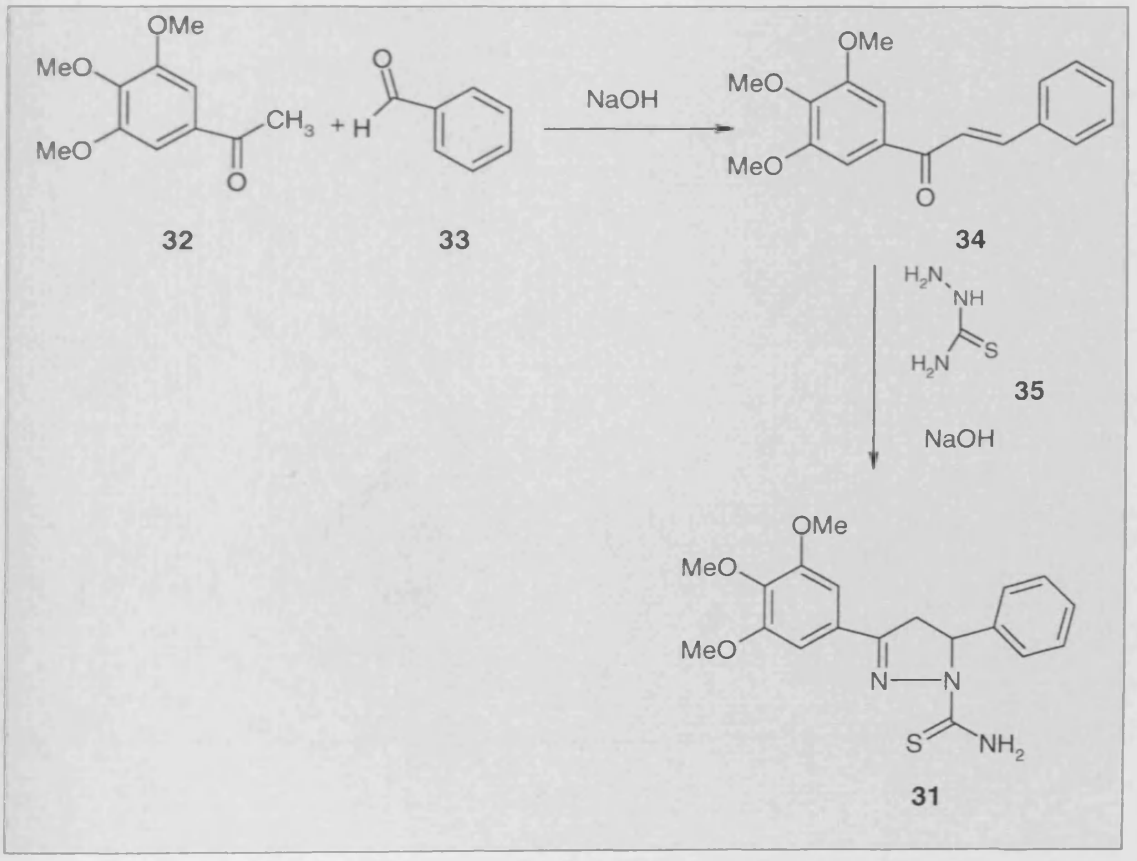
30

Figure 4.12 Chemical structure of hit **30.**

The first scaffold that was found in both Zinc and MOE databases was found as a thiadiazole ring system substituted with a 3, 4, 5-trimethoxy phenyl ring, acetyl and acetamido functional groups. By observing the hit within the model we found the sulphur atom in the thiadiazole ring did not contribute to the features of the model. Also, both the acetyl and acetamido side chains were not involved neither in the model nor in the docking interactions. As found in the highly ranked pose in the docking output the methoxy groups in 3, 4, 5-trimethoxyphenyl moiety may be of interest in interacting with Ser 135. In order to stabilize the interactions from the other side of the molecule an amino group may be useful for hydrogen bond formation with –C=O of Asp 129. That could be achieved by substitution on the N-3 of the thiadiazole ring with $-NH_2$ group that is separated by one carbon from the N atom of thiadiazole in order to be at a suitable distance from –C=O of Asp 129. The sulfur atom could be replaced by carbon as it does not have any contribution and both the acetamido and acetyl groups in position 2 and 4 respectively could be removed.

4.6.2- Synthesis of 5-phenyl-3-(substituted-trimethoxyphenyl)-4, 5-dihydro-1H-pyrazolecarbothioamide.

The synthesis of 1-thiocarbamoyl-3,5-diaryl-4,5-dihydro (1H)- pyrazole was achieved by the reaction of 1 equivalent of Chalcone [3-phenyl-1-(3,4,5-trimethoxy phenyl)-1,2-propene-1-one] **34** and 2 equivalent of thiosemicarbazide **35** in the presence of a base (NaOH) and by using ethanol as a solvent and the mixture was heated (scheme 4.1).[12] According to the reported procedure 5-phenyl-3-(substituted-trimethoxyphenyl)-4,5-dihydro-1H-pyrazolecarbothioamide was synthesized and the yield was 49 % and the product was recrystalised from ethanol.



Scheme 4.1 synthesis of pyrazole carbothioamide derivatives.

The mechanism of the first step of the Chalcone formation could be explained by Aldol condensation [13] in which the 3, 4, 5-trimethoxy acetophenone and benzaldehyde were stirred together for 40 minutes in basic medium to give **34** and the formed product was filtered, washed and dried then refluxed with thiosemicarbazide **35** that would act as a nucleophile in the basic medium to attack **34.** The reaction mechanism involved the formation of hydrazone that resulted from the attack of thiosemicarbazide on **34** as an intermediate and was then followed by the addition of –NH to the olefinic C=C of Chalcone and by further condensation it will close the pyrazole ring.

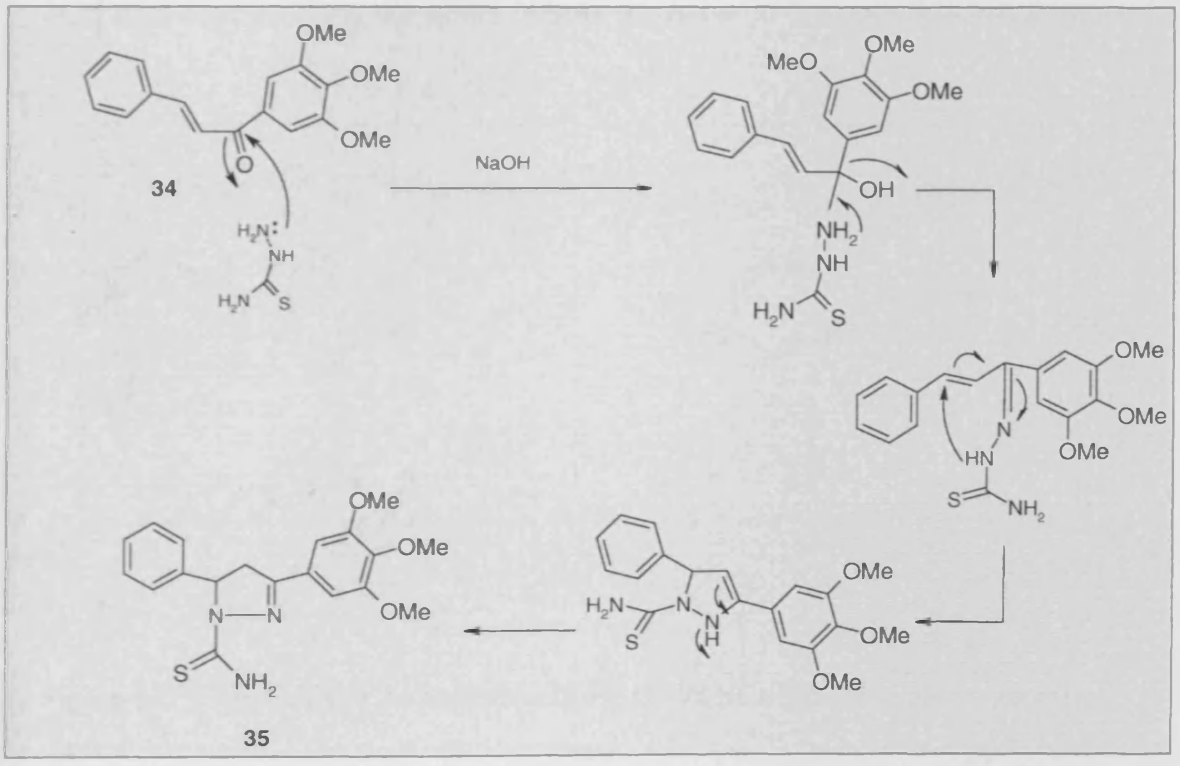
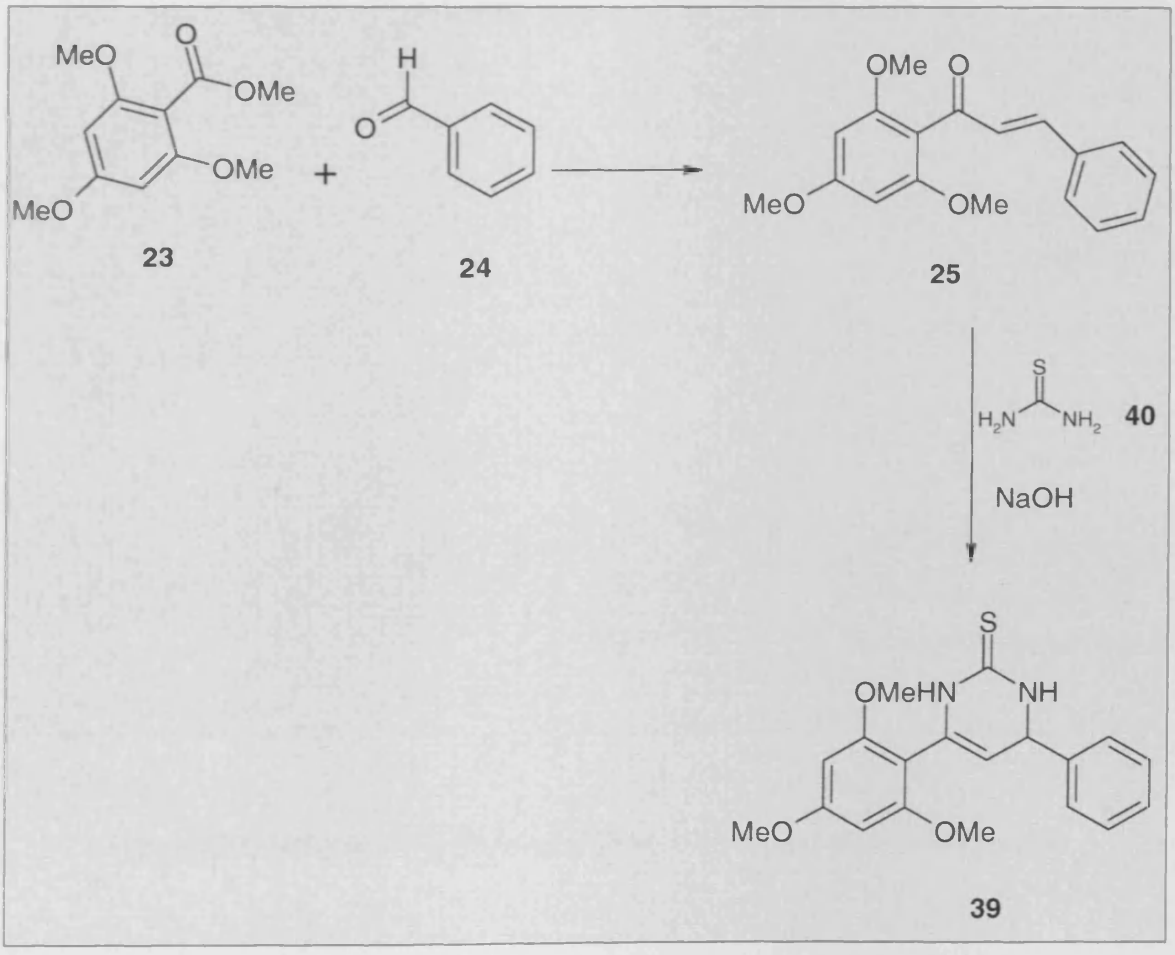


Figure 4.16 reaction mechanism of pyrazole carbothioamide formation.

4.6.4- Synthesis of 4-phenyl-6-(2, 4, 6-trimethoxyphenyl)-1, 2, 3, 4-tetrahydro-2-pyrimidinethione

The synthesis of 4-phenyl-6-(2, 4, 6-trimethoxyphenyl)-1, 2, 3, 4-tetrahydro-2-pyrimidinethione **39** was achieved in two steps (scheme 4.2). The first step was the Aldol condensation of 2, 4, 6-trimethoxy acetophenone with benzaldehyde in the presence of NaOH to give yellow crystals that were recrystalised from ethanol. The second step was the cyclization and formation of the pyrimidine ring by the reaction of the chalcone formed in the first step with Thiourea using NaOH as a base in ethanol to afford **39** in 45 % yield.



Scheme 4.2 Synthesis of compound **39**.

The mechanism of the cyclization of the pyrimidine ring from **25** could be explained as the **Michael addition reaction** [14] in which a nucleophile can react with α, β-unsaturated carbonyl compounds so as to give the conjugate addition product **I** (Figure .4.22). In this case the amino group of thiourea is activated in the presence of a base to act as a nucleophile. This reaction is a good reaction for the formation of C-C bonds, and it can be performed in ethanol and the base used was sodium hydroxide.

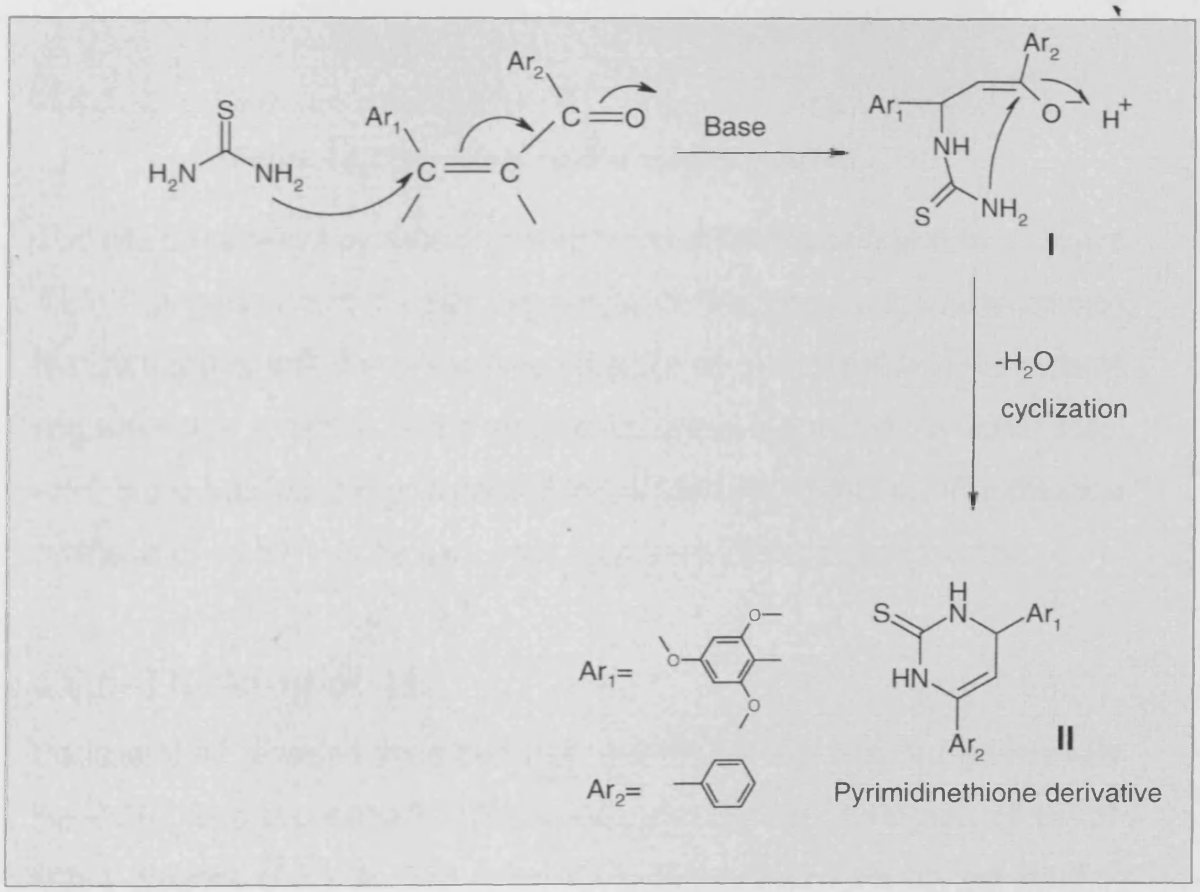
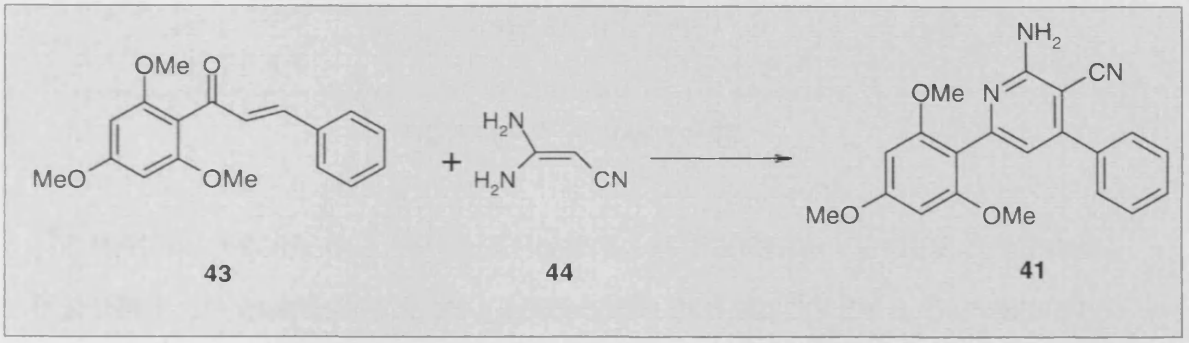


Figure 4.22 Mechanism of Michael addition reaction in pyrimidinethione synthesis

4.6.7- Synthesis of 2-amino-4-phenyl-6-(2, 4, 6-trimethoxyphenyl)-3-pyridyl cyanide.



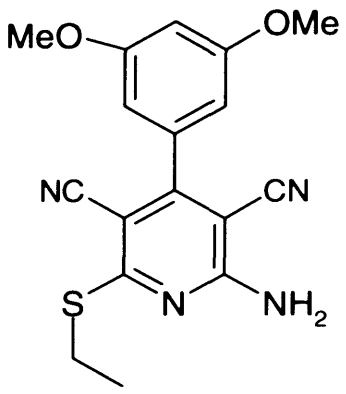
43 44 41

Scheme 4.3 showing the proposed synthesis of **41.**

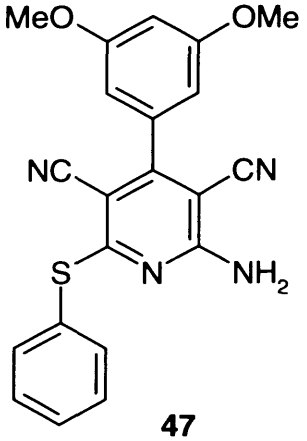
According to the reported method [15] the synthesis of this hit was supposed to be done by the reaction of chalcone **43** with enaminnitrile and here the enaminnitrile that could produce the desired product is 2-cyanoethanimidamine that could provide the amino group in the position 2 of the pyridine ring. Unfortunately **44** was not commercially available. A second pathway was to react **43** with 3-aminocrotononitrile **45** that was commercially available and the product was separated in 60 % after the reflux of both **43** and **45** in ethanol. Here the product **46** has a methyl group instead of the amino group found in **41.**

4.6.9-2-amino-4-(3,5-dimethoxyphenyl)-6-(ethylsulfanyl)-3, 5-pyridinedicarbonitrile 42.

The ethyl group in the 6 position of the pyridine ring of **42** was replaced by phenyl ring due to the availability of thiophenol to produce **47.** Docking of **47** against the DV NS3 protease **S1** pocket has shown some interesting interactions. For example the oxygen atoms of methoxy groups in the 3 and 5 positions of the phenyl ring have formed a hydrogen bond with –OH of Tyr 161 from one side and on the other side with –OH of Ser 135. The amino group at 2 position of the pyridine ring has shown a hydrogen bond with –C=O group of both Asp 129 and Phe130 at the same pose (Figure 4.28).



42



47

Figure 4.27 Chemical Structure of hit **42** and **47.**

The mechanism of this step involved the formation of a C=C double bond that could be observed in the intermediate **50.** The reaction mechanism in the first step is a base-catalyzed **Knoevenagel condensation reaction.**[18] The reaction is base catalyzed in order to activate the methylene in malononitrile to attack the carbonyl of 3, 5-dimethoxy benzaldehyde (Figure 4.29).

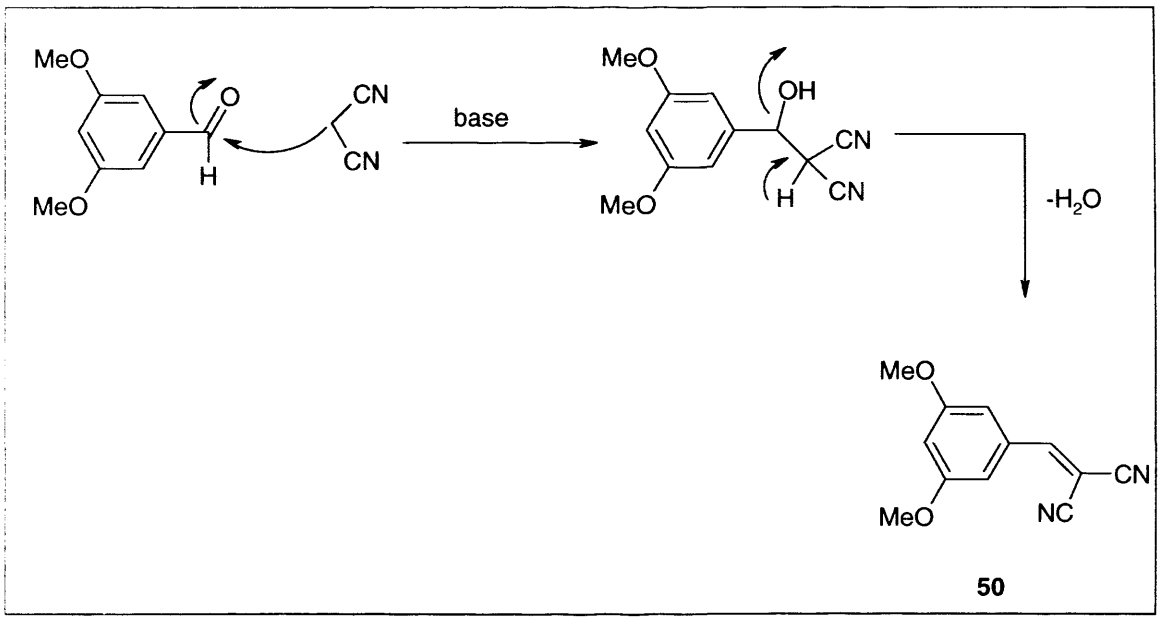
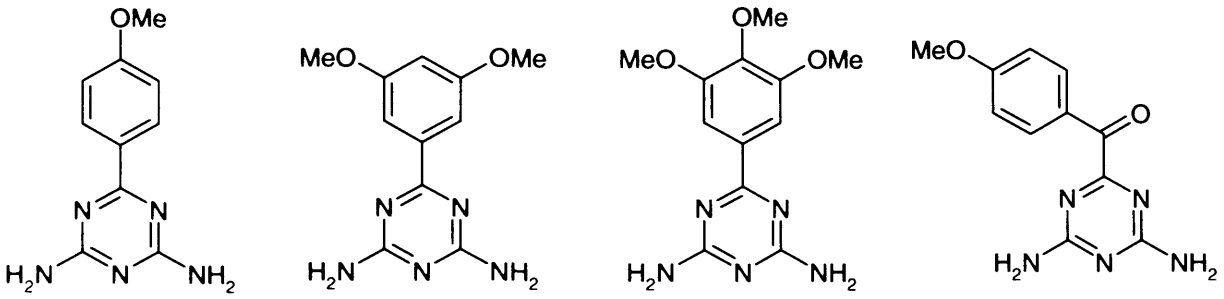


Figure 4.29 Mechanism of Knoevenagel adduct formation.

In the second step compound **50** was refluxed in malononitrile and thiophenol in the presence of triethylamine. After cooling, the precipitate was separated and dried to give **47.** Here thiophenol was added to one nitrile group of the first adduct, causing the formation of intermediate **II.** A second equivalent of malononitrile will be added to this intermediate in a 1,4- addition giving **III** which is expected to cyclise in the presence of a base to give a closed ring intermediate **IV,** the compound will then tautomerise to give **V** which will be oxidised to give the final pyridine-3,5-dicarbonitrile (Figure 4.30).[19]

Kambe et al [20] proposed that the final conversion from 1, 4-dihydropyridine to the final product occurred by the loss of molecular hydrogen from 1, 4-dihydropyridine. The oxidation is an aerobic oxidation that occurs after the reaction finished, cooled and stirred with exposure to air. **Evdokimov**[21], reported a different mechanism in which thiophenol, malononitrile, and the Knoevenagel adduct **I** will close a ring in one step. The product tautomerises to give 1, 4-dihydropyridine. And under base catalysis the Knoevenagel adduct **I** acts as an oxidizing agent for the 1, 4-dihydropyridine and Knoevenagel adduct **I that** will be reduced. The following scheme illustrates this mechanism (Figure 4.31).



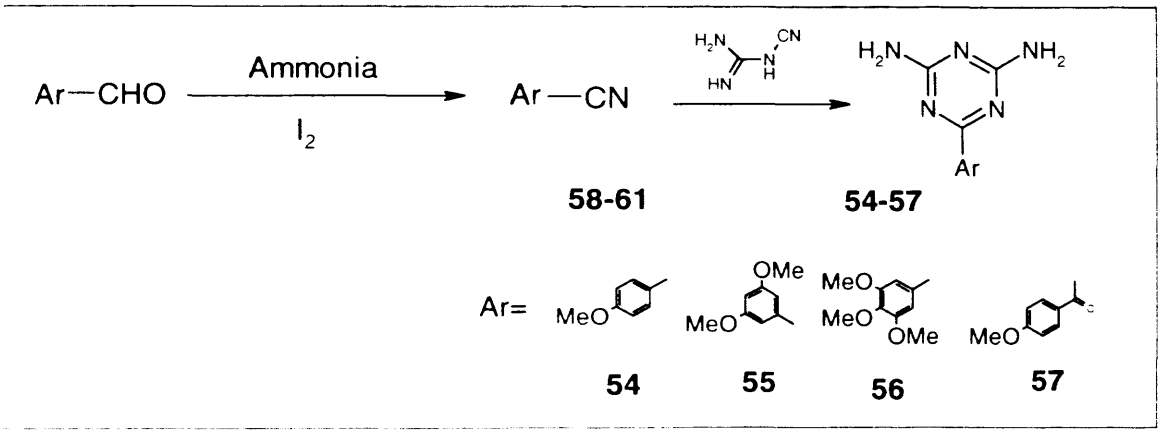
54 **55** **56** **57**

Figure 4.34 A group of compounds containing different substitutions of the methoxy group on the phenyl ring.

A number of compounds were built and docked using MOE. Docking of **54** showed an improved scoring and the expected interaction of the Ser 135 –OH group with the recently introduced 4-methoxy group was achieved with a score of 62.3 % and a distance of 2.87 A°. Also, the hydrogen bond formation between an amino group and the –C=O of Asp 129 with a score 10.5 % and a distance of 2.67 A° was observed (Figure 4.35).

4.6.12- Synthesis of 6-(substituted phenyl)-1, 3, 5-triazine-2, 4-diamine

According to the literature substituted 1, 3, 5-triazine-2, 4-diamine is prepared in a two step route. In the first step an aldehyde is reacted with ammonia in the presence of iodine to give a nitrile. This is then condensed with cyanoguanidine (dicyandiamide) to give the 1, 3, 5-triazine.[23-24]



Scheme 4.6 synthesis of 2, 4-diamino-(1, 3, 5-substituted)-triazines.

The previously reported method was tried for the synthesis of **54**, **55**, and **56** starting with 4-methoxy benzaldehyde, 3, 5-dimethoxy benzaldehyde and 3, 4, 5-trimethoxy benzaldehyde. But, the yield of this method was too low for the first step. The second choice was to begin from the aryl nitriles by reacting them with dicyandiamide. A mixture of dicyandiamide (1.1 mmole) and KOH (2.2 mmole) was added to 1mmole of the nitrile (4-methoxy benzonitrile **58**, 3, 5-dimethoxy benzonitrile **59**, 3, 4, 5-benzonitrile **61** and 4-methoxy benzoyl nitrile **60**) in THF and was refluxed and monitored with TLC. Upon cooling the formed precipitate was collected, washed and dried. The direct reaction between the nitrile and the cyanoguanidine was successful in a yield 73 %.

The second step of the mechanism involves the reaction of thionyl chloride with the previously formed alcohol in order to get the chloro derivative. This takes place in five steps. In step **I,** (Figure 4.46) the oxygen atom of the alcohol acts as a nucleophile to attack the highly electrophilic sulfur atom of thionyl chloride. Then in step **II,** the presence of a base will remove the proton from the alcoholic oxygen atom. In step **III** the intermediate is rearranged to form the S=O and a chloride ion is lost as a leaving group. In step **IV** an SN2 reaction occurs in which the chloride ion acts as a nucleophile on the electrophilic carbon from the alcohol to displace the SO_2 group that will be a leaving group and release another chloride ion and the final chloride derivative of our alcohol will be formed **V**.

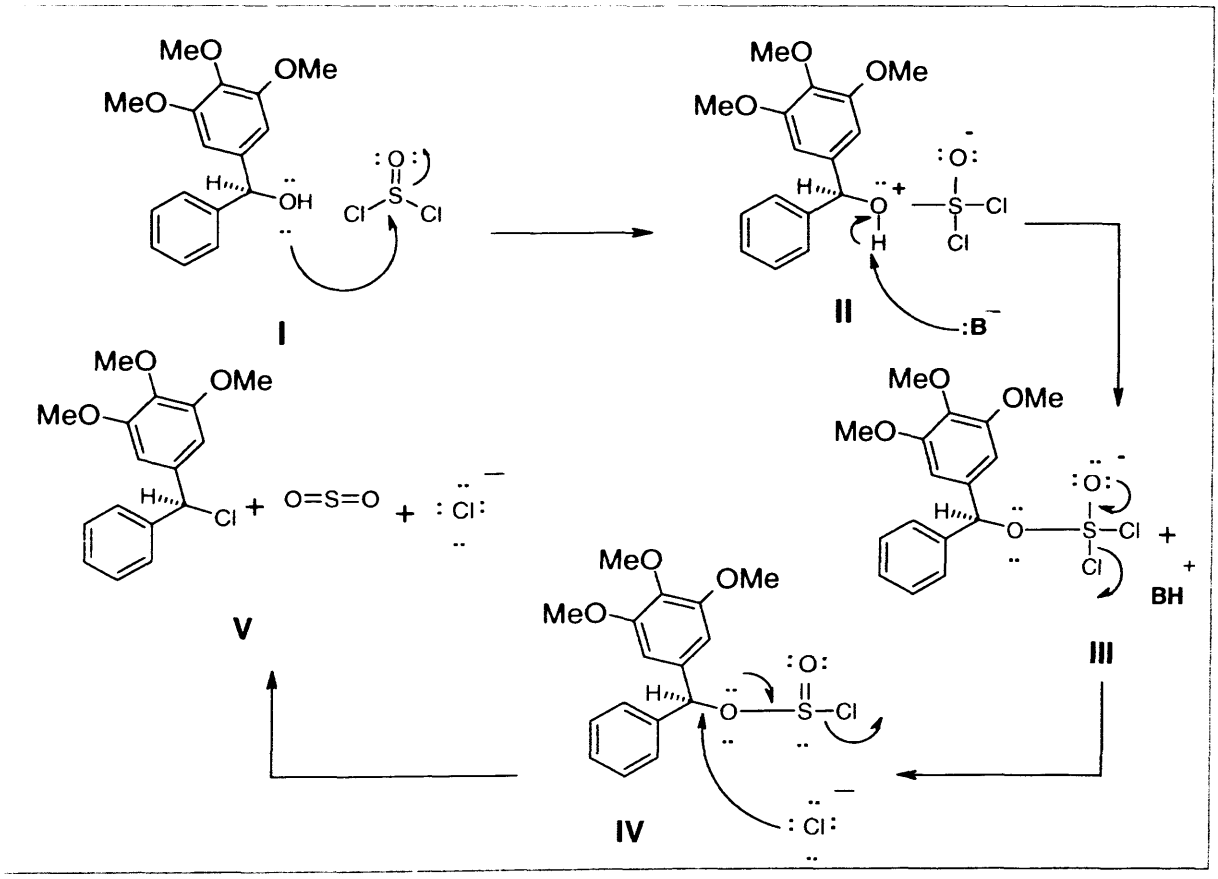


Figure: 4.46 Mechanism of thionyl chloride reaction with alcohols.

The mechanism of such a reaction could be explained by the reaction of aldehyde and primary amine to give the C=N bond of the Schiff's base which is attacked by the C=C at position 3 of the reagent to give an intermediate which is arranged so that the lone pair on the –NH will attack the ester –C=O carbon to close the pyrrolin-2-one ring and liberate the alcohol.[33]

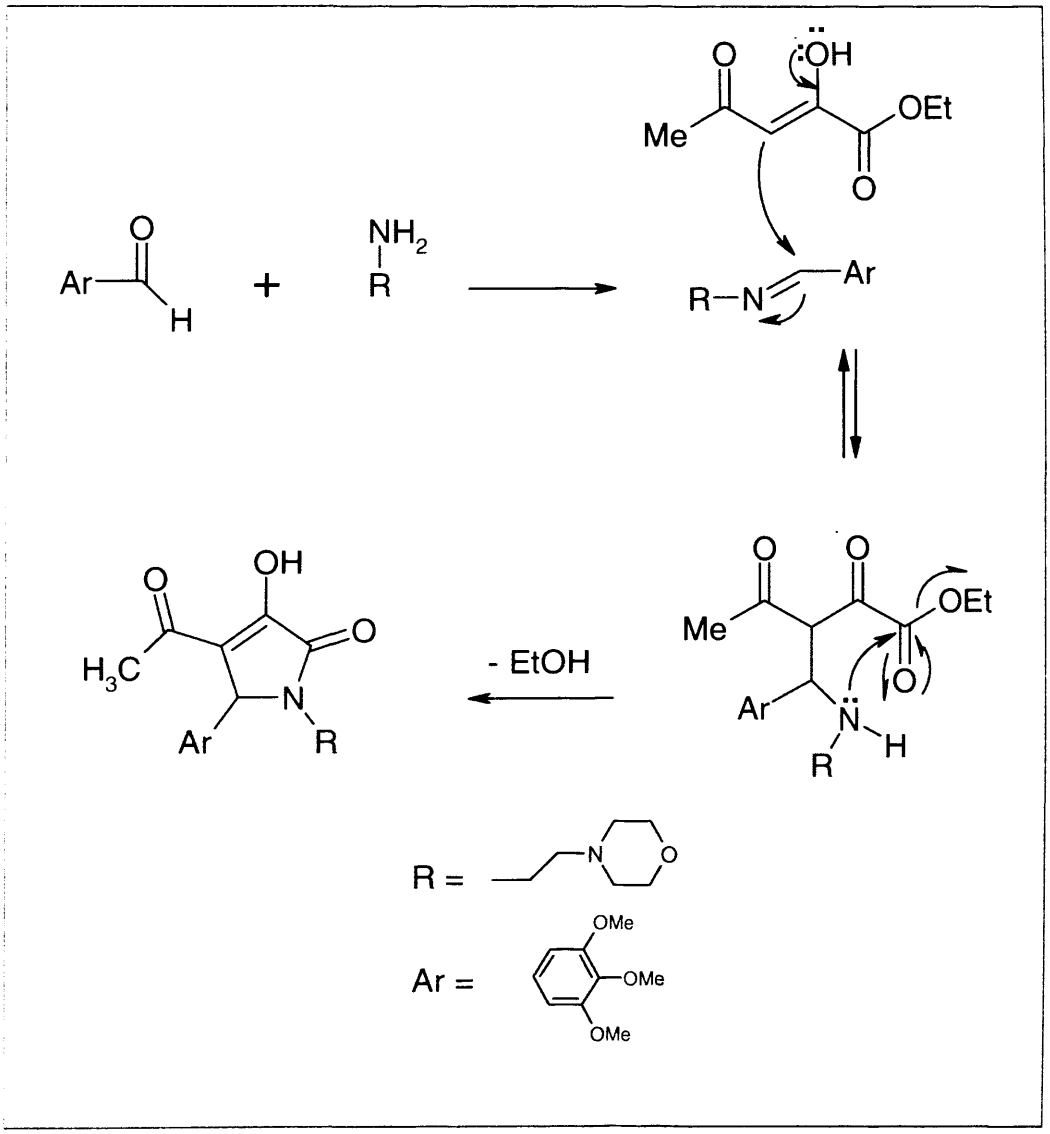


Figure 4.58 Possible mechanism for the synthesis of 5-(3, 4, 5-trimethoxy phenyl)-4-acyl-3-hydroxy-1-morpholinoethyl-3-pyrrolin-2-one

4.6.26- References.

1. Douglas, B. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discovery.* **2004,** *3,* 935-949.

2. Krishna, M.; Padmanabhan, R. Crystal structure of dengue virus NS3 protease in complex with a bowman-birk inhibitor: Implications for flaviviral polyprotein processing and drug design. *J. Mol. Biol.* **2000**, *301* (4), 759-767.

3. Vannakambadi, K. Identification and characterization of non-substrate based inhibitors of the essential dengue and west nile virus proteases. *Bioorg. Med. Chem.* **2005**, *13,* 257-264.

4. Erbel, P.; Schiering, N. Structural basis for the activation of flaviviral NS3 protease from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **2006,** *13,* 372-373.

5. Krishna, M.; Padmanabhan, R. *J. Biol. Chem.* **1999**, *274,* 5573-5580.

6. Krishna, M.; John, E.; Knox, Wai, Y. P.; Ngai, L. M. Yellow fever virus NS3 protease peptide-inhibition studies.*J. Gen. Virol.* **2007**, *88,* 2223-2227.

7. Zheng, Y.; Sejal, J. P.; Wei-Ling, W. Peptide inhibitors of dengue virus NS3 protease part 2: SAR studies of tetrapeptide aldehyde inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16,* 40-43.

8. John, E. K.; Ngai, L. M.; Zheng, Y.; Sejal, J. Peptide inhibitors of West Nile virus NS3 protease: SAR study of tetrapeptide inhibitors. *J. Med. Chem.* **2006**, *49,* 6585-6590.

9. Molecular Operating Environment. (MOE), Chemical computing group. Inc. Montreal, Quebec, Canada.

10. Niklaus, H.; Nagarajan, P.; Camilo, A. S.; Prasanth, V. Padmanabhan, R. *Antimicrob. Agents Chemother.* **2008,** *52* (9), 3385-3393.

11. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.

12. Ahmet, O.; Gulhan, T.; Zafer, A. K.; Gilbert, R. Synthesis and antimicrobial activity of 1-(4-aryl-2-thiazolyl)-3-(2-thienyl)-5-aryl-2-pyrazoline derivatives.Eur. *J. Med. Chem.* **2007,** *42*, 403-409.

13. John McMurry. Organic Chemistry. **1999.**

14. Shivarama, H. B. One pot synthesis of thiazolodihydropyrimidinonesand evaluation of their anticancer activity. *Eur. J. Med. Chem.* **2004,** *39*, 777-783.

15. Thomas, L.; Andreas, P. Named organic reactions. Wiley. Volkswagen AG, Wolfsburg, Germany **2005.**

16. Bossert, F.; Meyer, H.; Wehinger, E. *Angew. Chem., Int. Ed.* **1981,** 93, 755-763.

17. Margot, W.; Beukers; Lisa, C. W.; Jacobien, K. New, non-adenosine, high potency agonists for the human adenosine A_{2B} receptor with an improved selectivity profile compared to the reference agonist N-ethylcarboxamidoadenosine. *J. Med. Chem.* **2004,** *47* (15), 3707-3709.

18. Paun, C. Supported and liquid phase task specific ionic liquids for base catalyzed Knoevenagel reactions. *J. Mol. Catal. A: Chem.* **2007,** *269*, 64–71.

19. Kai, G.; Mark, J. Mechanistic studies leading to a new procedure for rapid, microwave assisted generation of pyridine-3, 5-dicarbonitrile libraries. *Tetrahedron.* **2007,** *63*, 5300-5311.

20. Kambe, S.; Saito, K.; Sakurai, A.; Midorikawa, H. Activated nitriles in heterocyclic synthesis. *Synthesis* **1981**, *7*, 531–533.

21. Evdokimov, N. M.; Magedov, I. V.; Kireev, A. S.; Kornienko, A. One-step, three-component synthesis of pyridines and 1,4-dihydropyridines with manifold medicinal utility. *Org. Lett.* **2006**, *8*, 899–902.

22. Kai, G.; Mutter, R.; Heal, W.; Reddy, T. K. R.; Cope, H. Pratt, S. Thompson, M. J. Chen, B. Synthsis and evalution of a focused library of pyridine dicarbonitriles agonists prion disease. *Eur. J. Med. Chem.* **2008**, *43*, 93-106.

23. Jium, J. S.; Jim, M. F. Direct conversion of aldehydes to amides, tetrazoles, and Triazines in aqueous media by one-pot tandem reaction. *J. Org. Chem.* **2003,** *68*, 1158-1160.

24. Jan, J.; Ryszard, K. Transformation of one of two CN groups of o-dicyanobenzene in the presence of cyanoguanidine. Crystal and gas-phase structure of 2-(20-cyanophenyl)-4,6-diamino-1,3,5-triazine. *J. Mol. Struct.* **2005,** *749* (1-3), 60-69

25. Hiroshi, O. Benzylpiperazine derivatives II. Synthesis and cerebral vasodilating activities of 1-[(3-alkyl-3-hydroxy-3-phenyl) propyl]-4-benzylpiperazine derivatives. *Chem. Pharm. Bull.* **1987,** *35* (7), 2782-2791.

26. Webbers, V. J.; Bruce, W. F. The Leuckart reaction: A study of the mechanism. *J. Amer. Chem. Soc.* **1948**, *70,* 1422.

27. Staple, E.; Wagner, E. A study of the Wallach reaction for alkylation of amines by action of aldehydes or ketones and formic acid. *J. Org. Chem.* **1949**, *14*, 559.

28. Sandra, G.; Giusepe, C.; Stefania, B.; Gogan, K.; Salvatore, S. E. Clotrimazole scaffold as an innovative pharmacophore towards potent antimalarial agents design, synthesis, and biological and structural-activity relationship studies. *J. Med. Chem.* **2008,** *51*, 1278-1294.

29. McMurry, J. Organic Chemistry. 5th ed. Brooks/Cole, **1999.**

30. Chai, H. L; James, D. K; Harold, K. 3-oxo- and 3-amino-4-substituted-1, 2, 5-thiadiazolidine-1, 1-dioxides. Synthesis, spectral properties and selected chemistry. *J. Org. Chem.* **1989,** *54* (13), 3077-3083.

31. Coker, J. N.; Kohlhase, W. L.; Matrens, T. F. A General route to hydantoins. *J. Org. Chem.* **1962,** *27,* 3201.

32. Gein, V, L. Synthesis and biological activity of 5-aryl-4-acyl-3-hydroxy-1-morpholinoalkyl-3-pyrrolin-2-ones. *Pharm. Chem. J.* **2007,** *41* (5), 256-263.

33. PhD thesis; Synthesis van 5-aryloxopyrrool derivaten en exploratie ter vorming van potentiele NNRTI's, flouroforen en op APTRA gebaseerde Ca2+ indicatoren. **2005,** Katholieke University, Leuven.

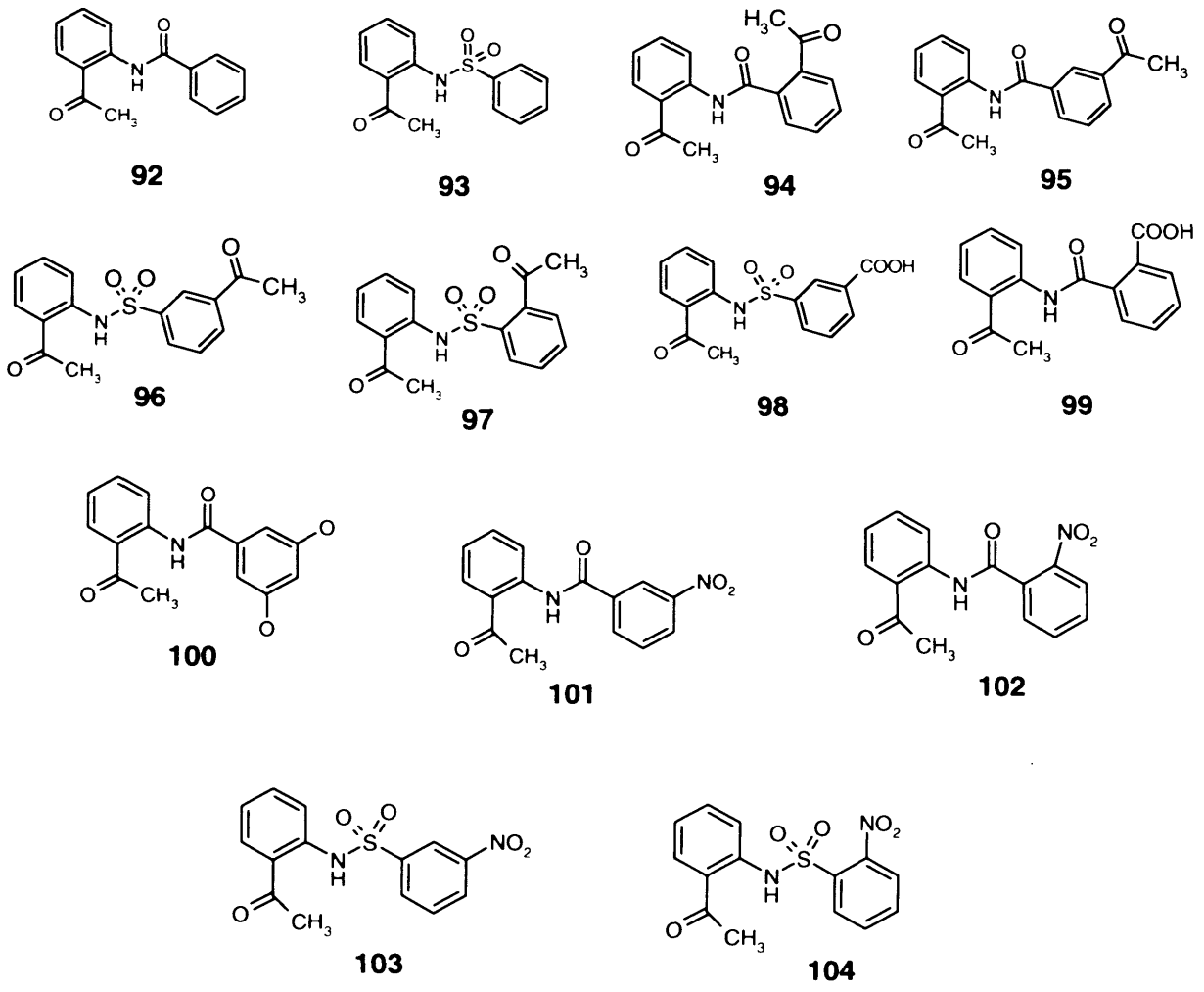


Figure 5.11 A library of 13 designed compounds.

5.4- Docking of designed compounds.

Docking of the designed library was done using the MOE in which the previously identified allosteric site was used for docking and the rescoring was done by London dG. After docking visualization of the best poses of each compound was carried out in order to indicate the best mode of binding. Table 5.1 contains the chemical structure, score of the best pose for every docked compound in Kcal/mol and an image for that pose.

5.5- Results and Discussion of Docking.

According to the docking results and visualization of the best pose of each compound from the designed library it was found that the carbonyl group of the 2-acetyl moiety found in the 2-amino acetophenone seed nucleus was involved in some interactions forming hydrogen bonding with $-NH_2$ of Arg 737 as in **94** and sometimes with –OH group of Ser 710, Thr 794, Trp 795 and Ser 796 as in case of **95, 96, 100** and **102** respectively. The presence of an electron accepting group like the carbonyl group in the substituted benzamide moiety and sulfonyl group in the substituted benzene sulfonamide moiety was important in making good interactions with some residues within the specified pocket. For example, the carbonyl group was found to form hydrogen bonds with Arg 737, 729, Thr 794 as found in compounds **92, 94, 95, 99, 100, 101, 102**. On the other hand, the sulfonyl group was found interacting with –OH group of Thr 794 as in **93**, the -OH of Ser 710 as in **96** and –OH group of Ser 796 and $-NH_2$ group of Arg 729 as in **98.** The carboxylic group in the meta position of benzene sulfonamide was important for the $-NH_2$ of both Arg 729 and 737 as in **98.** However, that in the ortho position was important for interacting with the –OH of Thr 794 and the –NH of Trp 795 as in **97**. The oxygen atom of 3, 5-dimethoxy groups in **100** formed a hydrogen bond with –OH of Ser 796. The nitro group in **101, 102, 103,** and **104** was not involved in any kind of interactions according to our docking. According to these results and according to the availability of commercial reagents we decided to synthesize compounds **92, 93, 98, 99** and **100** aiming to find a potential selective inhibitor for DV RdRp.

5.10- References.

1. Nittoli, T. Identification of anthranilic acid derivatives as a novel class of allosteric inhibitors of hepatitis c NS5B polymerase. *J. Med. Chem.* **2007,** *50*, 2108-2116.
2. Yap, T, L. Crystal structure of the dengue virus RNA-dependant RNA-polymerase catalytic domain at 1.85 A°. *J. Virol.* **2007,** *81*, 2753.
3. Molecular Operating Environment (MOE). Chemical computing group Inc. Montreal, Quebec, Canada.
4. Steven, M. J.; Stephen, C.; Jan, A. W. Toward optimization of the second aryl substructure common to transthyretin amyloidogenesis inhibitors using biochemical and structural studies. *J. Med. Chem.* **2009,** *52*, 1115-1125.
5. Xiaohu, D.; Neelakandha, S. M. A facile environmentally benign sulfonamide synthesis in water. *Green Chem.* **2006**, *8*, 835-838.
6. Sarala, N. Mild and eco-friendly chemo selective Acylation of amines in aqueous medium. *ARKIVOC.* **2004 (i),** 55-63.
7. Zheng, Y. Yen, L. C. N-Sulfonylanthranilic acid derivatives as allosteric inhibitors of Dengue viral RNA-Dependant RNA Polymerase. *J. Med. Chem.* **2009.** 52 (24), 7934-7937.

Chapter 6

Biological Assays

Chapter 6

6.1- HCV Replicon Assay.

Replicon is a genetic element, which can be either DNA or RNA that can replicate under its own control in a cell. [1] Since viruses are obligate intracellular parasites, the efficacy of an antiviral drug is usually evaluated in a cell culture system. Unfortunately, hepatitis C virus isolates taken from patients usually replicate poorly in cell culture. [1] Initially HCV replicons were replicating in genetically engineered HCV RNA "mini-genomes" in which the regions that encodes the core to NS2 is replaced by a selectable marker and an internal ribosome entry site (IRES) that mediates translation of HCV replicase (NS3-5B). Transfection of this RNA in cells of human hepatoma cell line Huh-7, followed by selection results in HCV clones. Recently, many different HCV replicons have been developed that allow screening of chemical compounds. [2] Replicon development had gone through the following stages;

6.1.1- Cell Culture Propagation of HCV.

In the ideal case, a virus can be propagated in the laboratory by infection of cultured cell lines that are readily available. For unknown reasons, propagation of HCV in primary human hepatocytes has been suffering from low reproducibility and efficacy. [3-4] this low efficacy made the specific detection of HCV viral antigens or RNA difficult. As a further complication, primary cells were not readily available, and the efficiency of infection depends on the quality of the cells, which is a parameter that is difficult to control. So, the usefulness of these systems for drug development is limited. [1]

6.1.2- Establishment of the first HCV Replicon.

However, encouraged by results from other positive-strand RNA viruses, which showed that the structural proteins are not essential for RNA replication, an alternative strategy based on the construction of subgenomic, selectable replicons was devised. [5-8] In these genetically modified HCV "minigenomes" the region that encodes the structural proteins by a selectable marker; the neo gene encoding the enzyme neomycin phosphotransferase (NPT), which inactivates the cytotoxic drug G418 (geneticin; an aminoglycoside antibiotic for eukaryotic cell selection) was used. [1]

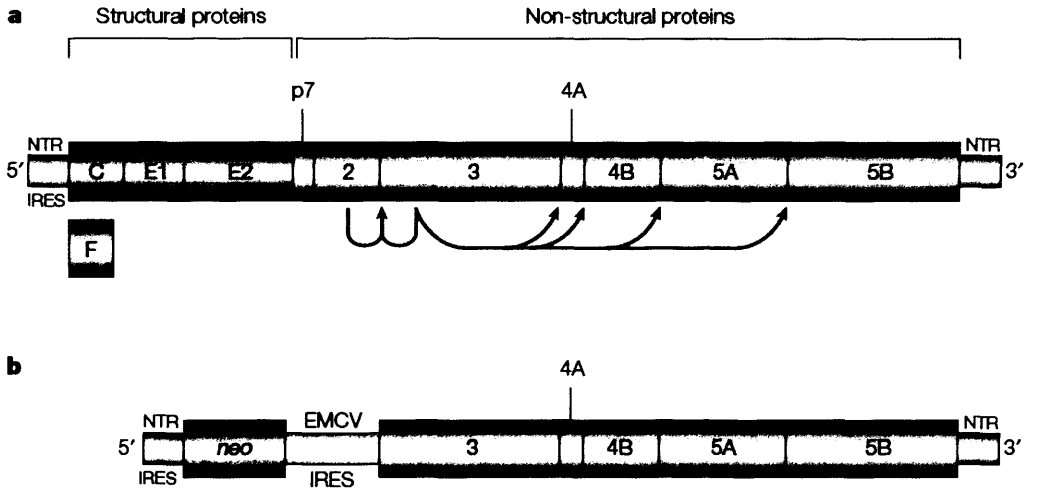


Figure **6.1** A) HCV genome encoding the core structural protein C and the envelope glycoproteins (E1 and E2), the non-structural (NS) proteins NS2 to NS5B is in color. B) A subgenomic replicon by replacing up to the NS2-encoding region by the neomycin phosphotransferase gene (neo) and the internal ribosome-entry site (IRES) of another virus (encephalomyocarditis virus; EMCV) [1]

These replicons are called bicistronic since they consists of two genetic units that are expressed as two proteins, Neomycin phosphotransferase (NPT) mediated by the HCV internal ribosome-entry site (IRES), whereas a second IRES of another virus (encephalomyocarditis virus; EMCV) is to direct the expression of the HCV replication proteins (NS3 to NS5B) [1] after Transfection of the human hepatoma cell line Huh-7 with the subgenomic replicon RNA and subsequent selection with G418, only cells in which the replicon was amplified to high levels expressed sufficient amounts of NPT, therefore survive into a colony that can be isolated and expanded. [1]

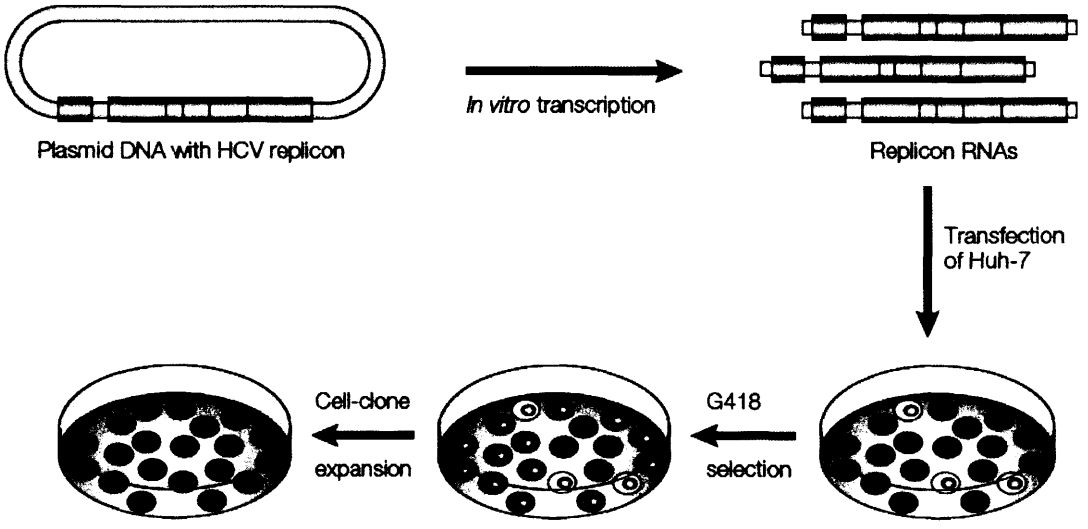


Figure 6.2 Establishment of the cell clones. Cells that did not take up the RNA (white) and cells that does not replicate (orange) will die because of the toxic effect of G418.

6.1.3- Breakthrough for HCV research.

In the past decade, HCV replicon systems have viral molecular biology and virus-host interactions to be probed [6-7] however, these systems can not replicate in vitro without acquiring adaptive mutations, nor do they produce infectious virions [9-11] more recently, a system that replicates a full-length RNA without acquired mutations was developed and hence is representative of the wild type infectious HCV virions. [12-14] Moreover, RNA replication was measured originally by quantifying the amount of HCV RNA or protein in a cell but, the insertion of a reporter gene, such as firefly luciferase has made this process much easier by measuring its activity without the time-consuming selection for stable cell clones. [15-22]

6.1.4- Application in drug development.

HCV replicon is a powerful tool to unravel the principles of HCV replications. Despite this several considerations should be made when using it. First, since replication of these RNAs depends on cell proliferation, compounds that interfere with cell growth lead to an apparent inhibition of the replicon; however, by using assays for cell metabolic activity, such false positive hits can be excluded. Second, by using subgenomic replicons, compounds that work by interfering with HCV structural proteins might be missed. With the availability of full-length genomes this possibly can be examined. [23-24]

As a result of the previous assay some compounds showed low micro molar inhibition of HCV replicon. The most potent compound **62** that has EC_{50} value of 1.1 µM that could prove that piperazine based scaffold may be a promising for antiviral activity. Compounds **76, 77** and **78** with the substituted-phenyl methylidene-1-hydrazine carbothioamide scaffold showed EC_{50} values of 5.5 µM and 17.9 µM and 22.6 µM respectively. Compound **31** showed EC_{50} value of 23 µM. Compound **98** showed EC_{50} value of 8.7 µM. Compound is **47** with the pyridine dicarbonitrile scaffold showed EC_{50} value of 5.6 µM.

6.2- Dengue Assay.

All compounds were tested for their antiviral activity against dengue virus. The assay for the prepared compounds was carried out in the Rega Institute for Medical Research, KULeuven, Leuven, Belgium, under the supervision of Professor Johan Neyts. Only five compounds showed low micro molar activity and the results are shown in table 6.2.

Compound	Observed activity µM
39	faint CPE reduction at 140
46	faint CPE reduction at 5
47	A delayed CPE reduction at 128
78	Normal growth at 209
88	a clear CPE reduction at 138
Ribavirin	Normal growth at 81

Table 6.2 Inhibition of DV replication in the dengue assay, showing the observed activity of 5 compounds in µM compared to Ribavirin as a standard.

Compound **39** has a faint cytopathic effect (CPE) reduction at 140 µM. Compound **46** has a cytostatic effect at 138 µM and a faint CPE reduction at 5 µM. Compound **47** showed a cytostatic effect at 128 µM and a delayed CPE reduction at 128 µM. Compound **78** showed a delayed CPE and the cells appear to have grown normally at 209 µM. Compound **88** has a clear CPE reduction and cells appear to have grown normally at 138 µM. Ribavirin has a cytostatic effect at 409 µM and cells appear to have grown normally at 81 µM. The clear CPE reduction of compound **88** can be used as a start point for developing a new potent inhibitor. Also, compounds **47** and **78** showed antiviral activity for both HCV and DV that could have a good effect on WNV as well. The presence of 3,4,5-trimethoxy phenyl moiety and 2,4,6-trimethoxy phenyl moiety was important for interacting with Ser 135, Tyr 150, and His 51 of the **S1** pocket. The cyano group found in compounds **46** and **47** may have good hydrogen formation with –NH of PHE 130. As a result pyrimidinethione, pyridine, imine and pyrrolin-2-one based scaffolds showed their ability to be used for the inhibition of potential compounds against DV.

6.5- References.

1. Bartenschlager, R. Hepatitis C virus replicon: Potential role for dug development. Nature reviews drug discovery. **2002,** 1, 911-916.
2. Mc Hutchion, J.; Bartenschlager, R.; Patel, K.; Pawlotsky, J. The face of future hepatitis C antiviral drug development: recent biological and virological advances and their translation to drug development and clinical practice. J. Hepatology. **2006,** 44, 411-421.
3. Bartenschlager, R. Novel cell culture systems for the hepatitis C virus. Antiviral Res. **2001,** 52, 1-17.
4. Kato, N.; Shimotohno, K. Systems to culture hepatitis C virus. Curr. Top. Microbiol. Immunol. **2000,** 242, 261-278.
5. Blight, K.; Kolykhalov, A.; Rice, C. Efficient initiation of HCV RNA replication in cell cuture. Science. **2000,** 290, 1972-1974.
6. Krieger, N.; Lohmann, V.; Bartenschlager, R. Enhancement of hepatitis C virus RNA replication by cell culture adaptive mutations. J. Virol. **2001,** 75, 4614-4624.
7. Lohmann, V.; Korner, F.; Bartenschlager, R. Mutations in hepatitis C virus RNAs conferring cell culture adaptation. J. Virol. **2001,** 75, 1437-1449.
8. Guo, J.; Bichko, V.; Seeger, C. Effect of α-interferon on the hepatitis C virus replicon. J. virol. **2001,** 75, 8516-8523.
9. Frese, M.; Pietschmann, T.; Moradpour, D.; Haller, O.; Bartenschlager, R. Interferon α inhibits hepatitis C virus subgenomic RNA replication by an MxA-independent pathway. J. Gen. Virol. **2001,** 82, 723-733.
10. Pietschmann, T. et al. Persistent and transient replication of full length hepatitis C virus genomes in cell culture. J. virol. **2002,** 76, 4008-4021.
11. Jones, S. Antiviral drugs: Breakthrough for hepatitis C research. Nature reviews drug discovery. **2005,** 4, 629.

12. Wakita, W. et al. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. Nature Med. **2005,** 11, 791-796.

13. Lindenbach, B. et al. Complete replication of hepatitis C virus in cell culture. Science. **2005,** 309, 623-626.

14. Zhong, J. et al. Robust hepatitis C virus infection in vitro. Proc. Natl Acad. Sci. **2005,** 102, 9294-9299.

15. Friebe, P.; Krieger, N.; Lohmann, V.; Bartenschlager, R. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. J. virol. **2001,** 75, 12047-12057.

16. Friebe, P.; Bartenschlager, R. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. J. virol. **2002,** 76, 5326-5338.

17. Cheney, I. et al. Mutations in the NS5B polymerase of hepatitis C virus. Virology. **2002,** 297, 298-306.

18. Frese, M. et al. Interferon-γ inhibits replication of subgenomic and genomic hepatitis C virus RNAs. Hepatology. **2002,** 35, 694-703.

19. Tardif, K.; Mori, K.; Siddiqui, A. Hepatitis C subgenomic replicons induce endoplasmic reticulum stress activating an intracellular signaling pathway. J. virol. **2002,** 67, 7453-7459.

20. Pflugheber, J. et al. Regulation of PKR and IRF-1 during hepatitis C virus RNA replication. Proc. Natl Acad. Sci. **2002,** 99, 4650-4655.

21. Bartenschlager, R. Efficient hepatitis C virus cell culture system. Proc. Natl. Acad. Sci. USA. **2005,** 102, 9739-9740.

22. Yi, M.; Bodola, F.; Lemon, S. Subgenomic hepatitis C replicons including the expression of a secreted enzymatic reporter protein. Virology. **2002,** 304, 187-210.

23. Pietschmann, T.; Lohmann, V.; Rutter. G.; Bartenschlager, R. Characterization of cell lines carrying self replicating hepatitis C virus RNAs. J. Virol. **2001,** 75, 1252-1264.

24. Young, S. Inhibition of HIV-1 integrase by small molecules: the potential for a new class of AIDS chemotherapeutics. Curr. Opin. Drug Discov. Dev. **2001,** 4, 402-410.

25. Paeshuyse, J.; Vliegen, I.; Coelmont, L..; Leyssen, P.; Neyts, J. Comparative In vitro Anti-hepatitis C virus activities of a selected series of Polymerase, Protease, and Helicase inhibitors. Antitumor agents and chemotherapy. **2008,** 52, 9, 3433-3437.

Chapter 7

Conclusion & Future Work

Chapter 7

The first aim of this work was to modify the structure of Panduratin A and 4-hydroxypanduratin A [1] aiming to find a better inhibition of the dengue virus replication. Docking of these compounds revealed some interactions and some suggestions for the modification of the structure. The 2-methyl-1-propenyl side chain was suggested to be excluded from the structure. 4, 6-Dihydroxy groups were replaced by methoxy groups. The resulting compound was predicted to be synthesized in a two step reaction. The first step involved the synthesis of 3-phenyl-1-(2, 4, 6-trimethoxyphenyl) -2-propen-1-one by Aldol condensation. The second step involved the reaction of 3-phenyl-1-(2, 4, 6-trimethoxyphenyl) -2-propen-1-one with 1, 3-dimethyl butadiene to close the cyclohexenyl ring via the Diels Alder reaction. However, the second step failed.

The second aim of this work was to design small potential inhibitors of DV NS3 serine protease. A similarity search was done to find a similar crystal structure for DV NS3 protease. The crystal structure of WNV NS3 protease was found to be the highly similar to DV NS3 protease. A superimposition of WNV NS3 protease complexed with a peptide inhibitor (pdb code = 2FP7) [2] together with the DV NS3 protease (pdb code = 2FOM) was done. The peptide inhibitor was fixed within the DV NS3 protease crystal structure and all possible interactions with **S1** pocket residues were visualized. A pharmacophore model was built that has two cationic/donor features and two acceptor features. Structure-based VS was applied by searching drug-like databases. Filtration of hits was carried out according to the Lipinski rule of five, and RMSD values. As a result a number of hits that had a variety of chemical scaffolds were docked and selected for chemical synthesis after some modifications required for some hits.

7.1- Compounds with antiviral activity against the Dengue virus.

All compounds were tested against the dengue virus. Compounds **39, 46, 47, 78,** and **88** showed reduction of cytopathic effect of 140 μM, 5 μM, 128 μM, 209 μM, and 138 μM respectively against DV. As a result pyrimidinethione, pyridine, imine and pyrrolin-2-one based scaffolds showed their ability to be used for the inhibition of DV.

7.2- Compounds with antiviral activity against HCV.

Compounds **62, 76, 98,** and **47** showed some activity against HCV replicon assay of 1.1 μM, 5.5 μM, 8.7 μM, and 5.6 μM respectively. The activity of compound **98** may be explained due to its similarity with anthranilic acid derivatives that act on the HCV RdRp. Compound **47** was found to be of an interesting antiviral activity against both the Dengue virus and HCV.

7.3- Conclusion.

The use of molecular modeling approaches such as structure-based virtual screening, molecular alignment and docking was successful in obtaining some compounds of low micro molar activity against HCV and DV. This could be a successful starting point for further development of potent inhibitor. Compound **98** has a similar scaffold to anthranilic acid derivatives that inhibit RdRp and it showed an inhibition of HCV = 8.7 μM. Pyridine dicarbonitrile scaffold found in compound **47** showed an activity against HCV = 5.6 μM and against DV = 128 μM. The (substituted-phenylmethylidene)-1-hydrazine carbothioamide scaffold in **76, 77,** and **78** showed an activity against HCV of 5.5 μM, 17.9 μM and 22.6 μM respectively. In addition, **78** showed 209 μM activity against DV. Compound **88** showed a clear inhibition of DV = 138 μM. The presence of 2, 4, 6-trimethoxy phenyl moiety was observed as a common feature in compound

39, and **46** that showed activity against DV. However, the 3,5-trimethoxy phenyl moiety was a common feature in **47** and **78.**

7.4- Future work.

Further study and modification may be required for the active compounds that showed activity against DV to get a potent inhibitor. In addition to the previously studies DV NS3 protease enzyme and DV RdRp the importance of the helicase enzyme in the virus maturation makes it an interesting target for drug discovery. The discovery of potent HCV helicase inhibitors [3] was successful. In the future work we will be looking for a comparative study between different flaviviral Helicases in order to study the main difference between them and trying to dock some of the HCV helicase inhibitors against the other crystal structures. The crystal structure of the helicase enzyme of HCV, DV, and YFV were reported in pdb. While that of WNV was not published to date. A homology model of WNV NS3 Helicase was built using MOE program. [4] The future work will include this comparative study aiming to find a suitable design of potential inhibitors for DV, WNV Helicases.

7.5- References.

1. Tan, S. K.; Richard, P.; Rohana, Y.; Halijah, I.; Norzulaani, K.; Noorsaadah, A. Inhibitory activity of cyclohexenyl chalcone derivatives and flavanoids of fingerroot, Boesenbergia rotunda (L.) towards dengue-2 virus NS3 protease. *Bioorg. Med. Chem. Lett.* **2006,** *16:* 3337-3340.
2. John, E. K.; Ngai, L. M.; Zheng, Y.; Sejal, J. Peptide inhibitors of West Nile virus NS3 protease: SAR study of tetrapeptide inhibitors. *J. Med. Chem.* **2006**, *49*, 6585-6590.
3. Kandil, S.; Brancale, A. Discovery of a novel HCV helicase inhibitor by a de novo design approach. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2935-2937.
4. Khedr, M.; Brancale, A. West Nile Virus Helicase; Homology modeling and docking study. *Antiviral Res.* **2009**, *82* (2), 60.

Chapter 8
Experimental Work.

General Information.

All chemicals, reagents and solvents were purchased from Aldrich or purified by standard techniques.

Thin layer chromatography (TLC).

Silica gel plates (Merck Kieselgel 60F254) were used and were developed by the ascending method. After solvent evaporation, compounds were visualized by irradiation with UV light at 254 nm and 366 nm.

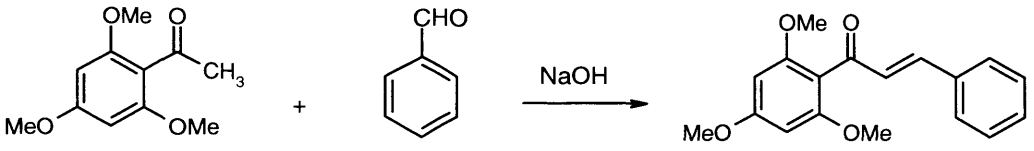
NMR Spectroscopy.

^{1}H, ^{13}C NMR spectra were recorded on a Bruker AVANCE 500 spectrometer (500 MHz) and auto calibrated to the deuterated solvent reference peak. Chemical shifts are given in δ relative to the tetramethylsilane (TMS). The spectra were recorded in $CDCl_3$ or DMSO at room temperature. TMS served as an internal standard (δ = 0 ppm) for ^{1}HNMR and $CDCl_3$ was used as an internal standard (δ = 77.0 ppm) for ^{13}CNMR.

Computational Studies.

All molecular modeling studies were performed using Molecular Operating Environment (MOE) version 2007.09, 2008.10 and FlexX 3.3.1 for molecular docking. Molecular docking was performed setting the first scoring function to the default **London dG** retaining the number of output poses to 30. All docked poses and scores were written to "dock.mdb" output database. Before docking the protein was prepared by addition of all hydrogens, and energy minimization.

8.1- Synthesis of (E)-3-phenyl-1-(2, 4, 6-trimethoxyphenyl)-2-propen-1-one (42).



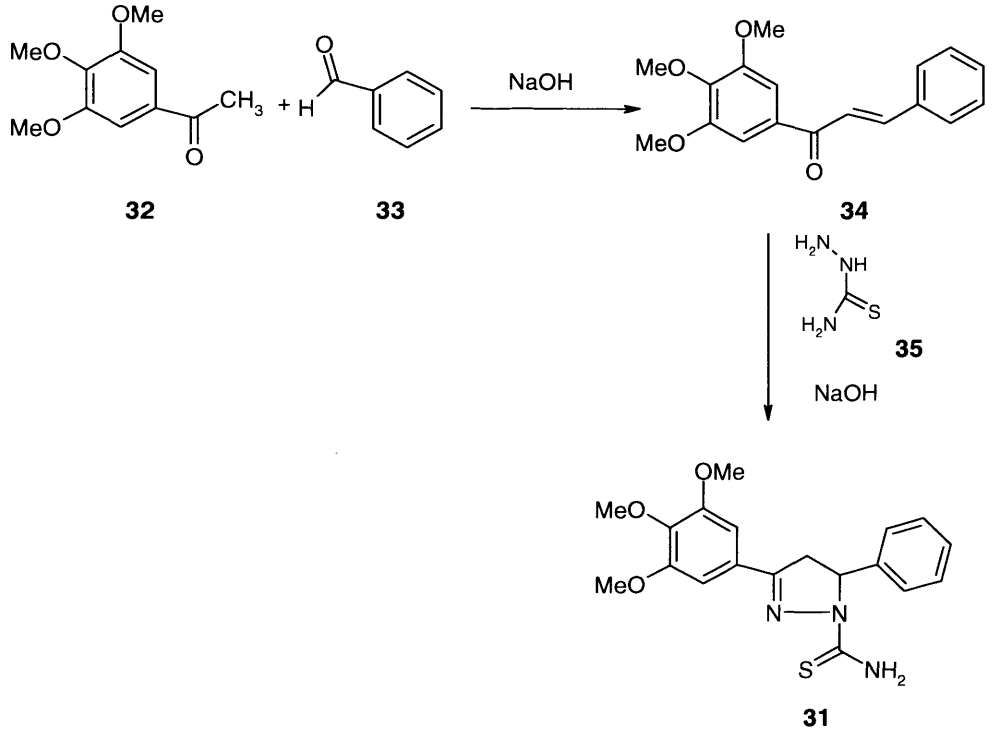
41 42

2, 4, 6-trimethoxyacetophenone (0.9 gm, 0.0043 mol) was added To NaOH (2.2 g) in 20 ml water and 12 ml absolute ethanol. The mixture was cooled in crushed ice and stirred for 15 minutes. Then benzaldehyde (0.455gm, 0.0043 mol) was added drop wise. The mixture was stirred vigorously till the reaction mixture became thick. The stirrer was removed and the contents were left in the refrigerator overnight, filtered, washed with cold water, and crystallized from ethanol. Over yield: 1.0g (80 %). Yellow crystals. M.P = 110-115 °C

^{1}H NMR (500 .0 MHz, CDCl3) δ 8.0 (s, 1H, CH), 7.9 (s, 1H, CH), 7.8 (dd, 1H, CH), 7.5 (m, 5H, C_6H_5), 6.6 (dd, 1H, 2CH), 3.8 (s, 6H, OCH_3), 3.5 (s, 3H, OCH_3].

^{13}C NMR (CDCl3) δ 189.23 (carbonyl), 162.2 (Quaternary carbon), 160.0 (Quaternary carbon), 144.7 (CH), 133.5 (Quaternary C), 130.5 (CH aromatic), 128.9 (CH aromatic), 121.8 (CH), 103.6 (1C, Quaternary carbon), 91.1 (CH), 55.7 (OCH_3), 55.4 (OCH_3).

8.2- Synthesis of 3-phenyl-5-(3, 4, 5-trimethoxyphenyl)-4, 5-dihydro-1H-1-pyrazole carbo thioamide (31).

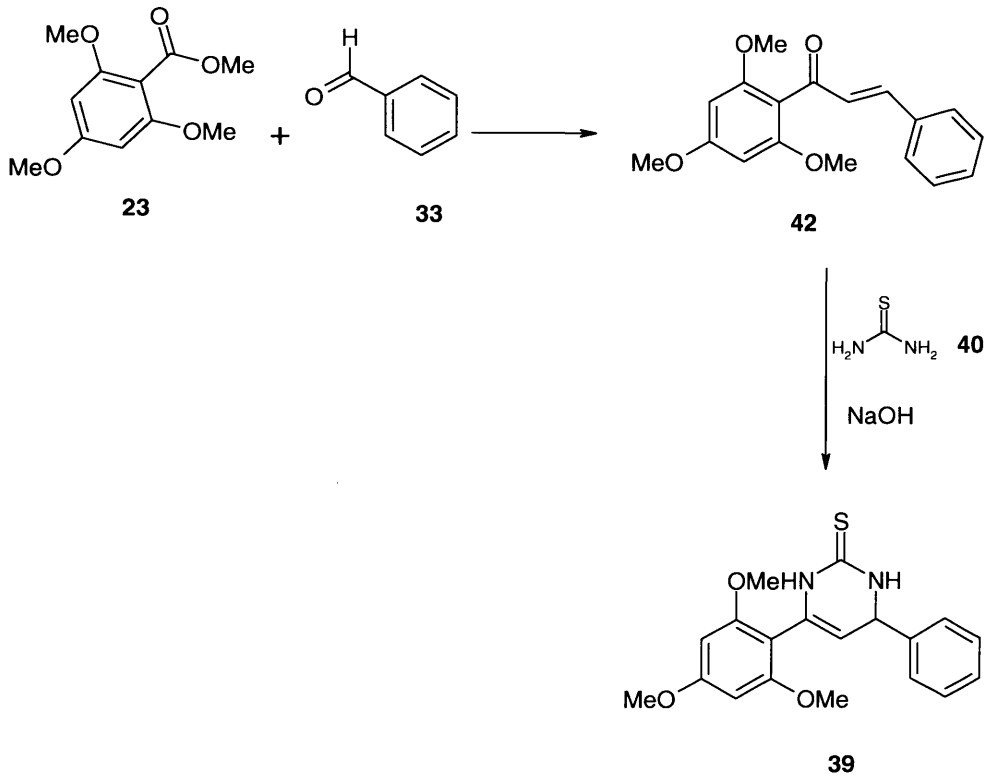


General procedure for chalcones was used to synthesize compound **34.** Thiosemicarbazide (1.09 gm, 0.012 mol) was added to a suspension of **34** (2.98 gm, 0.01 mol), NaOH (1gm, 0.025 mol) in ethanol, and the mixture was refluxed for 8 hours. The product was poured on crushed ice and the solid was separated and crystallized from ethanol. Yield: 1.8g (49 %). M.P = 250-255 ^{0}C.

^{1}H NMR (500 .0 MHz, CDCl3) δ 3.2 (s, 1H), δ 3.7 (s, 1H), δ 4.0 (s, 9H), δ 6.2 (s, 1H), δ 7.0 (d, 2H), δ 7.2 (s, 2H), δ 7.4 (m, 5H).

^{13}C NMR (CDCl3) δ176.6 (C=S), 155.8 (Quaternary carbon), 153.5 (CH), 141.7 (Quaternary carbon), 140.9 (Quaternary carbon), 128.9 (aromatic CH), 127.7 (aromatic CH), 125,4 (aromatic CH), 125,9 (Quaternary carbon), 104.4 (CH aromatic), 63.6 (Quaternary carbon), 61.0 (OCH_3), 56.4 (OCH_3), 43.2 (CH).

8.3- Synthesis of 4-phenyl-6-(2, 4, 6-trimethoxyphenyl)-1, 2, 3, 4-tetrahydro-2-pyrimidine –thione (39).

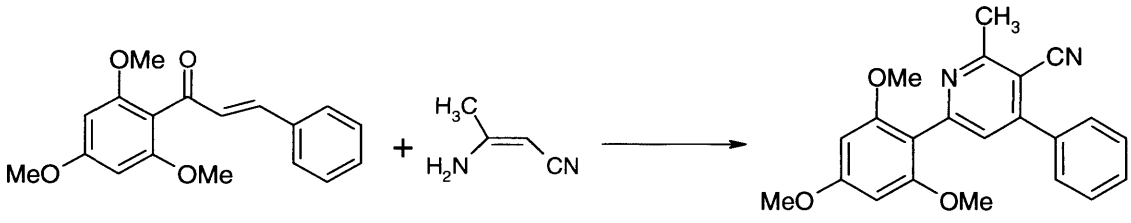


23 33 42 39

Thiourea (0.19 gm, 2.5 mmole) was added to (E)-3-phenyl-1-(2,4,6-trimethoxyphenyl)-2-propen-1-one (0.298gm, 1mmole) of in absolute ethanol. Then K_2CO_3 (0.345 gm, 2.5mmole) was added to the reaction mixture. The mixture was refluxed overnight then poured on ice, Filtered and crystallized from ethanol. Yield: 0.16g (45%). M.P = 240-245 ^{0}C

^{1}H NMR (500 .0 MHz, CDCl3) δ 7.5 (m, 5H), 6.5 (s, 1H), 6.1 (s, 2H), 5.3 (d, 1H), 5.0 (d, 1H), 4.0 (s, 6H), .5 (s, 3H).

8.4- Synthesis of 2-Methyl-6-phenyl-4-(3,4,5-trimethoxy phenyl)-3-pyridyl cyanide (46).



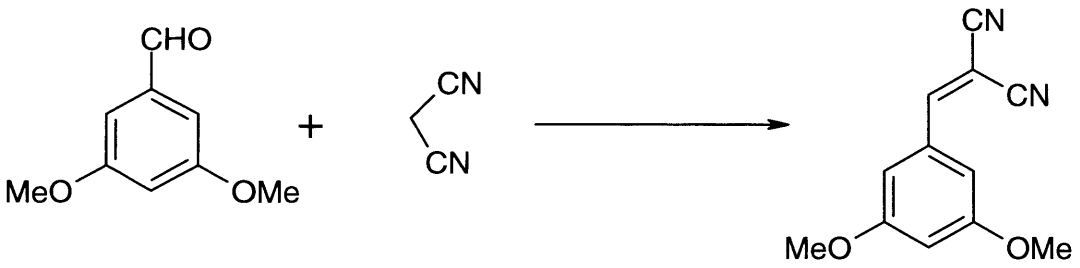
42 45 46

NaOH (0.4 gm, 10 mmole) was added to (E)-3-phenyl-1-(2,4,6-trimethoxyphenyl)-2-propen-1-one (2.98 gm, 10 mmole) in 40 ml absolute ethanol. A solution of 3-aminocrotononitrile (0.025 M) (20 ml) in absolute ethanol was added and shaked well. The mixture was refluxed at 70 °C for 6 hours, then cooled. A solution of 0.05 M HCl (10 ml) was added and shaked then the solution mixture was concentrated under reduced pressure. The resulted white powder was crystallized from absolute ethanol. Yield: 2.1g (60%). M.P = 170-175 °C

^{1}H NMR (500 .0 MHz DMSO-d6) δ 7.4 (m, 5H), 7.2 (s, 1H), 6.3 (s, 2H), 3.9 (s, 3H), 3.8 (s, 3H), 3.4 (s, 3H), 1.9 (s, 3H).

^{13}C NMR (DMSO-d6) δ 161.30 (quaternary), 148.0 (CH), 128.7 (CH), 128.3 (CH aromatic), 127.7 (quaternary carbon), 126.4 (CN), 121.7 (CH aromatic), 106.7 (CH aromatic), 103.3 (Quaternary carbon), 90.8 (quaternary carbon), 77.5 (CH aromatic), 55.8 (OCH3), 55.4 (OCH3), 39.5 (OCH3), 17.8 (CH3).

8.5- Synthesis of 2-[(3,5-dimethoxyphenyl) methylene] malononitrile (50).

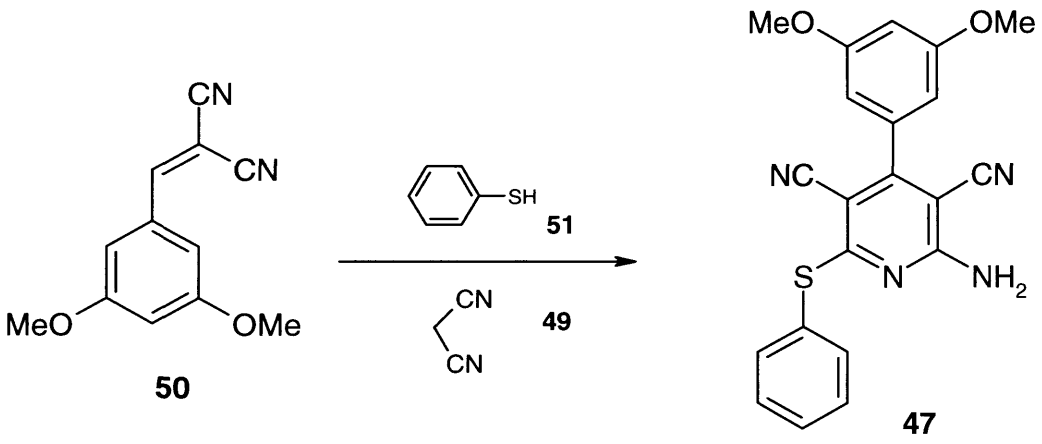


48 **49** **50**

Malononitrile (0.021gm, 1.5 mmole) was dissolved in 14 ml ethanol. 3, 5-Dimethoxy benzaldehyde (0.166gm, 1mmol), 3 drops of piperidine were mixed and the mixture was refluxed for 1 hour. Then cooled to room temperature. The formed brown precipitate was filtered off and dried. Yield: 0.15g (73 %). M.P =115-120 °C.

^{1}H NMR (500 .0 MHz DMSO-d6) δ 7.8 (s, 1H), 7.1 (s, 2H), 6.8 (s, 1H), 3.8 (s, 6H).

8.6-Synthesis of 2-Amino-4-(3,5-dimethoxypheny)-6-(phenylsulfanyl)-3, 5-pyridine dicarbonitrile (47).

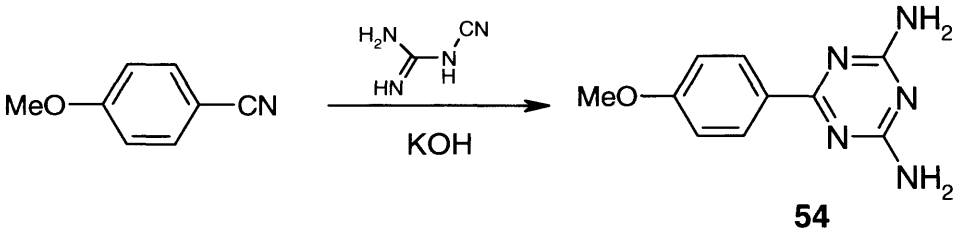


2-[(3,5-dimethoxyphenyl) methylene] malononitrile (0.214gm, 1mmole) was dissolved in absolute ethanol (5ml). Malononitrile (0.014gm, 1mmole) and thiophenol (0.11gm, 1 mmole), and triethylamine (5ml) were added and the mixture was refluxed for 4 hours. Then cooled and filtered. The resulted greenish precipitate was crystallized from ethanol. Yield: 0.15g (40 %). M.P = 210-215 °C.

^{1}H NMR (500 .0 MHz DMSO-d6) δ 7.8 (s, 2H), 7.7 (m, 5H), 7.5 (s, 1H), 6.7 (s, 2H), 3.7 (s, 6H).

^{13}C NMR (DMSO-d6) δ 165.9 (quaternary carbon), 160.4 (quaternary carbon), 159.5 (quaternary carbon), 135.7 (Quaternary carbon), 134.8 (CH aromatic), 129.6 (CH aromatic), 129.4 (1C, CH aromatic), 127.1 (quaternary carbon), 115.1 (CN), 114.8 (CN), 106.5 (CH aromatic), 101.7 (CH aromatic), 93.3 (quaternary carbon), 87.1 (quaternary carbon), 55.5 ($2OCH_3$).

8.7- Synthesis of 2-(4-methoxyphenyl)-1, 3, 5-triazine-2, 4-diamine (54).



54

A mixture of dicyandiamide (0.09g, 1.1 mmole) and KOH (0.12g, 2.2 mmole) was added to 4-methoxybenzonitrile (0.133g, 1mmole) in THF (10ml). The mixture was refluxed for 6 hours. Upon cooling the formed precipitate was collected, washed and dried. Yield: 0.14g (65 %.). M.P = 270-275 ^{0}C.

^{1}H NMR (500 .0 MHz D2O) δ 7.7 (dd, 2H), 7.0 (dd, 2H), 3.6 (s, 3H).