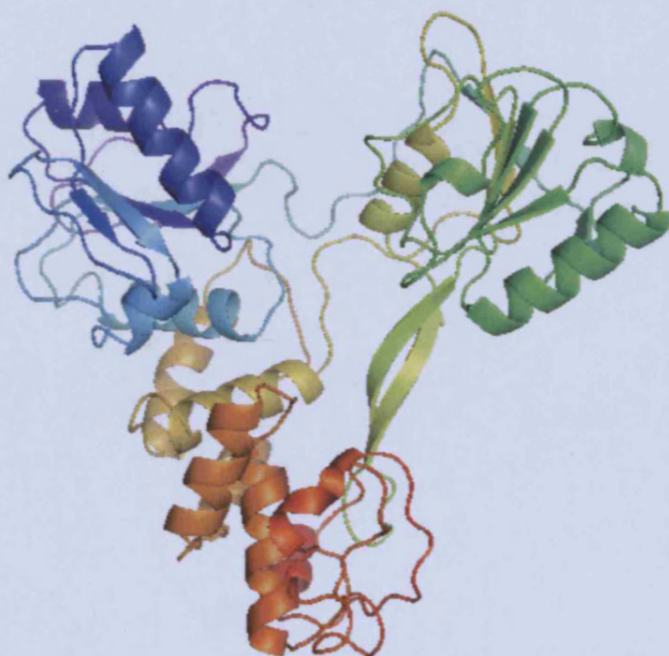




Computer Aided Drug Design



A thesis submitted to Cardiff University for the degree of doctor
in philosophy in the school of pharmacy

Sahar Kandil

Welsh School of Pharmacy, Cardiff University

2009

UMI Number: U584380

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U584380

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



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... Sahar Kandil... (candidate) Date... 22/12/2009

Statement 1

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

Signed... Sahar Kandil... (candidate) Date... 22/12/2009

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... Sahar Kandil... (candidate) Date... 22/12/2009

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... Sahar Kandil... (candidate) Date... 22/12/2009

Abstract

Hepatitis C virus (HCV) chronic infection represents one of the major and still unresolved health problems. HCV infecting 3% of the world population, leading to chronic hepatitis, liver cirrhosis and hepatocellular carcinoma in addition to the extrahepatic manifestations.

No efficient therapy exists; the standard dual treatment with peg IFN- α and ribavirin is effective only in 55% of the selected cases with substantial side effects in addition to the high cost. To date, there is no vaccine against HCV due to the high variability of the RNA genome. NS3 helicase is one of the non-structural proteins whose activity is indispensable for viral RNA replication and its inhibition is estimated to arrest viral proliferation and indirectly stimulate a cellular antiviral response against ds RNA. In our project we proposed to use structure based knowledge of the x-ray crystal structure of helicase enzyme to design and synthesise different scaffolds of novel potential HCV NS3 helicase inhibitors. Using different computer software packages, we manage to design a number of small focused libraries of compounds, which were used for docking simulations. The results obtained *in silico* guided the selection of two series of promising compounds for synthesis. In the first series; several quinazoline derivatives were prepared and evaluated for antiviral activity in subgenomic replicon assay showing EC₅₀ in the low μ M range with relatively high selectivity index. In the second series of pyrrole or phenyl based compounds, irreversible inhibition of helicase is assumed through addition to the electrophilic warheads of the α β unsaturated ketones, thiols or 1,2,4 thiadiazoles based inhibitors. Among the synthesised compounds a number showed a sub μ M activity in the helicase enzyme assay.

These promising findings are considered to be a starting point for further optimisation of structure, activity and toxicity relationships.

Acknowledgement

A debt of gratitude must be acknowledged towards those who have in various ways, been of particular and valuable assistance. First and foremost, my sincere appreciation goes to my supervisor, Dr. Andrea Brancale for the joy and pride of working under his kind understanding and supporting guidance throughout my PhD.

I am very thankful to Prof. Chris McGuigan for letting me attend his group meetings, which were of great help, especially in the initial stages of this thesis. I wish to express my sincere thanks to Dr. Andrew Westwell for his advices and constructive discussions. My deepest gratitude goes to Dr. Claire Simons for her generous help and support. I would like to acknowledge the cooperative attitude of Prof. Johan Neyts, Dr. Dimitrios Vlachakis and Dr. David Frick. I cannot fully express my gratitude to the exceptional team of the Medicinal Chemistry group at both the chemistry and the modelling labs. Thank you especially to Helene Bruyere and Nicola Zonta for their generous incomparable extraordinary and trusted friendship. I am also grateful to many friends and colleagues Sonia, Stephania, Anna, Elisenda, Bhattacharya, Onel, Priscilla, Lijun, Michael, Jaqui, Munirah and David for sharing all the ups and downs over my PhD.

I would be remiss if I did not mention that without the warm climate of kindness and friendship, which pervades the Welsh School of Pharmacy, this thesis would perhaps have never seen the light of the day. And finally, to my family, for their love and support at all times.

Finally, this thesis is dedicated in loving memory to my first teacher and guide, my source of trust and pride and the finest kind of utopian humans, my father, whom I owe everything. You were my encouraging protective support all the times.

List of Contents

	Page	
Chapter 1	Introduction	1
<hr/>		
1.1 HCV: The Best of a Bad Virus.....		2
1.2 Clinical spectrum of the "Silent Killer".....		3
1.3 Limitations for HCV research.....		4
1.4 Current treatment; Interferon (IFN- α)/ Ribavirin.....		6
1.4.1 Discovery.....		6
1.4.2 Structure and Classification.....		6
1.4.3 Mechanism.....		7
1.4.4 Drawbacks.....		7
1.4.5 Ribavirin (RBV).....		7
1.4.6 Pegylation.....		8
1.5 Flaviviridae family and HCV life cycle.....		10
1.6 HCV genome organisation.....		12
1.7 HCV vaccine development.....		14
1.8 Viral targets for future therapies.....		15
1.8.1 Protease inhibitors.....		15
1.8.2 Polymerase inhibitors.....		16
1.8.3 Immune modulators.....		18
1.8.4 Further novel investigational agents.....		18
Chapter 2	Helicase Enzyme; Overview	20
<hr/>		
2.1 Helicases (the Unwinding Machine).....		21
2.2 Classification, Structure and function.....		22
2.3 Viral Helicases as antiviral drug targets.....		24
2.4 HCV NS3 Helicase.....		26
2.4.1 Structure of HCV NS3 helicase.....		26
2.4.2 Mechanism of action.....		27
2.4.3 Host vs Viral Helicases (Selectivity Window).....		28
2.4.4 HCV NS3 Helicase as Drug Target.....		28
2.4.4.1 Inhibition of NTPase Activity Direct/Allosteric.....		29

2.4.4.2 Inhibition of Unwinding through Intercalation.....	33
2.4.4.3 Competitive Inhibition of RNA Binding.....	33
2.5 Reasearch objectives.....	37
Chapter 3	Results and Discussion
	38
<hr/>	
3.1 Molecular Modelling.....	39
3.1.1 Molecular mechanics (MM).....	41
3.1.2 Force fields.....	42
3.1.3 Energy Minimisation	45
3.1.4 Docking Simulation	47
3.1.4.1 Docking Algorithms.....	48
3.1.4.2 Shape complementarity	50
3.1.4.3 Protein Flexibility.....	51
3.1.5 Scoring Functions	51
3.1.6 Structural based Virtual Screening.....	54
3.1.7 <i>De Novo</i> Drug Design	55
3.1.8 Molecular Dynamics (MD).....	58
3.1.9 Ligand-Based Drug Design.....	60
3.2 HCV Helicase; Mechanistic insights.....	61
3.2.1 "Propulsion-by Repulsion" model.....	64
3.3 Quinazoline based structures.....	66
3.3.1 Introductory remarks.....	66
3.3.2 Inhibitor Structure Design.....	68
3.3.3 Bioisosterism of carboxylate group	76
3.3.4 Synthesis of Quinazoline based series.....	78
3.4 α , β Unsaturated ketones.....	90
3.4.1 Introductory remarks.....	90
3.4.2 Previous work overview.....	92
3.4.3 Inhibitor Structure Design.....	94
3.4.4 Pharmacophore based virtual screening.....	98
3.4.5 Synthesis of Pyrrole derivatives.....	101
3.4.6 Preparation of Phenyl analogues.....	106
3.4.7 Preparation of N-MethylPyrrole derivatives.....	109
3.5 Thiol derivatives.....	114
3.5.1 Inhibitor Structure Design	114

3.5.2 Synthesis of Thiol derivatives.....	117
3.6 (1,2,4)-Thiadiazole derivatives.....	123
3.6.1 Cysteine; A special amino acid.....	123
3.6.2 (1,2,4)-Thiadiazoles (THDs).....	124
3.6.3 Mechanism of Cysteine enzymes Inhibition.....	125
3.6.4 Proof of Principal.....	126
3.6.5 Inhibitor Structure Design.....	127
3.6.6 Selected compounds for chemical synthesis.....	129
3.6.7 Conclusion.....	133
Chapter 4	Experimental Chemistry
<hr/>	
4.1 Quinazoline based structures.....	136
4.2 α , β Unsaturated ketones.....	160
4.3 Thiol derivatives.....	181
Chapter 5	Assays: overview and findings
<hr/>	
5.1 Replicon.....	191
5.2 Cell culture propagation of HCV.....	191
5.3 Establishment of the first HCV replicon.....	192
5.4 Breakthrough for HCV research.....	193
5.5 Application drug development.....	194
5.6 Subgenomic replicon assay.....	194
5.7 Helicase enzyme assay.....	194
5.8 Anti-HCV assay of Quinazoline structures.....	195
5.9 Anti-HCV assay of α , β unsaturated ketones.....	197
5.10 Anti-HCV assay of thiol derivatives.....	199
Chapter 6	Bibliography
<hr/>	
References	201

List of abbreviations

AcOH	Acetic acid
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
CADD	Computer aided drug design
Cys	Cysteine
DCM	Dichloromethane
DMAP	Dimethylaminopyridine
DMF	Dimethylformamide
DMF DMA	Dimethylformamide dimethyl acetal
DMSO	Dimethylsulfoxide
DNA	Deoxy ribonucleic acid
DTT	Dithiothreitol
EDCI	1-ethyl-3(dimethylaminopropyl) carbodiimide hydrochloride
EMCV	Encephalomyocarditis virus
EtOAc	Ethylacetate
EtOH	Ethanol
GA	Genetic algorithms
Gln	Glutamine
Glu	Glutamic acid
HCV	Hepatitis C virus
His	Histidine
HIV	Human immunodeficiency virus
IFN	Interferon
IMPD	Inosine monophosphate dehydrogenase
IRES	Internal ribosomal entry site
LBDD	Ligand based drug design
Leu	Leucine
MD	Molecular dynamics

MeOH	Methanol
MF	Molecular formula
MM	Molecular mechanics
MMFF	Merck Molecular Force Field
mp	Melting point
Mw	Molecular weight
NMR	Nuclear magnetic resonance
NTR	Non translated region
PDB	Protein data bank
PEG	Polyethylene glycol
Phe	Phenylalanine
QM	Quantum mechanics
QSAR	Quantitative structure activity relationship
RBF	Round bottom flask
RBV	Ribavirin
RDRP	RNA dependent RNA polymerase
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SBDD	Structure based drug design
Ser	Serine
SVR	Sustained virological response
TEA	Triethylamine
THD	1,2,4 thiadiazole
THF	tetrahydrofuran
Thr	Threonine
TLC	Thin layer chromatography
TLR	Toll like receptors
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

Chapter 1

Introduction

Chapter 1

Introduction

1.1 HCV: The Best of a Bad Virus

In the mid-1970s, it was noticed that the world's supply of blood was contaminated with an unidentified agent causing post-transfusion non-A, non-B hepatitis. Yet it was not until 1989 that the first sequences of hepatitis C virus (HCV) were identified^{1,2}.

Hepatitis C virus (HCV) infection is a global health problem, being the second most common chronic viral infection in the world is the cause of emerging worldwide epidemic of chronic liver disease with serious medical, social and economic burden. Hundreds of thousands of people die each year from liver failure and cancer caused by this infection³. Currently, it is estimated that HCV has a global prevalence of about 3% (about 200 million people), nearly five times more than human immunodeficiency virus (HIV) infected individuals, and 3-4 million new infections are added each year⁴ (Figure 1.1).

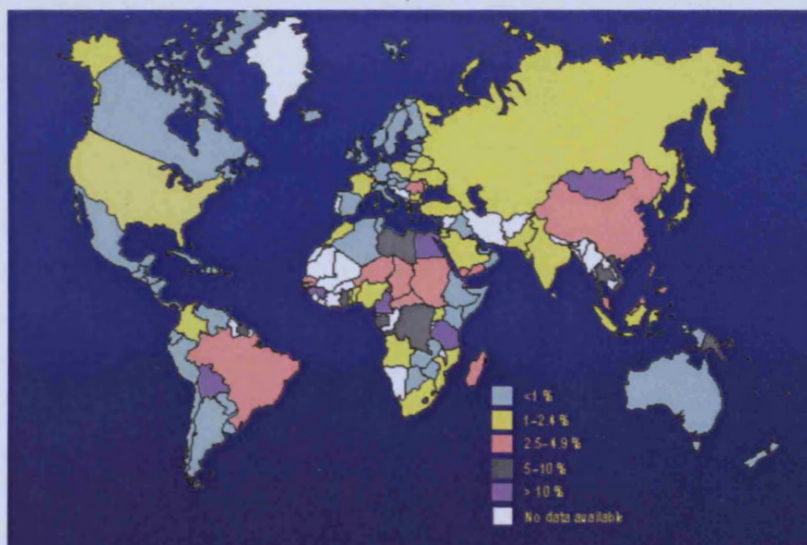


Figure 1.1 Global prevalence of hepatitis C Virus³

1.2 Clinical spectrum of the “Silent Killer”

HCV is a blood borne virus, transmitted by infected body fluids, in most of the infected individuals (85%). This remarkable RNA virus evades the immune system. HCV establish a chronic infection that can lead to both a hepatotropic and a lymphotropic manifestations causing on one hand, chronic hepatitis, cirrhosis and hepatocellular carcinoma and on the other hand several extrahepatic diseases like endocrine dysfunction ⁵.

The most frequently occurring of these endocrine disorders are thyroid autoimmunity, hypothyroidism and thyroid cancer. Also, HCV infection could lead to the development of type 2-diabetes mellitus, possibly as a result of HCV-induced metabolic disturbances. In fact, the association between hepatic (hepatocellular carcinoma) and extrahepatic (lymphoma, thyroid cancer) malignancies has justified the inclusion of HCV among human cancer viruses. The pathogenesis of HCV-related sequelae (hepatic or extrahepatic) is not fully understood and hence representing a challenge of prime importance ^{5,6}.

HCV is sometimes called a “silent killer” because many patients do not realize the infection until chronic liver damage is manifested 1-3 decades after the initial contraction of the virus ⁵. However, high viral loads are observed during the first few weeks of HCV infection and inflammatory processes leading to liver injury are usually occurring after 2–3 months ⁷.

Liver transplant recipients generally have favourable short-term outcomes despite efficient allograft reinfection and high levels of viraemia owing to immunosuppression. These observations have led to the idea that HCV is relatively noncytopathic and that liver disease is mainly immune-mediated. Nevertheless, studies on infected livers indicated that this organ is a veritable battleground of ongoing viral replication and host antiviral defences ⁸⁻¹¹.

Although the liver (the body metabolic engine) is the major site of HCV replication, evidence exists for systemic reservoirs including lymphoid cells¹²⁻¹⁴, epithelial cells in the gut¹⁵ and the central nervous system¹⁶.

This extrahepatic infection contributes to the immune-mediated pathogenesis of chronic liver disease and the development of autoimmune diseases, including mixed cryoglobulinaemia, which is the presence of abnormal proteins called cryoglobulins in blood that can cause a syndrome marked by varying combinations of fatigue, muscle and joint aches, besides arthritis, skin rash, neuropathy and glomerulonephritis, which is kidney disease affecting the capillaries of the glomeruli characterised by oedema, raised blood pressure and excess protein in the urine¹²⁻¹⁴.

1.3 Limitations for HCV research

Until recently the development HCV therapy has been severely hampered by several factors. One of these is the lack of small animal model for the assessment of the preclinical efficacy of novel drugs. While the only non-human animal that can be reliably infected with HCV is chimpanzee (*Pan troglodytes*), whose genome has a 98.77% sequence homology with the human genome, the use of this endangered species, as a laboratory animal is for ethical and economical reasons not evident¹⁷.

Recently, an important achievement has been made by the propagation of HCV in mouse models using either immunocompromised mice engrafted with human hepatocytes isolated from fresh livers¹⁷ or otherwise using human liver fragments infected with HCV *ex vivo* before being transplanted into immunocompromised mice¹⁸.

Nevertheless, the greatest limitation for HCV research has been the unsuccessful cell culture system that supports the efficient and reliable propagation of the virus¹⁹. However, the advent of the subgenomic replicon system, first reported by Lohman *et al* in 1999, established persistent HCV replication in a human hepatoma cell line (Huh-7).

The inefficiency of the RNA replication initiation in this system limited its utility, but this breakthrough provided a basis for further optimisation^{20,21}. This will be discussed in more details in chapter 5.

Moreover, the extremely exceptional genetic diversity of HCV as it is currently classified into six major genotypes, which diverge by about 30% at the nucleotide sequence level with more than 30 subtypes throughout the world. These six distinct genotypes show marked differences in geographic distribution, disease progression and response to therapy²².

In chronically infected patients viral loads typically range from 10^3 – 10^7 genomes/mL of serum. Mathematical modelling of viral dynamics indicates that HCV virions turn over rapidly (with a half-life about 3 hours), and up to about 10^{12} viruses are produced per day. This is about 100-fold greater than the rate reported for HIV^{23,24}. This unusual high replication rate together with lack of proof reading function of RNA dependent RNA polymerase (RDRP) result in explosive expansion of the virus and evolution of numerous viral *Quasispecies*. *Quasispecies* is a family of closely related, but slightly different, viral genomes that differ mainly in the hypervariable regions (HVR1 and -2) of the E2 gene. Thus, the virus produces a constant stream of escape variants with no corresponding receptors in the immunological repertoire²⁵⁻²⁸. Despite all these obstacles, there is a considerable progress in HCV study in the last years (Figure 1.2).

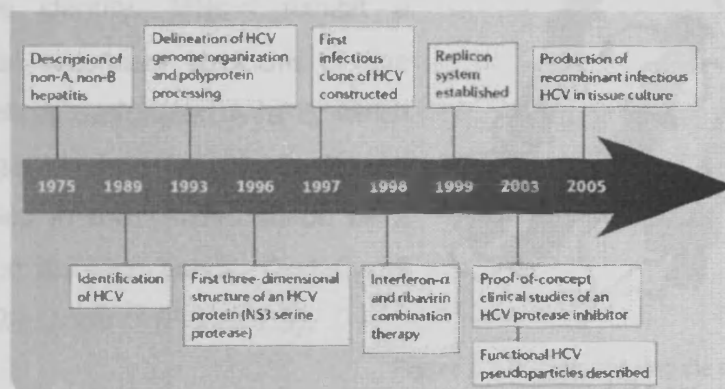


Figure 1.2 Milestones in hepatitis C virus (HCV) research²⁹

1.4 Current treatment; Interferon (IFN- α) / Ribavirin

In the absence of an effective antiviral or vaccination strategy against HCV, the single FDA approved therapy is Interferon- α (IFN- α).

1.4.1 Discovery

Interferon was discovered in 1957 during the analysis of the effects of viral infection on cells in a tissue culture. It was noticed that cells already infected with a virus appeared to be resistant to infection by other viruses for a certain period of time. The first infection was said to "interfere" with the second and the protein isolated from these cell cultures was therefore given the name interferon (IFN). It is now known that these substances belong to a class of proteins produced by leucocytes as soon as the body is exposed to attack by viruses, other microorganisms, or tumour cells. Being a crucial mediator of the antiviral immune response IFN- α was a natural choice for the treatment of HCV³⁰.

1.4.2 Structure and Classification

IFNs composed of 165 / 166 amino acid residues can be classified on the basis of their structure into three types. Among type I (IFN), IFNs- α occupy a central position which include at least 15 subtypes that share between 70 and 80 % amino acid sequence identity. IFNs- α exhibit a conserved secondary structure of five alpha helices designated A to E, which pack together as a helical bundle and maintained in their folded shape by 2 conserved disulfide bonds: Cys1-Cys99 and Cys29-Cys39³⁰ (Figure 1.3).

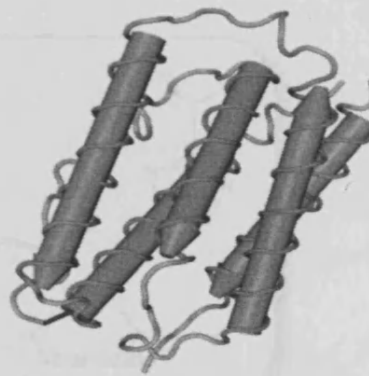


Figure 1.3 Diagram showing the secondary structure of Interferon alpha

1.4.3 Mechanism

Basically, IFN- α is a cytokine that facilitates recognition of viral antigens by the immune system, activates natural killer cells and macrophages, and directly inhibits viral replication³⁰⁻³³.

1.4.4 Drawbacks

Unfortunately, therapy with IFN- α alone had only 15% sustained virological response (SVR), which is the continued lack of detectable serum HCV RNA six months after the completion of treatment³⁴. Furthermore, IFN- α monotherapy is limited by adverse side effects, these range from nausea through to flu-like symptoms; fever, headache, chills, and muscle pain, blood changes; leukopaenia (reduced leucocyte count in blood), thrombocytopenia (reduced platelets number in blood). Nevertheless, autoimmune diseases, depression and hair loss are common³⁵.

1.4.5 Ribavirin (RBV)

A major advance came with the addition of the broad-spectrum antiviral agent ribavirin, which is a synthetic guanosine nucleoside analogue, to IFN- α treatment (Figure 1.4). This combination therapy resulted in SVR of 40% but with other side effects including haemolytic anaemia due to accumulation of RBV in erythrocytes³⁶⁻³⁸.

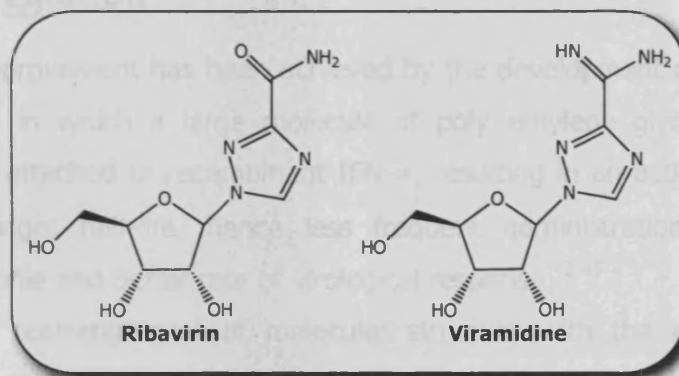


Figure 1.4 Chemical structures of ribavirin and viramidine

In search for ribavirin analogues with better safety profile, Viramidine was identified (Figure 1.4). Viramidine is a prodrug, which is converted to ribavirin by adenosine deaminase, an enzyme abundant in the liver³⁹.

The exact mode of action of ribavirin remains unknown, but multiple mechanisms have been proposed:

- *Direct inhibition of HCV replication;* As a guanosine analogue, RBV is phosphorylated and misincorporated into HCV RNA by RNA polymerase causing chain termination, and inhibition of replication.
- *Inosine-monophosphate-dehydrogenase inhibition;* Intracellularly, RBV monophosphate (RMP) is a competitive inhibitor of inosine monophosphate dehydrogenase (IMPDH), which in turn leads to depletion of intracellular levels of GTP required for HCV RNA synthesis.
- *Mutagenesis and error catastrophe;* A recent study suggested that while the diversity of HCV Quasispecies (virions with minor genomic differences) is caused by the high frequency of mutations that occur during viral replication owing to the lack of proofreading activity of the HCV RNA dependent RNA polymerase (RDRP), RBV acts as a viral mutagen, causing a higher frequency of mutations, pushing viruses toward the threshold of 'error catastrophe' and dramatically reducing viral infectivity^{40,41}.

1.4.6 Pegylation

Further improvement has been achieved by the development of pegylated interferon, in which a large molecule of poly ethylene glycol (PEG) is covalently attached to recombinant IFN- α , resulting in an active molecule with a longer half-life, hence less frequent administration, improved toxicity profile and better rate of virological response⁴²⁻⁴⁵.

Generally, rearrangement of molecular structure with the aid of PEGs allows the molecular properties of a drug to be altered. Protein pegylation is achieved via stable covalent bonds between an amino or sulfhydryl

group on the protein and a chemically reactive group (carbonate, ester or aldehyde) on the PEG. The resulting structures can be linear or branched (Figure 1.5).

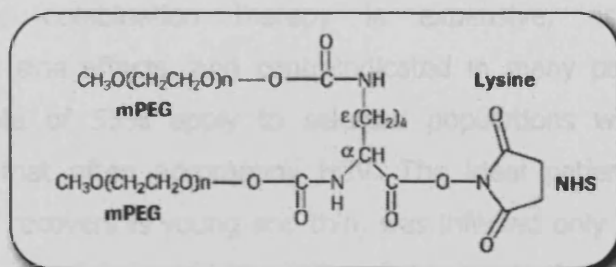
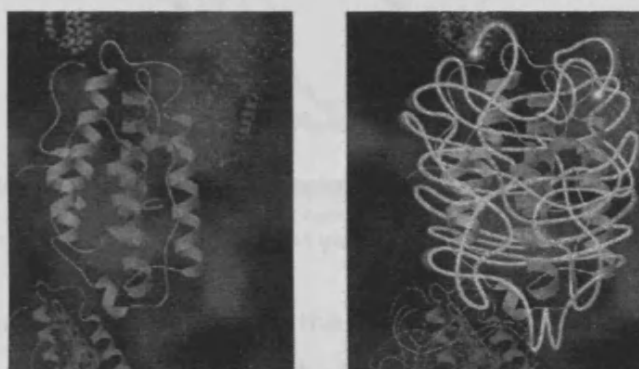


Figure 1.5 A branched PEG molecule created by coupling a monofunctional PEG (mPEG) to lysine, conjugation of which to IFN- α results in better pharmacokinetic profile. NHS, *N*-hydroxysuccinimide

This technique improves stability, biological half-life, water solubility, and immunologic characteristics. Pegylation increases the size of a small molecule to such an extent that it is less readily excreted through the kidneys and therefore persists in the body for longer.

In addition, as they are more or less surrounded by the attached PEGs (Figure 1.6, Roche), pegylated proteins are less rapidly broken down by the body's enzymes than are unmodified proteins these changes ultimately lead to increased half-life and hence reduce the frequency of injections a patient requires. They also reduce the rapidity and intensity of the body's immune reaction against the interferon molecules.



Standard interferon

Pegylated interferon

Figure 1.6 3D structure of standard IFN molecule and pegylated one

The combination of pegIFN- α with ribavirin yields sustained response rates of 55%. These results are heartening, but they also mean that about 40% of patients do not have lasting improvement with treatment⁴⁶⁻⁴⁸.

Furthermore, combination therapy is expensive, associated with troublesome side effects, and contraindicated in many patients *i.e.* this response rate of 55% apply to selected populations without the co-morbidities that often accompany HCV. The ideal patient in terms of likelihood of recovery is young and thin, was infected only recently, has a low concentration of virus, shows little cell damage in the liver biopsy, and is infected with virus of a favorable genotype. Unfortunately, these criteria are met by only 10 - 20 % of patients.⁴⁹⁻⁵¹

1.5 Flaviviridae family and HCV life cycle

HCV is an enveloped positive-stranded RNA virus (Figure 1.7) possessing a genome of approximately 9000-9300 nucleotides and is classified as Hepacivirus genus of Flaviviridae family⁵².

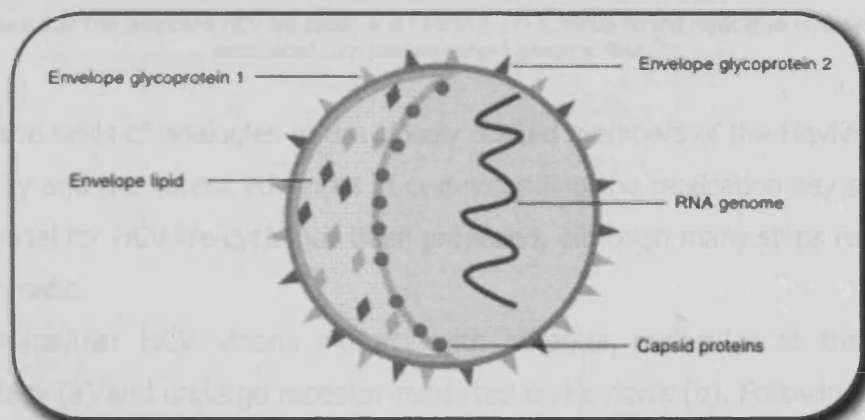


Figure 1.7 Model structure of HCV; left-hand side; viral surface of envelope lipids and glycoproteins; right-hand side; RNA genome encased by capsid proteins⁵³

The virus family was named after the jaundice occurring in the course of Yellow fever virus (YFV) infection, the first identified virus of the family⁵⁴.

Flaviviridae family of viruses is associated with both human and animal

diseases and comprises three distinct genera: pestiviruses, which cause diseases in cattle and pigs; flaviviruses, which are the most important cause of diseases such as dengue fever and yellow fever; and hepaciviruses, whose sole member is HCV⁵⁵.

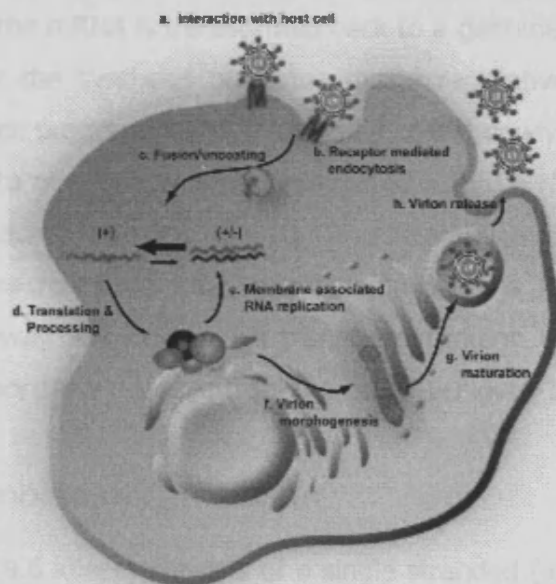


Figure 1.8 The proposed HCV life cycle; + is (+RNA); +/- is minus-strand replicative intermediate associated with positive-strand genomic RNA⁵⁶

On the basis of analogies to the closely related members of the Flaviviridae family and the recent advances in understanding the replication key steps, a model for HCV life cycle has been proposed, although many steps remain enigmatic.

Extracellular HCV virions interact with receptor molecules at the cell surface (a) and undergo receptor-mediated endocytosis (b). Following HCV glycoprotein-mediated membrane fusion, the viral single-stranded (ss), positive-sense RNA (+RNA) genome is released into the cytoplasm (c). The genomic RNA acts as mRNA for the translation of the viral single large polyprotein that is processed into the 10 mature HCV proteins in association with a virus-derived membrane structure termed the

membranous web, the translation of the polyprotein is directed by an internal ribosomal entry site (IRES) (d). Translation is interrupted by an unknown signal that activates a reaction leading to the formation of a replication complex on membranes of the endoplasmic reticulum. As a consequence, the mRNA is transformed back to a genome that now acts as a template for the synthesis of minus strand replicative intermediate (-RNA) to produce progeny (+RNA). A portion of this newly synthesized RNA is packaged into nucleocapsids and associated with the HCV glycoproteins, leading to budding into the ER (f). Alternatively, the rest of (+RNA) undergo a new round of RNA amplification. Virions follow the cellular secretory pathway and during this transit, maturation of particles occurs (g). Mature virions are released from the cell (h) (Figure 1.8).⁵⁶⁻⁵⁸

1.6 HCV genome organisation

HCV genome (9.6 kbase) consists of a single stranded (+RNA) encoding a single large open reading frame (ORF), flanked by structured 5' and 3' NCRs (noncoding regions)⁵⁹.

The translation of ORF, via the activity of an internal ribosomal entry site (IRES) in the NCR 5', generates a large multidomain polyprotein of roughly 3,000 amino acids that is organized with structural proteins in the amino-terminal of the polyprotein, followed by the non structural (NS) replication proteins. The structural proteins are used for the assembly of new-progeny virus particles, whereas most of the nonstructural (NS) proteins participate in the replication of the viral genome. The polyprotein undergoes a complex co- and post-translational series of cleavage events, catalysed by both host and viral proteases, to produce 10 individual HCV proteins (Figure 1.9).⁵⁶

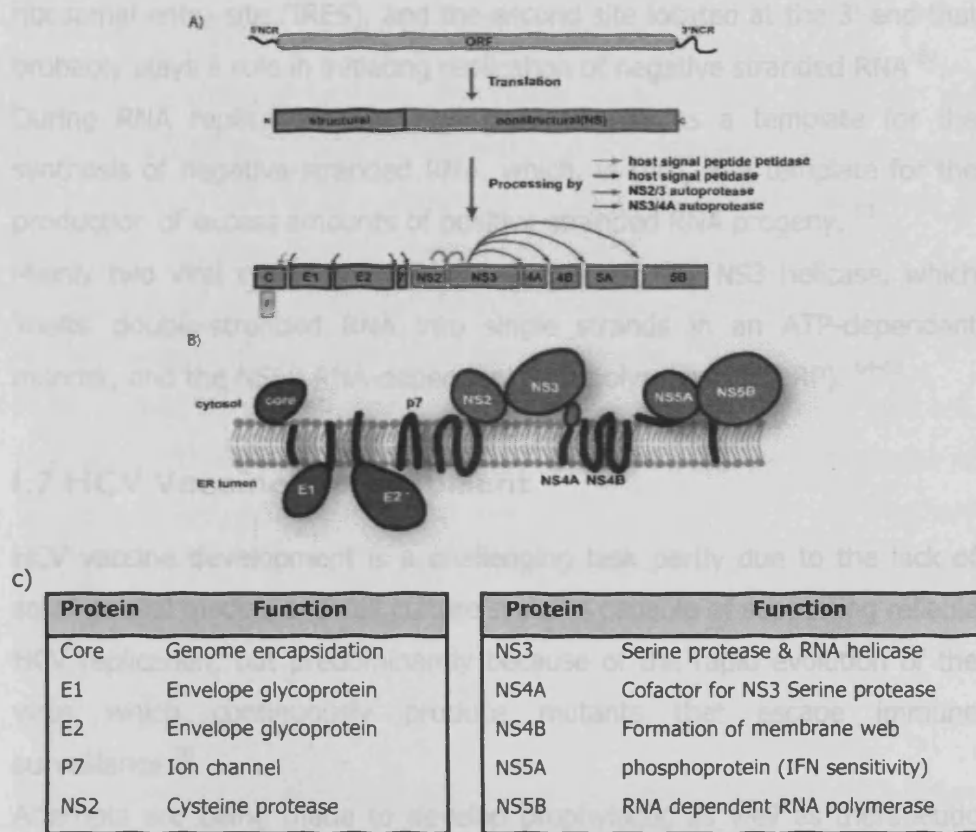


Figure 1.9 A) HCV genome and encoded viral proteins B) Topology of the HCV proteins relative to the ER membrane C) HCV encoded proteins and their respective functions ⁵⁶

The amino-terminal one third of the polyprotein encodes the virion structural proteins: a core (C) protein, and two envelope glycoproteins E1 and E2. After the structural region comes a small integral membrane protein, p7, as an ion channel. Then six non-structural (NS) proteins, NS2, NS3, NS4A, NS4B, NS5A and NS5B, which are essential enzymes or accessory factors for viral replication, translation, and polyprotein processing and thus believed to be potential drug discovery targets. ⁵⁹⁻⁶²

In addition, the genome contains two nontranslated/noncoding regions (NTRs/NCRs); the first is the 5'-domain that functions as an internal

ribosomal entry site (IRES), and the second site located at the 3'-end that probably plays a role in initiating replication of negative stranded RNA⁶².

During RNA replication, the viral genome acts as a template for the synthesis of negative-stranded RNA, which, in turn, is a template for the production of excess amounts of positive-stranded RNA progeny.⁶³

Mainly two viral enzymes catalyse this process; the NS3 helicase, which 'melts' double-stranded RNA into single strands in an ATP-dependent manner, and the NS5B RNA-dependent RNA polymerase (RDRP).⁶⁴⁻⁶⁹

1.7 HCV Vaccine Development

HCV vaccine development is a challenging task partly due to the lack of small animal models and cell culture systems capable of supporting reliable HCV replication, but predominantly because of the rapid evolution of the virus which continuously produce mutants that escape immune surveillance⁷⁰.

Attempts are being made to develop prophylactic as well as therapeutic vaccines against HCV infection. Therapeutic vaccines work by stimulating a suppressed immune system to produce more protective T-cells and antibodies⁷¹.

As yet, no prophylactic vaccine is available for HCV, but extensive studies of a recombinant vaccine in chimpanzees showed encouraging results.⁷²

Several approaches are also being taken to develop therapeutic vaccines. For example, a clinical-grade HCV E1 protein has been evaluated in clinical trials^{73,74}. Cellular immune responses were boosted, including a significant T-cell response. However, these cellular immune responses were not accompanied by any significant reductions in serum HCV RNA.⁷³⁻⁷⁶ Other approaches to therapeutic HCV vaccines include the use of the recombinant core protein, which elicited an unusually strong T-cell response to HCV in rhesus macaques.⁷⁷

1.8 Viral targets for future therapies

Recently, a number of advances have come about, owing to the development of HCV-replicon systems and the availability of the three-dimensional structures of several of the key virally encoded enzymes, in some cases in complex with substrates, cofactors and inhibitors.⁷⁸

In vitro assays have been developed to examine viral enzymatic activities for the testing and development of antiviral agents.⁷⁸

Potential processes for viral inhibition include virus entry into the host cell, proteolytic processing, RNA replication, and the assembly and release of the new virions. Among the most promising new agents in development are the protease and polymerase inhibitors. The structural viral envelope proteins E1 and E2, as well as their assembly, represent other potential antiviral targets⁷⁹.

1.8.1 Protease inhibitors

HCV possesses two proteases NS2/3 and NS3/NS4A. However, NS2/3 has received little attention as an antiviral target because of incomplete knowledge of the NS2/3 cleavage process. Furthermore, NS2/3 processing is partially mediated by host cell proteases, making it a less attractive target for drug development.^{80,81}

On the other hand, the non-structural protein NS3 has a protease domain that is responsible for polyprotein processing and is a potential target for antiviral intervention. Despite the catalytic site being a shallow and largely hydrophobic groove, making it difficult to target, several compound inhibitors of the NS3 protease have been successfully designed and reached preclinical and clinical stages, for example, BILN 2061, Intermune (ITMN-191), telaprevir (VX-950) and boceprevir (SCH503034) BI12202 (Figure 1.10).

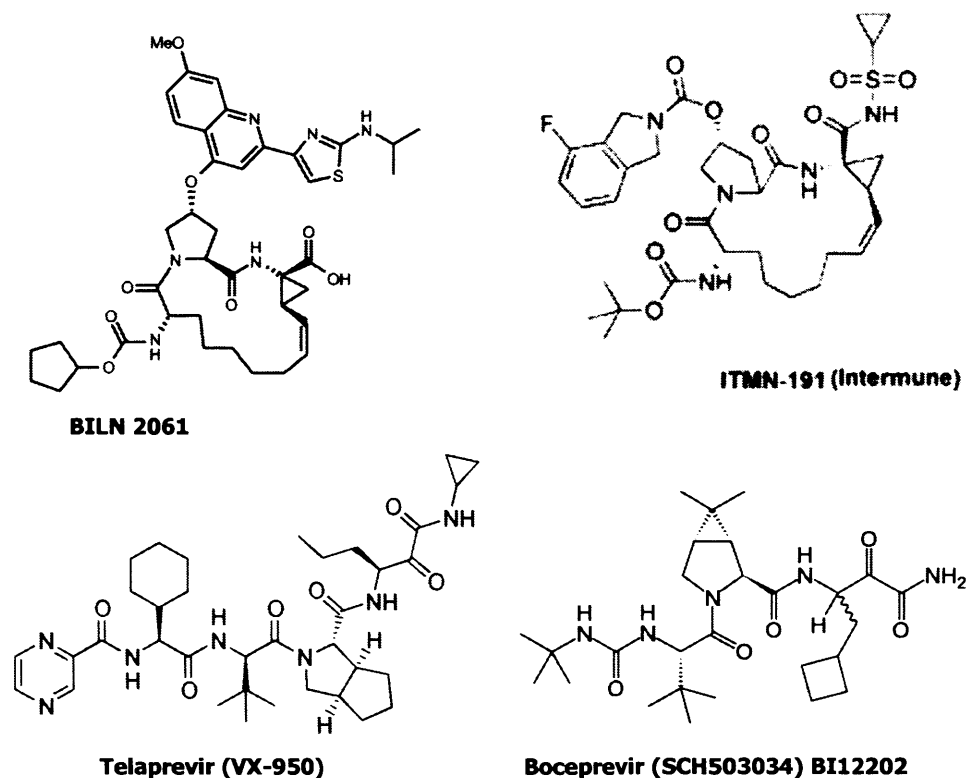


Figure 1.10 Structures of some NS3 protease inhibitors

The proof-of principle for this class of compounds was provided by BILN 2061, an NS3 protease inhibitor that provides at least a 2–3 \log_{10} decrease in HCV load within 48 hours⁸². However, the clinical development of BILN 2061 was stopped owing to significant side effects^{83–87}.

1.8.2 Polymerase inhibitors

NS5B protein functions as a RNA-dependent RNA polymerase (RDRP), which is virus specific with no functional homologues in the host and therefore its inhibitors expected to be less toxic.⁸⁸

RDRP is the key enzyme for synthesis of a complementary minus strand (-RNA), using the genome (+RNA) as a template, and the subsequent synthesis of genomic (+RNA) from this (-RNA) template.⁸⁸

Two separate classes of compounds have shown inhibitory effects on the NS5B through two distinct mechanisms:

- First, nucleoside polymerase inhibitors, which directly inhibit the active site causing chain termination for example, valopicitabine (NM-283), PSI-6130, MK-0608 and R1626⁸⁷ (Figure 1.11).
- Second, non-nucleoside polymerase inhibitors, which cause allosteric inhibition resulting in a conformational change of the protein for example HCV-796⁸⁷ (Figure 1.11).

Preclinical and clinical studies have shown that agents targeting the HCV RNA polymerase are associated with significant reductions in serum HCV RNA when used either as monotherapy or in combination with pegIFN- α . However, due to the unfavourable risk-benefit profiles, the development of several polymerase inhibitors is on hold.^{89–92}

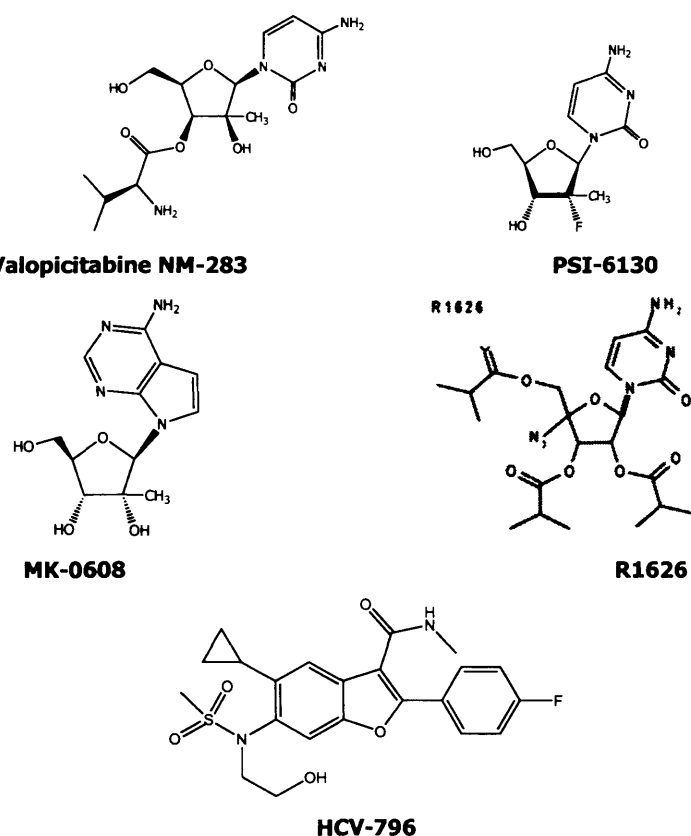


Figure 1.11 Structures of some NS5B polymerase inhibitors

1.8.3 Immune modulators

Other mechanisms that are under investigation include immune modulators targeting the cellular immune response, which plays a major role in HCV infection. Examples include agents that promote an effective immune response, such as the toll-like receptor (TLR) agonists, which mediates induction of interferon IFN- α production. However, improved SVR results have not been reported so far and the clinical development is currently on hold owing to preclinical safety issues; induction of a general inflammatory response in animals.^{93,94}

1.8.4 Further novel investigational agents

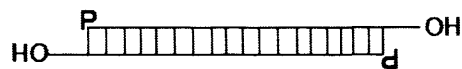
The effectiveness of inhibitors of cyclophilin B, a host factor involved in viral replication, is being evaluated in patients with HCV. A cyclosporin A analogue, suppresses HCV genome replication in a cell culture system and may provide a novel strategy for anti-HCV treatment⁹⁵⁻⁹⁷.

Recently, it has also been reported that NS4A, a cofactor for the NS3 protease, is a valid therapeutic target for chronic HCV infection, inhibiting the correct proteolytic processing of the HCV polyprotein and thereby the formation of a functional replication complex, consequently decreasing viral RNA synthesis, although reversible nephrotoxicity precluded further development.⁹⁸

Finally, the 5' NTR region is an internal ribosomal entry site (IRES) that binds to the ribosomes of the infected cell allowing the proviral RNA translation.⁹⁹ The IRES is highly conserved from isolate to isolate, with greater than 85% sequence identity among the various genotypes, making it an outstanding target for drug development.¹⁰⁰

Nucleic acid based therapies, designed to target the HCV IRES, have shown substantial success at inhibiting the HCV life cycle. Examples of this strategy are;

- Antisense oligonucleotides, which are synthetic nucleic acid polymer whose base sequence is complementary to the target RNA sequence (sense strand), thus it would bind to it and effectively turning it "off". This is because RNA has to be single stranded to be translated e.g. a sense segment " 5'-AAGGUC-3' " would be blocked by the anti-sense RNA segment " 3'-UUCCAG-5' ".¹⁰¹
- Ribozymes, which is also known as ribonucleic acid enzyme, RNA enzyme or catalytic RNA, is an RNA molecule that catalyses the hydrolysis of RNAs phosphodiester bonds.¹⁰²
- Small interfering RNA (siRNA) also known as short interfering or silencing RNA, is a natural process used by eukaryotes to recognize and destroy abnormal or exogenous RNA. siRNA is triggered by double stranded RNA regions (dsRNA) and in average consist of 20-25 nucleotide-long double-stranded RNA (dsRNA) with 2-nt overhangs on either end:



Each strand has a P (5' phosphate group) and a OH (3' hydroxyl group) to bring about the specific knockdown of a gene of interest.¹⁰³

However, some of the main issues regarding the success of the RNA-based therapies are the delivery of big synthetic RNA polymers to the appropriate cells besides the stability towards nucleases.¹⁰¹⁻¹⁰³

Chapter 2

Helicase Enzyme; Overview “The Unwinding Machine”

Chapter 2

Helicase Enzyme; Overview

2.1 Helicases (The Unwinding Machines)

The discovery of the structure of DNA double helix, proposed by Watson and Crick was a pivotal event in modern biology, showing two paired DNA strands that are tightly base-paired with the complementary nucleotides on the opposite strands. However, during DNA replication or gene transcription, a single stranded DNA is at least transiently required to convert a closed duplex into two open single strands so that other protein machinery can access the sequence of bases buried in the interior of the double helix. In fact, the unwinding of double-stranded polynucleotides is catalyzed by helicases that are present ubiquitously in all kingdoms of life from virus to man. Viruses, like all other organisms, synthesize their DNA or RNA genomes in a template-dependent manner. In addition to DNA or RNA polymerases, a helicase is therefore required to displace the single-stranded genome after replication, thus leading to the formation of progeny viral particles¹⁰⁴ (Figure 2.1).

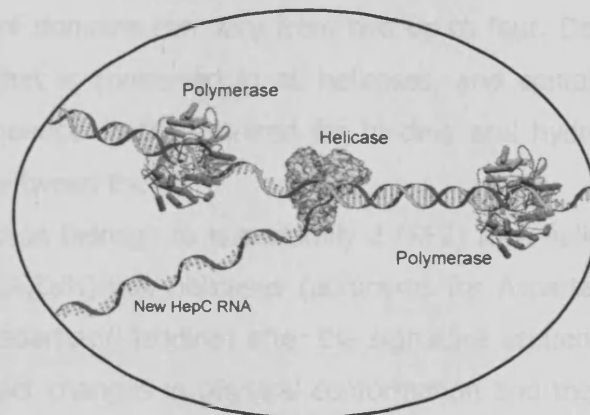


Figure 2.1 The polymerase synthesizes new RNA (red) from old RNA (gray). The helicase unwinds new RNA from old RNA so more RNA can be produced (adapted with permission from Dr. D. Frick)

2.2 Classification, Structure and function

Since the discovery of the first helicase in *E. coli* in 1976 many other similar enzymes have been isolated from viruses, prokaryotic and eukaryotic cells. At the most basic level, Helicases function as molecular motors that couple the consumption of chemical energy from NTP hydrolysis to mechanical force to unwind double stranded (ds) nucleic acids ¹⁰⁵⁻¹⁰⁸.

Further classification depends on a number of biochemical criteria; whether they can bind single stranded (ss) nucleic acid, unwind RNA, DNA, or both, the polarity of unwinding (3' to 5' or 5' to 3') and the presence of certain conserved motifs in the primary sequence. Generally, helicases have been divided into three superfamilies (SF1, SF2 and SF3) based on sequence comparisons. There are seven highly conserved motifs between the two largest families (SF1 and SF2) ¹⁰⁹.

Although, the conserved motifs are associated with conserved helicase function, structural analyses have identified crucial residues that are conserved only in space, but not in the primary sequence ¹¹⁰.

The structure of a number of helicases has been solved to show that the total number of domains can vary from two up to four. Domains 1 and 2 form a core that is conserved in all helicases, and contain most of the conserved sequence motifs required for binding and hydrolysing NTP at the interface between them ¹¹¹.

NS3 HCV helicase belongs to superfamily 2 (SF2) RNA helicases, which is also called DEA(D/H)-box helicases (acronyms for Aspartate, Glutamate, Alanine and Aspartate/Histidine) after the signature sequence of Walker B motif. The exact changes in physical conformation and the path the RNA or DNA takes as one strand is destabilised and 'unwound' from the other are not well understood, but they do not seem to be conserved between different helicases. Complicating matters further, some helicases seem to

function as monomers (HCV helicase), others as dimers and yet others as hexamers^{112,113}. Functionally, Helicases show a number of intrinsic biochemical activities:

2.2.1 NTPase activity

The hydrolysis of NTP to convey the chemical energy into mechanical force to move the helicase unidirectionally along the DNA is dependent on the presence of an NTP and a divalent cation, commonly Mg^{2+} . Basically, all helicases bind NTP using two structural elements at the interface between domains 1 and 2;

- Phosphate binding, P-loop, motif I or Walker A motif.
- Mg^{2+} co-factor binding loop, motif II or Walker B motif.

It is still not clear whether the energy derived from NTP hydrolysis is used for translocation or destabilising the initial ds DNA segments or both¹¹⁴.

2.2.2 Nucleic Acids Binding

To date, the most striking feature of helicases is that DNA or RNA binding is allosterically regulated by nucleotides (ATP or ADP)¹¹⁵. Also apparently there is a mechanistic distinction between SF1 and SF2 helicases in the nature of their interactions with polynucleotide substrates. In SF1 the ss polynucleotide binds mainly through a series of hydrophobic contacts formed by aromatic side chains stacking against the polynucleotide bases. However, a different mode is utilised by SF2 helicases where the protein polynucleotide interaction is predominantly through the phosphate backbone (sequence independent)¹¹⁶.

2.2.3 Unwinding Activity

The process of separating the complementary polynucleotide strands can be described in a number of terms;

Polarity is an intrinsic character of a given helicase and described as the directionality of unwinding that could be either 3'-5' or 5'-3'.

Processivity is the ability to repetitively continue the catalytic function without dissociating from the DNA substrate. Some helicases exhibit a high processivity with a separation of 42.3 kilo base pairs of DNA per binding event; other helicases can only separate several bases.

Unwinding Step-Size; A helicase unwinds DNA through a number of steps resulting in the unwinding of certain number of base pairs during each reaction cycle in which chemical energy is consumed and coupled to a mechanical energy ^{117,118}.

2.2.4 Translocating & Removing Proteins

An increasing body of evidence has shown that many helicases can translocate unidirectionally along ss polynucleotide track without unwinding activity. In fact, the removal of proteins from nucleic acid appears to be a core function of enzymes of the helicase families ¹⁰⁴.

2.2.5 Annealing Activity

In recent years, several DNA and RNA helicases have been shown to possess strand-annealing and even strand exchange activities in addition to their helicase activities ¹⁰⁴.

2.3 Viral Helicases as antiviral drug targets

Many viral pathogens, including hepatitis C virus (HCV), herpes simplex virus (HSV), human papillomavirus and severe acute respiratory syndrome (SARS), encode helicase enzymes, which are essential for viral replication and pathogenesis ¹¹⁹⁻¹²³.

The necessity of unwinding activity mediated by the NTPase/helicase enzymes in the viral life cycle has recently been reported in "knock-out" experiments that demonstrated unambiguously that the switch-off of the helicase activity abolishes the virus propagation of flaviviridae viruses like bovine diarrhoea virus (BVDV) and of dengue fever virus (DENV) ^{124,125}.

Recently, a preclinical proof of concept for helicase inhibitors as antiviral

agents has been obtained for Herpes simplex virus HSV¹²⁰⁻¹²³. Using high-throughput screening (HTS) and optimization of the screening hits resulted in compounds that inhibited HSV growth in cell culture with little cytotoxicity and which were also orally active in animal models¹²⁶⁻¹³². Over the past decade, significant progress has been made in the development of selective helicase inhibitors. Figure 2.2 shows a number of viral helicase inhibitors that have been developed and used in clinical studies as antiviral and anti-cancer drugs¹⁰⁴.

Virus species	Helicase name	Helicase family	<i>In vitro</i> activity	Inhibitor
Herpesvirus	UL5	1	DNA 5'-3' DNA helicase	T-157602 ¹²⁰ , BAY 7561293 ¹³⁰ , BILS-179BS ¹³²
papillomavirus	E1	3	DNA helicase	Biphenyl-4-sulfonyl acetic acid ¹³³
SARS	Nsp13	1	5'-3' RNA/DNA helicase, RTPase	Bananin, iodobananin, vanillinbananin and eubananin ¹³⁴
HCV	NS3	2	3'-5' RNA/DNA helicase	RDP, RTP ¹²⁶ , TBBT ¹³⁵ and QU663 ¹³⁷

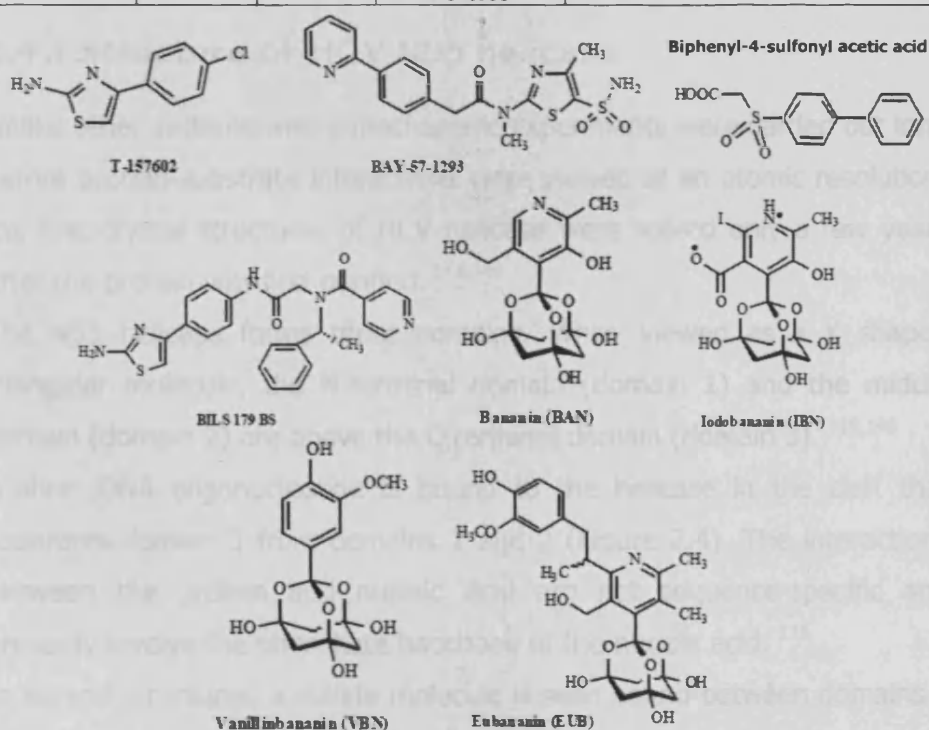


Figure 2.2 Examples of different Viral Helicase Inhibitors

2.4 HCV NS3 Helicase

The first HCV protein crystallised was the non structural protein (NS3) helicase (Figure 2.3), whose activity is thought to be required for melting secondary structures prior to initiation of RNA synthesis either to separate ds RNA intermediates formed during viral RNA synthesis or as a translocase that can remove proteins bound to viral RNA.¹³⁸

Strand separation is an energy-dependent reaction driven by ATP hydrolysis. Therefore, all RNA helicases have ATPase activity; this activity of NS3 RNA helicase can hydrolyse non-specifically any nucleoside triphosphate (hence known as "NTPase").¹³⁹⁻¹⁴²

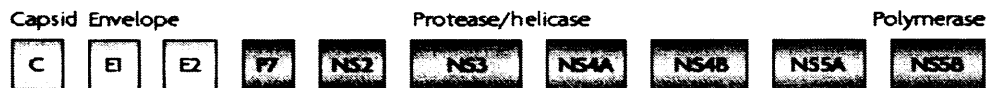


Figure 2.3 The encoded 10 proteins of HCV genome

2.4.1 Structure of HCV NS3 helicase

Unlike other systems where mechanistic experiments were carried out long before protein-substrate interactions were viewed at an atomic resolution, the first crystal structures of HCV helicase were solved only a few years after the protein was first purified.¹⁴³⁻¹⁴⁶

The NS3 helicase forms three domains. When viewed as a Y shaped triangular molecule, the N-terminal domain (domain 1) and the middle domain (domain 2) are above the C-terminal domain (domain 3).^{116,146}

A short DNA oligonucleotide is bound to the helicase in the cleft that separates domain 3 from domains 1 and 2 (Figure 2.4). The interactions between the protein and nucleic acid are not sequence-specific and primarily involve the phosphate backbone of the nucleic acid.¹¹⁶

In several structures, a sulfate molecule is seen bound between domains 1 and 2, in a position where ATP has been seen in high-resolution structures of similar helicases.¹⁴⁷⁻¹⁴⁹

The enzyme contains the seven conserved motifs associated with the Superfamily 2 (SF2) class of NTPases and RNA helicases.¹⁵⁰⁻¹⁵² These include the phosphate-binding P-loop, Walker motif A or motif I, located on the surface of domain 1, which binds the terminal phosphate group of the NTP and the Mg^{2+} co-factor binding loop, Walker motif B or motif II also on domain 1, and, responsible for the chelation of the Mg^{2+} ion.¹⁴³ On the other hand, Domain 2 possesses a highly conserved arginine-rich motif, which has been implicated in RNA binding; however, it has also been suggested that it may have a direct involvement in the binding of ATP. Unlike both domains 1 and 2, domain 3 is predominantly α -helical possessing no helicase-conserved motifs.^{152,153}

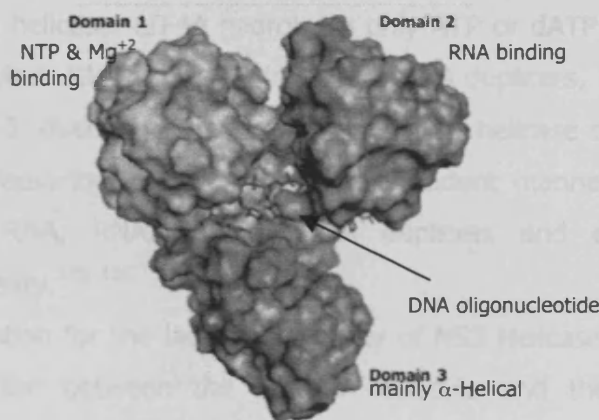


Figure 2.4 Surface diagram of the HCV RNA helicase showing Domain 1 (pink), Domain 2 (orange), Domain 3 (green) and DNA oligonucleotide (blue).

2.4.2 Mechanism of HCV NS3 Helicase action

Two alternative mechanisms of the unwinding reaction have been postulated. Both models predict that the enzyme binds and subsequently hydrolyzes NTPs in a well-defined NTP binding pocket: *In the active mechanism*; the energy released from the reaction is utilised in the translocation of the enzyme along the DNA or RNA structures, and the unwinding reaction results from capturing single strand regions that arise

due to fluctuations at the fork. Alternatively, *the passive mechanism*; the energy could be transferred to the fork and used for disruption of the hydrogen bonds that keep the strands together. ^{116,143,150-154}

2.4.3 Host vs Viral Helicases (Selectivity Window)

Comparison of the primary sequence of two SF2 helicases e.g. HCV helicase and human DDX3 reveals a paucity of identical residues and a tremendous divergence in sequence outside of the conserved motifs. The differences in primary sequence and tertiary structure between the helicase of a viral pathogen and that of cellular helicases can be exploited to confer specificity to an antiviral inhibitor. ¹¹⁰

Another example is the Eukaryotic initiation factor 4A (eIF4A), which is a human RNA helicase. eIF4A hydrolyses only ATP or dATP in the presence of RNA, unwinds RNA or RNA/DNA but not DNA duplexes, and acts in both 3'-5' and 5'-3' directions. In contrast, HCV NS3 helicase can hydrolyse all ribo- and deoxyribo-NTP in an RNA-independent manner, is capable of unwinding RNA, RNA/DNA, or DNA duplexes and exhibits a 3'-5' unidirectionality. ^{119, 155}

The explanation for the lack of specificity of NS3 Helicase is the fact that the interaction between the protein molecule and the DNA or RNA substrate is mediated by phosphate groups and not by the nucleotide base or sugar moieties. ^{116,143}

It is believed that the observed structural and functional differences between the human and viral helicases may prove to be useful in designing selective inhibitors.

2.4.4 HCV NS3 Helicase as Drug Target

The HCV NS3 helicase has been a more challenging target for drug development than other non-structural proteins (NS5B polymerase or NS3-NS4A protease), mainly because its mechanisms of action are not well

understood. There is also a problem with selectivity for compounds that inhibit via the ATP binding site and these compounds are likely to be cytotoxic to the host. At the technical level, traditional assays for screening helicases are very time consuming. New assays using DNA substrates have recently been developed in a high-throughput format but few options are available for assays using RNA as a substrate.¹²³

Moreover, NS3 helicase has broad substrate specificity, unwinding DNA-DNA, RNA-RNA and RNA-DNA heteroduplexes. This peculiar lack of substrate specificity of NS3 NTPase and helicase activities makes identification of a specific inhibitor a really challenging task.¹³⁸

In principle inhibition of the helicase enzyme has the potential not only to terminate the proliferation of the virus but also indirectly stimulate a cellular antiviral ds RNA sensing machinery because of the expected build up of intracellular ds RNA intermediates¹²⁶⁻¹²⁸. But because of the above-mentioned shortcomings, there are only a few HCV helicase inhibitors reported so far. Generally, strategies of helicase inhibition can be tentatively classified into three categories:

2.4.4.1 Direct/Allosteric Inhibition of NTPase Activity

It is believed that the ATPase cycle not only supplies the energy for the helicase unwinding reaction but also causes conformational changes in the nucleic acid binding site to drive the movement of the helicase along the length of the nucleic acid chain. Therefore, a reduction of accessibility to the NTP-binding site may lead to decreased NTPase activity and hence to a respective reduction of the unwinding rate.¹²³

A wide range of competitive NTPase inhibitors such as ribavirin 5'-triphosphate (RTP) and ribavirin 5'-diphosphate (RDP) (Figure 2.5) have been tested as potential helicase inhibitors. Surprisingly, although the IC₅₀ values possessed by each of these compounds against NTPase activity lie

in the low micromolar range, they demonstrate only moderate inhibitory action against the unwinding activity of HCV helicase.⁶⁷

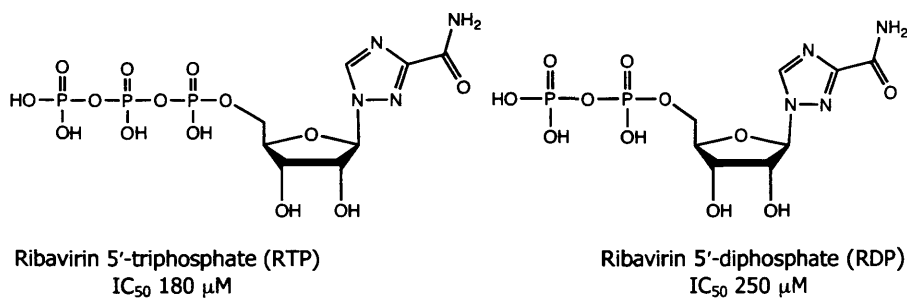


Figure 2.5 Structure of Ribavirin 5'-triphosphate and Ribavirin 5'-diphosphate

This phenomenon was also observed with paclitaxel (Figure 2.6), an antimitotic agent derived from the Western yew plant. Paclitaxel has been shown to compete with ATP at the ATP-binding domain, displaying an IC_{50} of 17 μ M against the NTPase activity of the enzyme. However, when tested for helicase inhibitory activity, paclitaxel was not capable of inhibiting the enzyme below 1 mM. This partial inhibition mediated by these competitive NTPase inhibitors is common to all members of the class, and the basis for the phenomenon remains unclear.¹⁵⁶

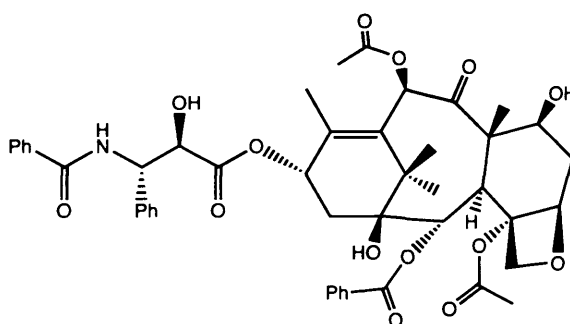


Figure 2.6 Structure of Paclitaxel (IC_{50} >1 mM)

It was recently reported that halogenated benzimidazoles and benzotriazoles were identified as inhibitors of NTPase/helicase activities of HCV during the course of a random screening study. Among these compounds, the 4,5,6,7-tetrabromobenzotriazole (TBBT), known as a

potent highly selective inhibitor of protein kinase 2 and 5,6-dichloro-1(β -D-ribofuranosyl)benzotriazole (DRBT) (Figure 2.7), displaying an inhibitory effect to the unwinding activity with IC_{50} values of 20 and 1.5 μ M, respectively, but not to the ATPase activity. Further kinetic study in combination with the structural comparison of ATP-binding sites has revealed that these compounds inhibit the unwinding activity through occupation of an allosteric binding site, rather than inhibition of the catalytic ATP-binding site.^{135,136}

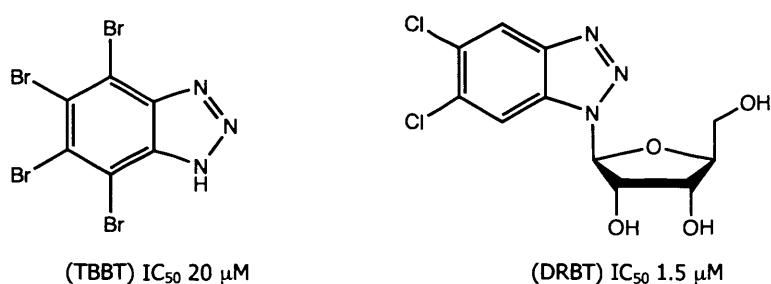


Figure 2.7 Structure of TBBT and DRBT

An additional class of nucleoside analogues known as ring-expanded (REN or "fat") nucleosides were reported to be active against the helicase unwinding reaction. A number of RENs such as **1** and **2** (Figure 2.8) displayed IC_{50} values in the micromolar range.¹⁵⁷

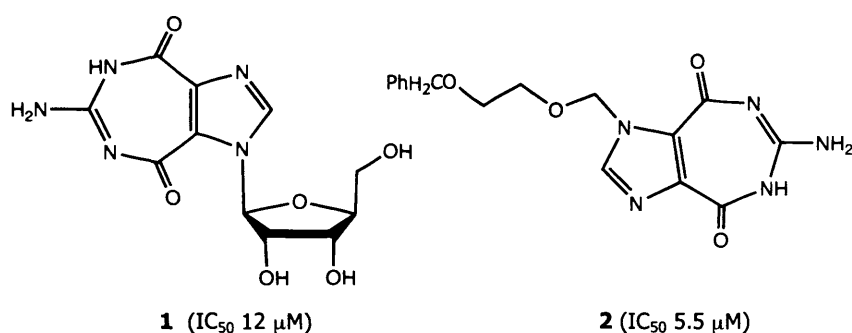


Figure 2.8 Structure of ring expanded nucleosides (REN)

The REN 5'-triphosphates such as **3** and **4** (Figure 2.9), on the other hand, did not influence the unwinding reaction but instead exerted their inhibitory effect on the ATPase activity of the enzyme ¹⁵⁷.

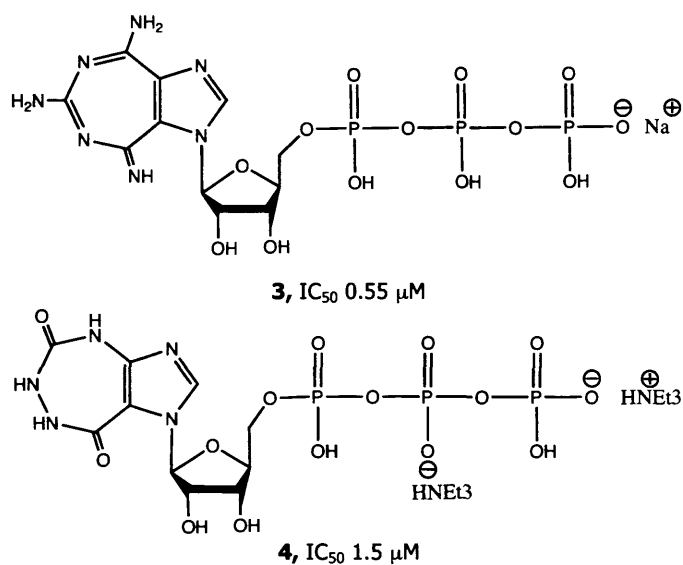


Figure 2.9 Structure of ring expanded 5-triphosphate nucleosides

The partial inhibition mediated by the competitive NTPase inhibitors may be avoided by utilising compounds chemically unrelated to NTP, which reduce the accessibility to the NTP-binding site in a non-competitive manner ¹⁵⁸. An example of such an inhibitor is the calmodulin antagonist Trifluoroperazine (Figure 2.10). Although the molecule is known to interact with domain 1 of HCV helicase, it is uncertain if inhibition results from conformational changes or from blockage of the ATP-binding site. ¹⁵⁶

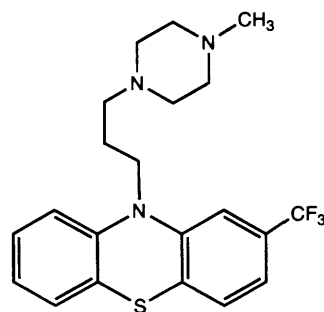


Figure 2.10 Structure of Trifluoroperazine (IC₅₀ 0.6 - 0.7 mM)

2.4.4.2 Inhibition of Unwinding through Intercalation

There is increasing evidence that DNA or RNA duplexes, together with an intercalated agent, are more stable than their unbound counterparts by increasing the energy required for duplex unwinding.¹⁵⁹⁻¹⁶¹

Therefore, DNA or RNA intercalating compounds such as Epirubicin and Nogalamycin (Figure 2.11), which are both members of the anthracycline family of anticancer/antibiotics, are effective inhibitors of the unwinding reaction catalysed by the helicase enzyme. Unfortunately the high cytotoxicity of such compounds and their weak penetration into the cell limit their application in the treatment of chronic viral infection.¹²⁶

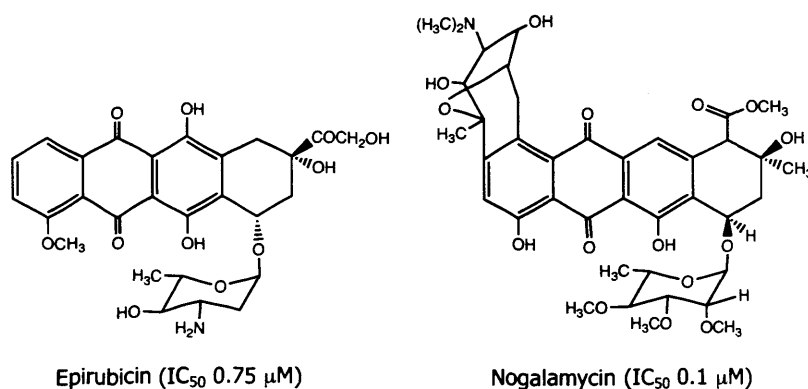


Figure 2.11 Structures of Epirubicin and Nogalamycin

2.4.4.3 Competitive Inhibition of RNA Binding

During the initial studies of HCV helicase activity, it was observed that numerous polynucleotides elicited an inhibitory response on the helicase enzyme. This is believed to result from the competition of the polynucleotides with DNA or RNA substrates, which could possibly be mimicked by synthetic macromolecules.¹⁵⁶ In an attempt to emulate this response, two series of compounds containing aminophenylbenzimidazole (**5**) and piperidinylbenzimidazole (**6**) moieties attached to symmetrical linkers were synthesised¹⁶² (Figure 2.12).

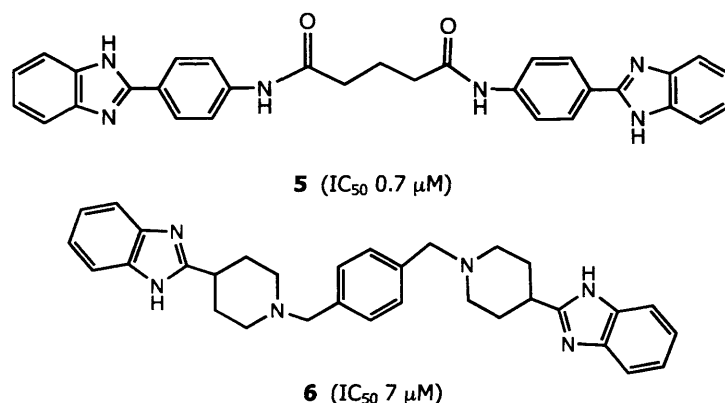
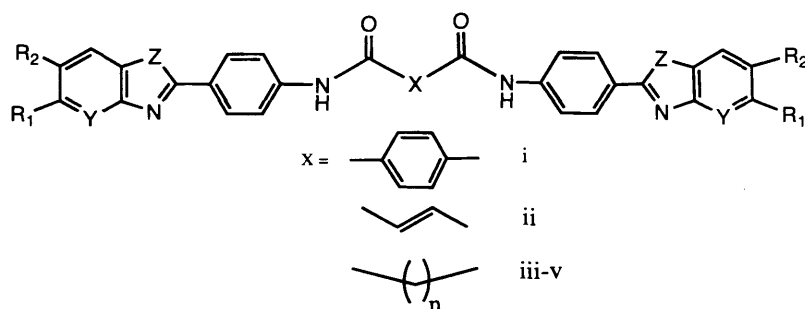


Figure 2.12 Structures of aminophenylbenzimidazole (**5**) and piperidinybenzimidazole (**6**)

Preliminary SAR studies of these compounds demonstrated a dramatic decrease in potency with the replacement of the benzimidazole moiety in **5** and **6** with the benzoxazole **7** a, c, d (i-v) and benzothiazole **7** b, e (i-v) moieties (Figure 2.12, 2.13). Similarly, the linker was also implicated in inhibitor activity because replacement of the diamide linkage in **5** with the diurea one, **9a** and **9b** (Figure 2.12), led to diminished potency. Thus, the SAR data indicate that the benzimidazole ring, the benzene group at the C2 position of the benzimidazole moiety, and the nature of the linker are essential for inhibitory activity.¹⁶³



Compound	Z	Y	R ₁	R ₂
7a	O	CH	H	H
7b	S	CH	H	H
7c	O	N	H	H
7d	O	CH	CH ₃	H
7e	S	CH	H	CH ₃

Figure 2.13 Structures of aminophenylbenzimidazole and benzothiadiazole derivatives

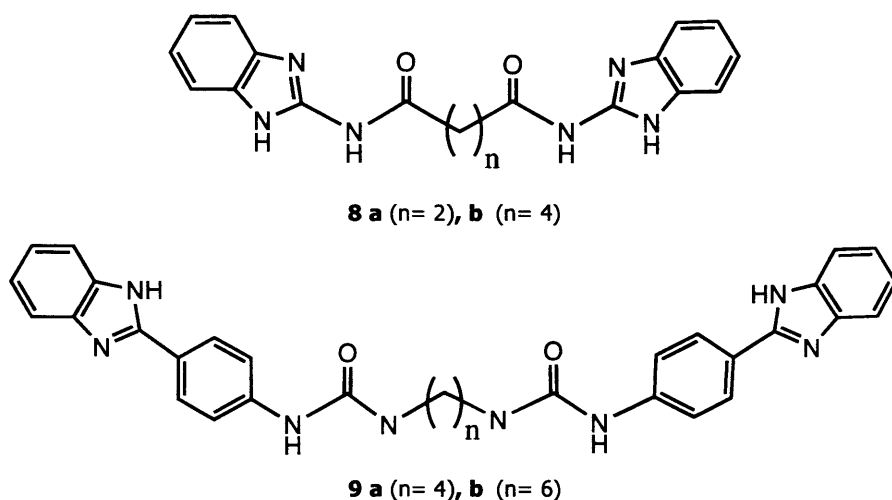


Figure 2.14 Structures of two aminobenzimidazole-derived diamides (**8a**, **b**) and two aminophenylbenzimidazole diureas (**9a**, **b**)

Recently, Maga *et al* have discovered QU663, a quinoline derivative (Figure 2.15) as a potent selective inhibitor of the NS3 helicase activity ($K_i=0.75 \mu\text{M}$). The study of the inhibition mechanism has revealed that QU663 is a competitive inhibitor of the strand-displacement activity, but without affecting the ability of NS3 helicase to hydrolyse ATP.¹³⁷

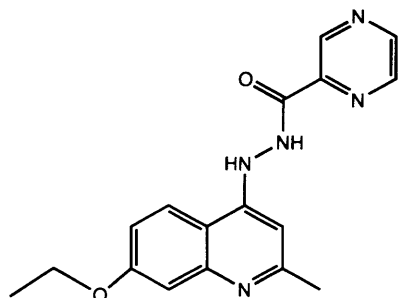


Figure 2.15 QU663 ($K_i = 0.75 \mu\text{M}$)

Chachulska and coworkers have reported a new class of tropolone derivatives that possess HCV NS3 helicase inhibitory activity. The most potent derivative was 3,7-dibromo-5-morpholinomethyl tropolone (DBMTr, Figure 2.16) ($\text{IC}_{50}=17 \mu\text{M}$). The mechanism of DBMTr may involve allosteric interaction with the enzyme but it is also probable that it competes with the enzyme for chelation of metal ions and ATP binding.¹⁶⁴

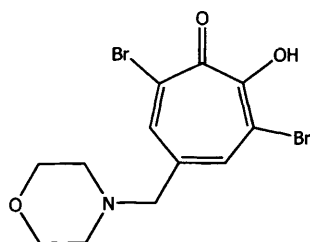
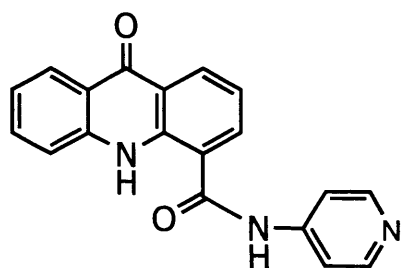
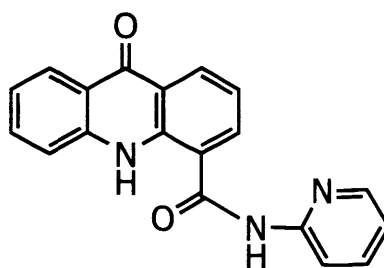


Figure 2.16 structure of 3, 7-dibromo-5-morpholinomethyltropolone (DBMTr)

More recently, a series of acridone-4-carboxylic acid derivatives (**12**, **13**, Figure 2.17) was reported to have anti-HCV replication activity, which was partially attributed to NS3 helicase inhibition in an uncertain mechanism. Nevertheless, intercalation of ds DNA is not completely excluded¹⁶⁵.



12 ($IC_{50}=9 \mu M$)



13 ($IC_{50}=3 \mu M$)

Figure 2.17 Structure of acridone-4-carboxylic acid derivatives

2.5 Research objectives

The standard treatment for chronic HCV infection so far is the dual pegylated IFN- α /ribavirin therapy¹⁶⁶. Although the mechanism of action of these drugs is debated, with both antiviral and immunostimulatory mechanisms being implicated, the sustained response rates are far from ideal. Moreover, there is substantial associated toxicity and the likelihood of success depends on viral and host factors that are often beyond the control of patients and physicians, in addition to the high cost¹⁶⁷⁻¹⁶⁹.

Among hepatitis C viral proteins, NS3 helicase is a valuable but yet largely unexplored target. NS3 helicase inhibition is a promising strategy because it has the potential not only to terminate the proliferation of the virus but also to indirectly stimulate a cellular antiviral response against double-stranded RNA^{126-128,170}.

NS3 helicase enzyme has multiple functional domains that present multiple potential mechanisms of action for the design of an inhibitor. RNA binding site was chosen because other strategies suffer from either lack of correlation with the unwinding activity or substantial cytotoxicity.

In this project, design and synthesis of a series of novel potential HCV NS3 helicase inhibitors was proposed. Starting from the structure of the small number of the reported compounds that target the RNA binding site, different computer software packages were used to design new families of compounds to block the action of the enzyme.

A small library of the newly designed compounds will be used for docking simulations. The results obtained *in silico* will guide the selection of the most promising compounds for synthesis then followed by subgenomic replicon and enzyme assays, starting a virtual "Design-synthesis-testing-optimisation" cycle.

Chapter 3

Results and Discussion

1. Molecular modelling
2. Mechanistic insights
3. Quinazolines
4. α , β Unsaturated ketones
5. Thiols
6. Thiadiazoles

Chapter 3

Part 1

Molecular modelling

During the 1960s, a limited number of scientists explored the use of mathematical techniques to derive relationships between observed activity and chemical structure. By 2007, computational chemistry was fully integrated into almost every aspect of the chemical research enterprise.

Molecular modelling science refers to a range of theoretical methods and techniques, which aim to simulate the behaviour of molecules using the equations of quantum and classical physics. Nowadays, molecular modelling is virtually synonymous with computational chemistry, given the diversity of different computational methods available to perform molecular modelling studies^{171,172}.

Computational chemistry thus comprises all computer-based methods such as quantum mechanics (QM), molecular mechanics (MM), molecular dynamics (MD), and conformational analysis, among others, used to understand and predict the molecular behaviour. Computational chemistry programs allow scientists to generate and present molecular data including geometries (bond lengths, bond angles, torsion angles), energies (heat of formation, activation energy, etc.), electronic properties (dipole moments, charges, ionisation potential, electron affinity), spectroscopic properties (vibrational modes, chemical shifts) and bulk properties (volumes, surface areas, diffusion, viscosity)^{173,174}.

It is a rather new branch of theoretical chemistry and is evolving very fast, following the evolution of computing hardware that provides the computational power to carry out more and more complex calculations. The drug discovery made by screening natural and synthetic compounds is an expensive and laborious process. In contrast, the impact of molecular

modelling on the process of drug discovery is significant, and destined to become even more important because virtual experiments are cheaper, faster, and safer than real experiments, and the data can help scientists to eliminate compounds that would not perform the required function¹⁷².

In the age of human genome being deciphered and the rapidly growing number of available three dimensional structure of therapeutically relevant macromolecules, coupled with the huge number of small non-peptide potential drug candidates easily available (over 7 million compounds), the need of using computer-aided techniques for the efficient identification and optimisation of novel hit compounds is highlighted^{172,173}.

In drug design, potential compounds may be conceptualised for the performance of required function based on essential characteristics (pharmacophores), including idealised structural and physical properties. Molecular models of compounds can be built, and virtual (*in silico*) tests may be run to assess their suitability before an expensive synthesis attempt is made. One successful example is the development of zanamavir (Relenza®), a potent Influenza neuroaminidase inhibitor¹⁷¹⁻¹⁷⁴.

Computer aided drug design (CADD) approaches are traditionally divided into two major categories: structure based (SBDD) and ligand based (LBDD). Structure based approaches are based on the key-lock metaphore for ligand-target interaction. They are increasingly used due to the large number of target proteins with known structure, mainly obtained by X-ray crystallography and NMR spectroscopy, becoming more available¹⁷³.

On the other hand, ligand based approaches are generally used when the structures of the target receptors are not known. They rely on chemometric functions to compute various descriptors from the 2D or 3D structure of a set of molecules and use statistical methods to build correlation models to predict biological properties (such as activity) of

compounds. Pharmacophore search and quantitative structure activity relationship "QSAR" studies are examples of LBDD¹⁷³.

As with all models however, the chemist's intuition and training is necessary to interpret the results appropriately as paraphrased by Henry Bent "A model must be wrong, in some respects, else it would be the thing itself, the trick is to see where it is right". Comparison to experimental data, where available, is also important to guide both laboratory and computational work¹⁷³.

The following sections summarise the background for the methods used in the present work, and also the specific application of molecular modelling techniques to drug design.

3.1.1 Molecular Mechanics (MM)

The term molecular mechanics (MM) refers to the application of Newtonian mechanics to the modelling of molecular systems. MM methods calculate the energy of a system as a function of the nuclear position of the atoms. This is possible due to the Born-Oppenheimer approximations, which states that due to the large difference in mass between nuclei and electrons, the latter move much faster and adjust instantaneously to the positions of the former¹⁷³⁻¹⁷⁶.

Thus MM methods ignore the electronic movements and allow calculations on large systems (e.g. proteins), which would be virtually impossible and extremely time-expensive were QM methods used. This is because the latter explicitly represent the electrons in the calculations, thus implying a much larger number of particles to be considered, when compared to MM methods¹⁷³⁻¹⁷⁶.

In order to lower its complexity, the physical mathematical description of molecules in a MM simulation adopts several simplifications, some of the most important are:

- Molecules are treated as balls-and-springs systems.
- Bonds are unbreakable.
- Electrons are not considered.
- Atoms are treated as solid spheres with fixed radius (typically the Van der Waals radius) and charge.
- Atoms differ among themselves for atom-type (depending on atomic number, hybridisation and chemical properties), a molecule is therefore reduced to a set of atom types and a bonding list.

Although most MM applications consider each atom as a separate entity, variations on this theme are possible; for example, many simulations have historically used a "united-atom" representation in which methyl and methylene groups were represented as a single particle, and large protein systems are commonly simulated using a "bead" model that assigns two to four particles per amino acid. Inside the wide environment of molecular mechanics several specific applications can be performed. The following sections describe some of the most important ones in computer aided drug design (CADD) ¹⁷³.

3.1.2 Force fields

MM force fields are sets of mathematical equations, which have been developed to compute various properties from atomic coordinates and atom types. Force fields are usually parameterised directly to reproduce potential energies.

Force fields compute both bonded and non-bonded interaction energies using both *ab initio* QM (from classical physics equations) and *empirical* functions and parameters (from experimental or calculated data, thus providing the total potential energy of a molecular system as a mathematical function of its inner variables (usually atom coordinates). Many of the force fields functions (the total molecular mechanics energy of a system; E_{MM}) currently in use can be calculated from the contributions of

a fairly simple four-component picture of the inter- and intra-molecular forces within the system such as bond stretching, bending and torsion, and also from interactions between non-bonded atoms such as electrostatic and van der Waals interactions (Figure & Equation 3.1) ¹⁷³;

$$E_{MM} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{non-bond}} + E_{\text{miscellaneous}} \quad (\text{Equation 3.1})$$

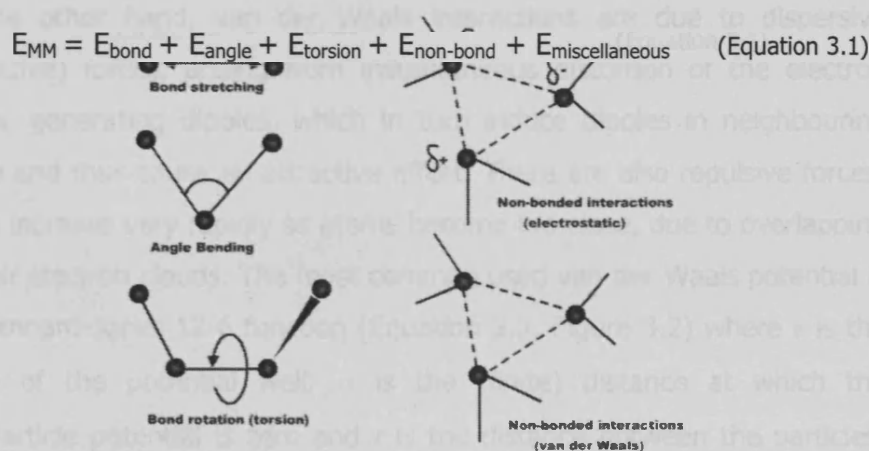


Figure 3.1 The four key contributors to a molecular mechanics force field ¹⁷³

The bond stretching term (E_{bond}) in the more simplistic form is given by the application of Hooke's law, where the bond is treated as a spring, the potential energy being proportional to the deviation from the equilibrium point. The bond angle energy (E_{angle}) can also be treated in a simplistic form where it is proportional to the square of the displacement from equilibrium. The torsion angle energy (E_{torsion}) models the change in energy with rotation of the bond, incorporating the energy barrier to rotation, dependent on the torsion angle defined by the four atoms involved.

The energy terms of the non-bonded parts of the system ($E_{\text{non-bond}}$) are modelled in two components; electrostatic and van der Waals interactions. Electrostatic interactions are calculated using Coulomb's law (Equation 3.2) following assignment of partial charges to atoms in a molecule. The electrostatic energy (E_{elec}) decreases when the distance r between two atoms of partial charges Q_1 and Q_2 increases. It is also dependent on the

dielectric constant ϵ of the medium. Partial charges are used due to the unequal distribution of charge in molecules, caused by the differences in electronegativity of atoms, and can be determined in numerous ways, varying significantly between different force fields.

$$E_{elec} = \frac{1}{\epsilon} \frac{Q_1 Q_2}{r} \quad (\text{Equation 3.2})$$

On the other hand, van der Waals interactions are due to dispersive (attractive) forces, arising from instantaneous distortion of the electron clouds, generating dipoles, which in turn induce dipoles in neighbouring atoms and thus cause an attractive effect. There are also repulsive forces, which increase very rapidly as atoms become too close, due to overlapping of their electron clouds. The most common used van der Waals potential is the Lennard-Jones 12-6 function (Equation 3.3, Figure 3.2) where ϵ is the depth of the potential well, σ is the (finite) distance at which the interparticle potential is zero and r is the distance between the particles. The $\exp(12)$ term of the equation is responsible for small distance repulsion, whereas the $\exp(6)$ provides an attractive term which approaches zero as the distance between the two atoms (r) increases. More sophisticated force fields may have additional terms ($E_{miscellaneous}$), but they invariably contain these four components^{173,174}.

$$E_{vdW} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (\text{Equation 3.3})$$

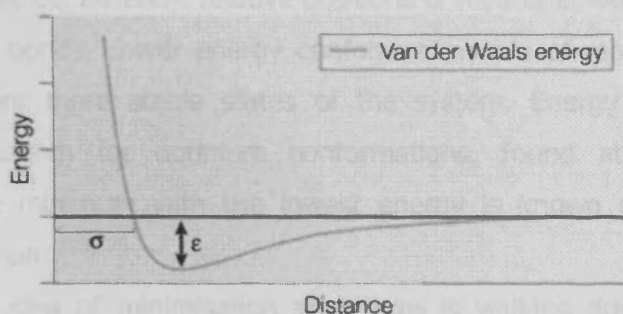


Figure 3.2 Lennard-Jones potential (12-6) function¹⁷⁴

Energetic penalties are associated with the deviation of bonds and angles away from their "reference" or "equilibrium" values. Several force fields have been developed to address specific problems, thus emphasising parameters, which are important for those systems. Therefore, there is not a single force field, which is best; rather there are force fields, which are more appropriate for specific study objectives than others^{173,174}.

Parameterisation and transferability are key issues in force fields. Force fields parameters are generally derived from experimental data (*empirical*) or from quantum mechanical calculations (*ab initio*) for small molecules. The transferability assumption implies that these parameters can be applied to larger molecules with the same functionalities^{176,177}. The choice of force field depends on the task to be done and the system to be studied. For example, the AMBER force field¹⁷⁸ was designed for the simulation of proteins and nucleic acids, whereas MMFF (Merck Molecular force field) aimed to achieve high accuracy for small organic molecules but also for large proteins¹⁷⁹.

3.1.3 Energy Minimisation

Energy Minimization is used synonymously with geometry optimization. Force fields compute the total potential energy of a molecular system (E_{MM}) but this has no much meaning if there is no knowledge about the state of the system in study. Molecules can assume different conformations, i.e. different relative positions of its atoms, due to rotations about single bonds. Lower energy conformations are of most interest as they represent more stable states of the system. Energy minimisation algorithms search for optimum conformations, found at the energy minima. The minimum with the lowest energy is known as the global energy minimum^{173,175}.

The general idea of minimisation algorithms is walking downhill on the energy surface from a starting point until you reach a minimum point. In a

simplistic way, minimisation algorithms involve consecutive small adjustments in the geometry of the molecules followed by energy calculation. Changes that result in increased energy are discarded as the method searched for lower energies, and the process continues to iterate until a minimum is found. This happens when any change made in the molecule's geometry invariably results in increased energy.

Obviously this method can just find one minimum point, the one nearest to the starting point. Sometimes, however, it is important to find the global minimum, for instance to predict the most stable conformation of a molecule under given conditions, therefore conformation search methods should be used in conjunction with energy minimisation algorithms using either systematic or stochastic approaches to generate starting points for several minimisation runs^{173,174}.

Analytical study of the energy function (known also as energy surface or energy landscape) provides important insight on the conformational states of a particular molecule. A minimum point in the energy function (local minimum) is associated with a stable state of the system. Flexible molecules generally contain more local minima than more rigid ones^{173,174}.

Minimisation algorithms are classified in two major groups: non-derivative methods (e.g. simplex method), which don't require the calculation of the derivative of the potential energy surface, and derivative methods (e.g. steepest descent, conjugate gradient, Newton-Raphson), which use these derivatives to guide the process to the next state (geometry) of the molecule.^{173,175}

Energy minimisation finds its application in several MM and MD experiments, such as in generating good quality conformations of a given molecule or in refining molecular docking results or homology models.¹⁷⁴

3.1.4 Docking Simulation

Docking of small molecules to protein binding sites was pioneered during the early 1980s. In molecular docking simulations modelling techniques are used to predict how macromolecules (typically a protein) interact with other molecules (may be other proteins, nucleic acids or small drug-like molecules). The process involves the prediction of ligand conformation and orientation (posing) within a binding site ¹⁸⁰.

Small molecule-protein docking (usually known as protein-ligand docking) is the most frequently used as the smaller dimensions of the system allow for a higher accuracy of the mathematical description and a higher reliability of the model. Therefore, it is used as the base of structure based drug design approaches (SBDD). Protein-protein docking is still at its beginning, as its high requirements of computational power has long made it too slow or too inaccurate ¹⁸¹.

Docking and scoring technology is applied at different stages of the drug discovery process for three main purposes: (1) predicting the binding mode of a known active ligand; (2) identifying new ligands using virtual screening; (3) predicting the binding affinities of related compounds from a known active series. Of these three challenges, the first represents the most straightforward and is the area where most success has been achieved. Docking can also contribute to the analysis of drug metabolism using structures such as cytochrome P450 isoforms ^{182,183}.

The docking process consists of two stages: the first stage is Searching and Filtering using docking algorithms for a step-by-step search through all the possible docking poses (not only the conformational, but also the orientational space) in the active site. This is followed by stringent filtering criteria, which retain only a fraction of the possibilities. The second stage is Scoring via scoring functions to evaluate the likelihood of each retained

solution of being an accurate model of the complex between compounds and potential targets^{173, 184}.

These are complex issues since docking a ligand into a binding site models several degrees of freedom; the six degrees of translational and rotational freedom of one body relative to another and then the conformational degrees of freedom of the ligand and of the protein¹⁷⁴. Sampling these degrees of freedom should be both accurate and fast enough to permit the evaluation of up to thousands of compounds in a given docking run^{184,185}.

It should also be noted that ligand-binding events are driven by a combination of enthalpic and entropic effects, and that either entropy or enthalpy can dominate specific interactions.

Enthalpic contributions result from ligand binding not in its global minimum, but its "bioactive" conformation, which might be significantly different and higher in energy. Furthermore, polar ligand groups have to break bonds to solvent molecules before binding to the protein. The entropic contributions are related to losses of rotational, translational and conformational freedom during the ligand binding process (entropy penalty for immobilisation of rotatable bonds).

This often presents a conceptual problem for contemporary scoring functions, because most of them are much more focused on capturing energetic rather than entropic effects. In addition some factors like, limited resolution of crystallographic targets, inherent flexibility, induced fit or other conformational changes that occur on binding, and the participation of water molecules in protein–ligand interactions add to this challenge¹⁷³.

3.1.4.1 Docking Algorithms

Docking algorithms are used as search methods to treat ligands flexibility and, to some extent protein flexibility. Treatment of ligand flexibility can be divided into three basic categories¹⁸⁶:

- **Systematic methods**

Incremental construction and Multiconformer database

Exploring all degrees of freedom in a molecule, will lead to the problem of combinatorial explosion because the exhaustive search calculation by varying systematically each of the torsional angles of the molecule in order to generate all possible conformations will yield huge amount of data ¹⁷³.

Therefore, ligands are often incrementally grown into active sites. A stepwise or incremental search can be accomplished in different ways for example, by splitting into different rigid fragments and perform an incremental construction of the solution by linking them covalently (used in *de novo* ligand-design strategy) or, alternatively, by dividing ligands into rigid (core fragment) and flexible parts (side chains). In the latter case, once the rigid cores have been defined, they are docked into the active site. Next, flexible regions are added in an incremental fashion ¹⁸⁷⁻¹⁸⁹. The docking of single fragments is much easier than the one of a full molecule, as they can be treated as rigid bodies. An approach of this kind is employed by DOCK and FlexX ¹⁹⁰. Another method of systematic search is Multiconformer docking, which is performed by pre-generating a library of low-energy conformations for the input molecule. Library conformations are typically only calculated once and the search problem is therefore reduced to a rigid body docking procedure as used in FRED ¹⁷⁴.

- **Random or stochastic methods**

Monte Carlo, Genetic Algorithm and Tabu search

These methods imply stochastic (random) steps in the search to speed up the process. They cannot guarantee to find the best solution, but are used to find a good one in a reasonable time. They operate by making random changes to either a single ligand or a population of ligands. A newly obtained ligand is evaluated by a pre-defined probability function. Two

popular random approaches are Monte Carlo, where the actual conformation is modified randomly to obtain a new one. This is repeated until the desired number of conformations is obtained (AutoDock)¹⁹¹ and Genetic Algorithms where the principles of biological evolution are adapted to search the conformational space (GOLD)¹⁹²⁻¹⁹⁵.

Another stochastic method is tabu search. The basic idea of a tabu search algorithm is to take into consideration already explored areas of conformational space. To determine whether a molecular conformation is accepted or not, the root mean square deviation (rmsd) is calculated between current molecular coordinates and every molecule's previously recorded conformation (PRO LEADS)¹⁹⁶⁻¹⁹⁸.

- **Simulation methods**

Molecular dynamics and Energy minimization

Molecular dynamics (MD) is currently the most popular simulation approach. However, MD simulations are often unable to cross high-energy barriers within feasible simulation time periods, and therefore might only accommodate ligands in local minima of the energy surface. Therefore, an attempt is often made to simulate different parts of a protein–ligand system at different temperatures^{199,200}. Another strategy for addressing the local minima problem is starting MD calculations from different ligand positions. In contrast to MD, energy minimization methods are rarely used as stand-alone search techniques, as only local energy minima can be reached, but often complement other search methods, including Monte Carlo²⁰¹. DOCK performs a minimization step after each fragment addition, followed by a final minimization before scoring¹⁷³.

3.1.4.2 Shape complementarity

The use of shape complementarity between the ligand and the active site as a pre filtering technique of large databases offers faster screening rates.

Studies with this tool can be initiated with only one bioactive lead or receptor site^{202, 203}.

3.1.4.3 Protein Flexibility

Available docking programs usually assume the protein structure to be rigid or allow limited flexibility, essentially in some amino-acid side chains. Flexible docking clearly outperforms rigid-body docking²⁰², however, the treatment of protein flexibility is less advanced than that of ligand flexibility, but various approaches have been applied to flexibly model at least part of the target²⁰⁴, including molecular dynamics and Monte Carlo calculations²⁰⁵, rotamer libraries²⁰⁶ and protein ensemble grids²⁰⁸. The idea behind using aminoacid side-chain rotamer libraries is to model protein conformational space on the basis of a limited number of experimentally observed and preferred side-chain conformations²⁰⁷. Another method of treating protein flexibility is to use ensembles of protein conformations (rather than a single one) as the target for docking²⁰⁸ and to map these ensembles on a grid representation²⁰⁹.

3.1.5 Scoring Functions

The general hypothesis of docking is that lower energy scores represent better protein-ligand bindings. The current pose is then accepted or rejected on the basis of the score for that pose. A new pose is then generated, and the search process iterates to an end point^{174,210}.

The evaluation and ranking of predicted docked poses is performed by using scoring functions which estimates the binding affinity of potential ligands to their targets. This is a crucial aspect of the docking process, particularly in virtual screening. Like force fields, scoring functions are mathematical functions that link atomic coordinates to energy values. They are usually parameterised to reproduce the interactions of intermolecular complexes. Scoring functions are usually much simpler than force field

functions, as they are the speed determining point in a docking algorithm. They usually focus on non-bonded interactions, and many of them consider bonded energies constant. Some scoring functions include empiric terms such as hydrogen bond energies or an estimate of the entropic loss of ligands upon binding or solvation effect. Essentially, there are three major kinds of scoring functions:

- **Force-Field-based scoring functions**

These functions are based on MM force fields, which usually quantify the sum of two energies, the receptor-ligand interaction energy and internal ligand energy (such as steric strain induced by binding). This is often described by using van der Waals (12-6 Lennard-Jones) and electrostatic (Coulomb's law) energy terms. Standard force-field scoring functions have major limitations, because they were originally formulated to model gas-phase and do not account for solvation and entropic effects, they only consider a single protein conformation and generally are computationally more costly than the other two types of scoring functions¹⁷⁴.

- **Empirical scoring functions**

The idea of empirical scoring functions is to reproduce experimental affinity data such as binding energies and/or conformations. This is based on the assumption that predicted binding energies can be approximated by a sum of individual uncorrelated terms. The parameters of the energy contributions are obtained from a regression analysis using experimentally determined binding energies and X-ray crystal structures of a training set complexes. Empirical scoring functions are commonly supplemented by entropy penalty terms, such as the number of rotatable bonds in ligands. Interaction types include, for example, hydrogen bonds, electrostatic interactions and hydrophobic interactions. The regression analysis requires both known structures and binding constants, and so the available

datasets are limited in size and often feature similar ligands and receptors. They are simpler and usually perform well, but they also show a strong dependence on the training set ²¹¹⁻²¹⁶.

- **Knowledge-based scoring functions**

Knowledge-based scoring functions have become popular in the field of docking during the past few years. Knowledge-based scoring is grounded on a statistical analysis of ligand–receptor complex structures. They are based on the assumption that atom configurations that are often seen in crystallographic structures are energetically favoured. The frequencies of each possible pair of atoms in contact to each other are determined. Interactions found to occur more frequently than would be randomly expected are considered attractive; interactions that occur less frequently are considered repulsive. A number of atom-type interactions are defined depending on their molecular environment. So, in common with empirical methods, knowledge-based scoring functions attempt to implicitly capture binding effects that are difficult to model explicitly. However, only structural information rather than binding energies is necessary to derive these frequencies so that a greater number and diversity of structures can be included in the analysis. Thus, less bias is expected compared with empirical scoring functions. A major attraction of many knowledge-based scoring functions is their computational simplicity, which permits efficient screening of large compound databases. A disadvantage is that their derivation is essentially based on information implicitly encoded in limited sets of protein–ligand complex structures ^{173,174}.

- **Consensus scoring**

Given the imperfections of current scoring functions, a recent trend in this field has been the introduction of consensus scoring schemes ²¹⁷. Consensus scoring combines information from different scores to balance

errors in single scores and improve the probability of identifying 'true' ligands. However, the potential value of consensus scoring might be limited, if terms in different scoring functions are significantly correlated, which could amplify calculation errors, rather than balance them¹⁷⁴.

3.1.6 Structural based Virtual Screening

Virtual screening experiments are based on the application of molecular docking algorithms and scoring functions techniques to large numbers of molecules when the structure of the target receptor is known. It has become common practice to screen hundreds of thousands or even millions of drug-like candidates in this way to find a drug lead for a particular target. Compared to traditional in-vitro high-throughput screenings, virtual screenings are orders of magnitude faster and less expensive representing a pre filter tool for selection of compounds to be experimentally tested. Successful candidates can thus provide new ideas for alternative structures, unrelated to natural or known ligands^{174, 202}.

Pharmaceutical companies generally search their own corporate database of compounds and can also screen libraries of commercially available compounds, from different suppliers, which are generally available in file formats that can be used directly. Recently, the ZINC (zinc is not commercial) database²¹⁸ was created by compilation of databases of commercially available compounds from numerous suppliers, and converted into 3D format, in various file types. This database is freely accessible for academic purposes. Another database is MMsINC, which is a free web-oriented database of non-redundant commercially available chemical compounds in 3D formats for virtual screening and chemoinformatic applications.²¹⁹ Before docking of the compounds of the database into the target receptor, filters can be applied to remove undesirable compounds, such as Lipinski's rule of five where molecules

with more than 5 H-bond donors, 10 H-bond acceptors, Molecular weight >500 Da and calculated partition coefficient (c log P) >5 are removed ²²⁰.

A virtual screening process can be divided into three phases:

- **Docking phase**

The binding pose for each molecule of the database is generated with molecular docking simulation.

- **Scoring phase**

Every molecule in the database is scored with one or more scoring functions and sorted accordingly to the predicted affinity.

- **Visual inspection**

The best ranking molecules are inspected by the user for specific interaction with the active site (e.g. specific H-bond), where the most promising ones are selected for synthesis and biological testing.

3.1.7 *De Novo* Drug Design

De novo drug design approaches use structural information to design bioactive compounds from scratch by incremental construction of a ligand model within the active site of the receptor or enzyme, the structure of which is known from X-ray or NMR data. This is to inspire medicinal chemists with new chemical motifs, support hit and lead identification efforts and widen the chemical horizon ^{221,222}. In this context, a medicinal chemist and *de novo* molecule design software are confronted with the challenge of cherry picking the most promising candidates in an infinite chemically feasible space of drug-like molecules ^{100, 223-225}.

The ultimate outcome of *de novo* design process is to identify isofunctional structures with different backbone architectures, to obtain new lead series with improved properties or to circumvent intellectual property constraints, what is described as "Scaffold Hopping". ²²²

Essentially, *De novo* design is faced with the problem of combinatorial explosion due to the huge number of different element types, the way they can be linked, the large number of theoretically possible topologies and the variety of conformations for a single topology. This renders exhaustive search for all solutions an impossible task. For this a number of constraints are to be applied during the *de novo* design process²²².

3.1.7.1 Primary target constraints

All information that is related to the ligand receptor interactions gathered from the 3D receptor structure and known ligands. First, the binding site is examined to derive shape complementarity constraints for a ligand, then specific non-covalent ligand–receptor interactions in the form of “hypothetical Interaction sites” are identified. Interaction sites are positions in space that is not occupied by the receptor and in which a ligand atom favourably interacts with the receptor. They are typically subdivided into H-bonds, electrostatic and hydrophobic interactions. Receptor groups capable of H-bonding are of special interest owing to its strongly directional nature. Key interaction sites define the explicit requirements for successful receptor–ligand binding and thus have a major role in reducing the vast number of possible structures^{222 - 228}.

3.1.7.2 Secondary target constraints

These include essential drug properties other than the binding affinity; such as “ADMET” properties where candidate compounds with suitable absorption, distribution, metabolism, excretion and toxicity can be predicted by filtering the designed structures with Lipinski’s rule of five²²⁰, however a lowered upper limit for the molecular mass of 350 and c logP of 3 might be more appropriate for the proposed molecules to be potential leads²²⁹. Also chemical stability is sometimes addressed as a list of undesirable substructures. Finally, synthetic accessibility is dealt with in

some programs via assembling the building blocks in accordance with a set of virtual organic reaction schemes.²²²

3.1.7.3 Structure sampling

The basic building blocks for the assembly of candidate structures can be either single atoms or fragments. Atom-based approaches are superior in terms of the structural variety. Fragment-based design strategies, on the other hand, significantly reduce the size of the search space by using fragments that are usually common in drug molecules²²².

There are several general concepts of structure sampling; the following are some of them:

a. Fragment placement and Connection (Linking approach)

Building blocks are placed at key interaction sites of the receptor either by the *de novo* program itself or by another program (pre-docked building blocks) then automatically connected by linkers to yield a complete molecule that satisfies all key interaction sites²²².

b. Sequential growth (Growing procedure)

A single building block point is selected by the program, the user or given by a predocked fragment and placed at one of the key interaction sites of the receptor. The structure is then grown to provide suitable interactions for the other key interaction sites of the receptor. Later, substituents are linked to the main scaffold with user-defined attachment sites²²².

c. Atomic lattice strategy

The placement of an atomic lattice of a non-physical atoms, which can be either sp^3 carbon atoms (diamond lattice)²³⁰, randomly and evenly distributed atoms²²⁷, or pre-docked fragments²²⁶ in the binding site. Lattice atoms in the vicinity of different interaction sites are joined through the shortest path. Atoms that are part of the shortest path are connected by newly formed bonds.²²²

Also few *de novo* design methods implement a MD simulation for structure sampling. Initially, building blocks are randomly positioned in the binding site. Covalent connections are formed among the building blocks in a stochastic and reversible manner to dynamically evolve candidate compounds.²³⁰⁻²³⁴

More recently, a haptic-driven molecular modeling software was introduced to the *de novo* drug design field, where the chemist can test his hypothesis by experiencing the energy profile of the protein ligand interactions in an intuitive way through the combined use of sight and touch²³⁵.

3.1.8 Molecular Dynamics (MD)

While Molecular Mechanics (MM) is concerned with atomic interactions, 3D geometries and energies in a given moment, Molecular Dynamics (MD) aims to predict the temporal behavior of a molecular system over time and thereby gives a view of the motion of the atoms. Such motion is inherent to all chemical processes from simple vibrations like bond stretching and angle bending, which give rise to IR spectra, to chemical reactions and ligand-receptor binding²³⁶. MD is based on the following simplified assumptions: (i) the nuclei can be treated as classical particles; (ii) the Born-Oppenheimer approximation holds; (iii) the electronic degrees of freedom can be integrated out.²³⁷

In MD, simulating the evolution of a system of N particles is done by solving Newton's second law of classical mechanics, $F = ma$ (F ; force, m ; mass, a ; acceleration and at same time the second derivative of displacement with respect to time) allowing the calculations of the position of each atom in the system as a function of time. MD simulation is capable of generating successive configurations of a system of molecules and propagating the positions and velocities forward in time. It mimics the way molecules search their conformational space in reality, being particularly

suitable for the study of protein or other large molecule's conformations, which cannot be simulated quantum-mechanically due to their size¹⁷⁵.

The starting point for MD simulation should be the energy minimised structure of the system in study. Although MD can be carried out in vacuum, unlike most molecular mechanics calculation it usually uses explicit solvent approaches. A box filled with water molecules is generated around the molecule to study, and periodic boundary conditions are applied (the box is surrounded by identical copies of itself, so whatever leaves it from one side enters back from the opposite side). Usually, MD simulation involves 3 steps:

- Energy minimisation of the system in study.
- Solvent equilibration.
- Setting simulation parameters (temperature, pressure, control algorithms and time step), computing the total force that acts on each atom, updating the coordinates and the velocities of the atoms by the integration of the Newton's equation of motion.

The simulated time span depends on the computational power used. With the availability of modern supercomputers, it has been possible to extend the length of the simulations up to hundreds of nanoseconds. Nevertheless, there is always a trade-off between size and complexity of the system under study, and the length of the trajectory. Data from MD trajectories can be used to calculate several structural and dynamic properties such as displacement and fluctuation;

- *Displacement*

The conformational stability of a macromolecule can be estimated by the Root Mean Square Displacement (RMSD) of a set of N atoms at time t , with respect to the initial conformation. An increase of the RMSD indicates that the protein moves to a conformation different from the initial

structure, which is usually used to estimate the stability of ligand-target complex^{173, 236,237}.

- Fluctuation

The Root Mean Square Fluctuation (RMSF) is computed for each atom providing information about the atomic fluctuations throughout the simulation and thus indicates the most flexible regions of the protein, which is often relevant to the protein function^{173,236,237}.

3.1.9 Ligand-Based Drug Design (LBDD)

If a 3D structure of the target is unavailable but one or more binding molecules are known, LBDD provides an alternative strategy.

One way is the derivation of a 3D ligand "Pharmacophore" model of the known actives. Pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal interactions with a specific biological target structure. Once established, it can be used to obtain a pseudo-receptor model to design structures that are complementary to the primary target constraints. The generalisation quality of a 3D pharmacophore model improves with the high degree of structural diversity in the training set of known ligands. Also, a common binding mode of all deployed ligands is a prerequisite for the building of a predictive model²³⁸.

A set of known ligands can also be used for the development of a target-specific Quantitative Structure-Activity Relationship (QSAR) model. QSAR is the mathematical relationships linking chemical structure and activity of a series of compounds in a quantitative manner²³⁹⁻²⁴⁴.

Chapter 3

Part 2

HCV Helicase; Mechanistic insights

There are 5 solved crystal structures of HCV helicase enzyme. Their corresponding PDB files are; 1HEI showing the closed conformation ¹⁴³, 8OHM showing the open conformation ¹⁴⁴, 1A1V showing the HCV helicase with a bound DNA oligonucleotide ¹¹⁶, 1CU1 which is a full-length NS3 complex with the central portion of NS4A ¹⁴⁵ and recently 2F55 showing two NS3 helicase fragments bound to the same oligonucleotide ¹⁴⁶. These crystal structures, along with high-resolution NMR structures of HCV helicase domain 2, have greatly influenced proposals explaining how helicases function and have guided experiments designed to test these ideas ^{245,246}.

The NS3 helicase enzyme forms a Y-shaped molecule with three domains, the N-terminal domain (domain 1) and the middle domain (domain 2) are above the C-terminal domain (domain 3). In structure 1A1V short DNA oligonucleotide is bound to the helicase in the cleft that separates domain 3 from domains 1 and 2 (Figure 3.3 B) ²⁴⁷.

The main difference between the available HCV helicase structures concerns the position of domain 2 relative to domains 1 and 3. Domains 1 and 3 share more of an interface than domain 2 shares with either of the other domains. Domain 2 is connected to domains 1 and 3 *via* flexible linkers, which allow domain 2 to freely rotate relative to domains 1 and 3. In some structures, domain 2 is rotated away from domain 1 in an "open" conformation, while in other structures domain 2 is closer to domain 1 in a "closed" conformation (Figure 3.3). The pivot point for these rotations is provided by additional contacts between domain 3 and an extended β -hairpin originating from domain 2. It is believed that switching between

the open and closed states is modulated by ATP binding and hydrolysis²⁴⁷.

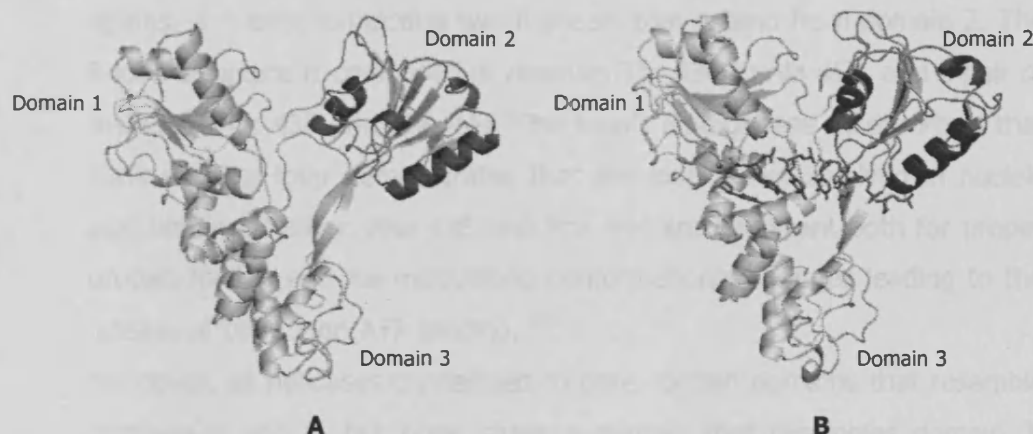


Figure 3.3 Comparison of HCV NS3 Helicase structures in the open (apo) conformation, PDB file 8OHM (A) and the closed conformation with the polynucleotide strand, PDB file 1A1V (B) showing the relative free rotation of domain 2 upon binding to the polynucleotide strand.

In addition to the known conserved seven motifs, which essentially form the motor that convert the chemical energy from ATP hydrolysis into a mechanical force and hence leading to the disruption of DNA or RNA base pairs, there are a number of other motifs which are found only in HCV helicase and closely related viruses, including the Arg-clamp, the Phe loop, and all motifs in domain 3.²⁴⁷

With a hope of discovering regions that might provide binding sites for novel anti-HCV therapeutics, the role of two motifs in domain 2 that are conserved in all HCV isolates but not related proteins, were examined by Frick *et al*¹³⁸.

The rationale was that compounds that bind such sites would be relatively nontoxic because similar sites are not present on related cellular helicases. The first motif identified was centered on Arg 393, a residue that contacts the nucleic acid backbone (Figure 3.4 A). When Arg 393 is changed to Ala (R393A mutation), the protein still catalyses RNA-stimulated ATP hydrolysis but does not unwind DNA or RNA, suggesting that this Arg-clamp motif tether the protein to the nucleic acid strand on which it is translocating²⁴⁷.

The second motif found to be characteristic of only helicases from HCV strains, is a loop connecting two β -sheets that extend from domain 2. The β -loop structure is composed of residues Thr 430 to Ala 452, and a pair of residues, Phe 438 and Phe 444 "Phe-loop". Mutagenesis of the Phe's that flank the Phe-loop demonstrates that the loop is not involved in nucleic acid binding. Rather, Phe 438 and Phe 444 are important both for proper protein folding and for modulating conformational changes leading to the release of DNA upon ATP binding.²⁴⁶

Moreover, all helicases crystallised to date contain domains that resemble domains 1 and 2, but none share a domain that resembles domain 3. Domain 3 is missing entirely; suggesting that domain 3 might not be required for HCV helicase movements. This is not the case, however, and although its role in unwinding is only beginning to be understood, domain 3 is clearly essential. Deletion of 97 amino acids from the C-terminus of NS3 results in an inactive helicase.^{248, 249}

Two key residues in domain 3 are Trp 501, which stacks against a nucleic acid base to act like a bookend²⁵⁰⁻²⁵², and Glu 493, which helps repel nucleic acids from the binding cleft upon ATP binding.²⁵³

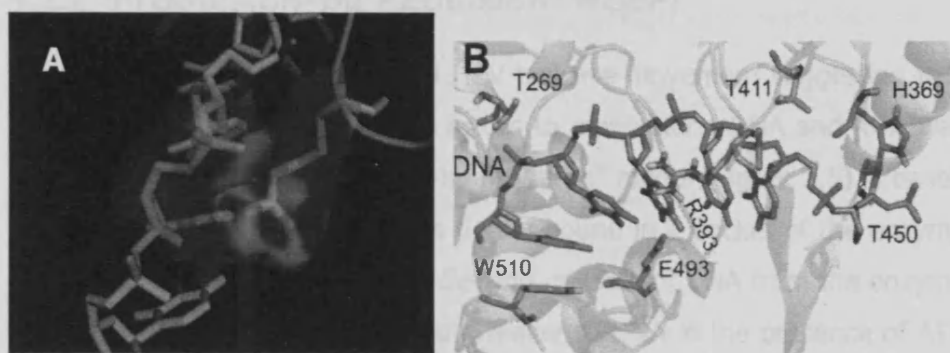


Figure 3.4 A) arginine 393 clamp (blue) in contact with phosphate group (orange). B) Key residues contacting the oligonucleotide bound to HCV helicase in PDB file 1A1V¹¹⁶

There is presently no consensus on exactly how the HCV helicase unwinds RNA. However, unlike SF1 helicases, which have many interactions with nucleic acid bases, most of the contacts of HCV helicase occur with protein side chains and the sugar-phosphate backbone of DNA¹¹⁶. A close-up of the DNA binding site of structure 1A1V with key amino acids is highlighted in (Figure 3.4 B). Two key hydrogen bonds are donated by residues Thr 269 from domain 1 and Thr 411 in domain 2. Mutagenesis of either residue affects both RNA binding affinity and unwinding rates¹³⁸.

Based on the observation that the oligonucleotide appears to be locked into the binding cleft because a residue in domain 3, Trp 501, is stacked against the 3'-terminal base, it was proposed that ATP binding, and the subsequent closure of the cleft between domains 1 and 2, will lead to a ratcheting of Trp 501 past 1 or 2 nucleotides. Consequently, the protein would move towards the 5'-end of the bound nucleic acid. After ATP is hydrolysed and Trp 501 is again locked into place acting as a bookend, the cleft opens and RNA slides through the other side of the protein. Similarly, it was proposed that the residue that acts as a 5'-bookend, analogous to the 3'-bookend Trp 501, might be Val 432 in domain 2²¹⁶.

3.2.1 "Propulsion-by Repulsion" model

More recently, a model to explain HCV helicase movement suggesting that HCV helicase utilises electrostatic forces to move along DNA and RNA was proposed^{253,254}. This "propulsion-by repulsion" model (Figure 3.5) is based on two observations. First, DNA is tightly bound in a pocket of the enzyme that is highly negatively charged. Second, release of DNA from the enzyme is pH dependent; the enzyme binds weaker to DNA in the presence of ATP at a higher pH. The first observation hints that there is a potential energy buildup when the protein is locked onto DNA in the absence of ATP. The second observation suggests that ionisable residues come in contact with DNA upon ATP binding. It was shown using mutagenesis that one of these

key residues is Glu 493 in the ss DNA binding cleft ²⁵³.

In this model, ATP binding leads to a conformational change such that the nucleic acid bases can clear the Trp 501 bookend ²⁵⁴. In the absence of ATP, RNA cannot exit the enzyme because it is blocked by Trp 501 and clamped in the cleft by the Arg clamp on domain 2 ²⁴⁷. When ATP binds, domain 2 rotates bringing with it the positively-charged Arg-clamp. The Arg-clamp attracts the negatively charged phosphodiester backbone so that RNA moves free from the bookend. The negatively charged RNA is then repelled by the negatively charged binding cleft, so it moves through the protein until ATP is hydrolysed, and the protein clamps it tightly again (Figure 3.5).

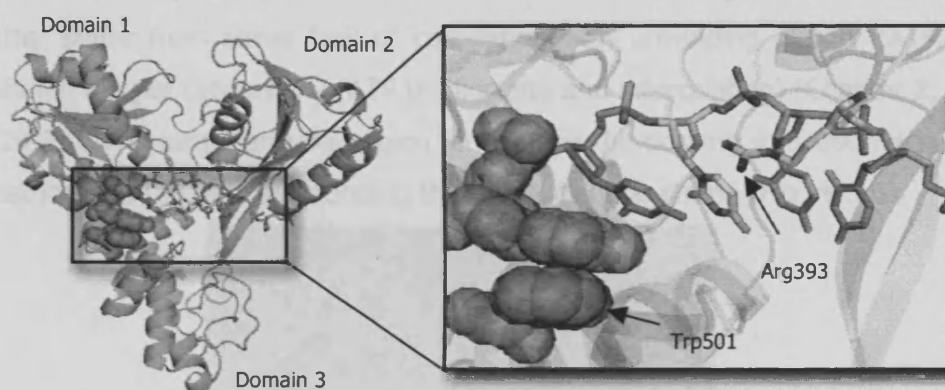


Figure 3.5 The propulsion-by-repulsion model; ATP binding rotates domain 2 so that a positively charged Arg-clamp moves the RNA so that it clears Trp 501, which is holding the RNA in a negatively charged cleft. When ATP is bound, the protein repels RNA past Trp 501 so that the protein moves in a 5' direction until ATP is hydrolysed and the protein returns to its original conformation.

Chapter 3

Part 3

Quinazoline based structures

3.3.1 Introductory remarks

In this work the 3D X-ray crystal structure of NS3 helicase co crystallised with a deoxyuridine octamer; PDB entry 1A1V¹¹⁶ with no water molecules was used as an input file for molecular modelling studies. Also the crystal structure of the PDB file 8OHM¹⁴⁴ showing the open conformation was occasionally used for comparison purposes. The RNA binding site was chosen to design the potential inhibitors over other strategies because the latter suffer from either lack of correlation with unwinding activity (ATP binding site) or cytotoxicity (ATP binding site and intercalation) [Chapter 2, p29-33]. The target enzyme region for running the docking simulation was restricted to the area surrounding the polynucleotide strand (Figure 3.6).

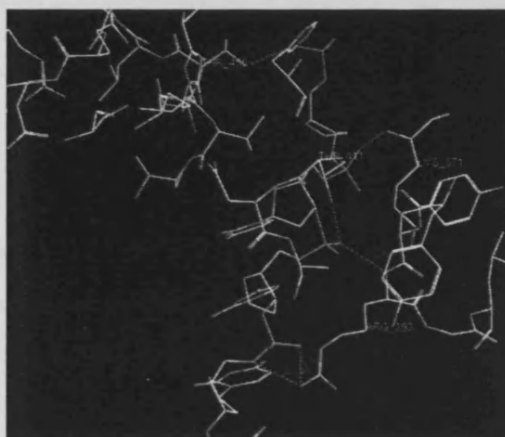
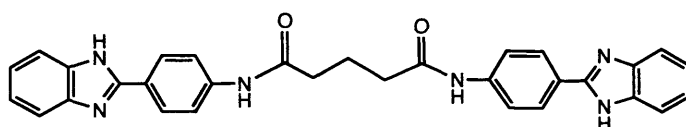


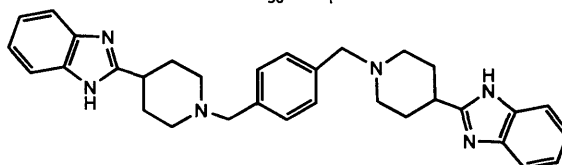
Figure 3.6 the interaction of the backbone phosphate groups of ss DNA with NS3 helicase

In this project, facing no much available data or leads, a SBDD approach was carried out. Starting from the structure of the small number of the reported compounds that target the RNA binding site, a thorough search

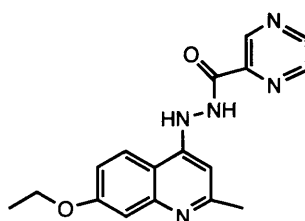
for promising binding regions was done using docking simulation modules within MOE²⁶⁰ and FlexX²⁶¹ software packages (Figure 3.7).



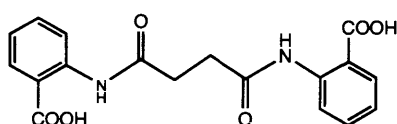
5 IC₅₀ 0.7 μM



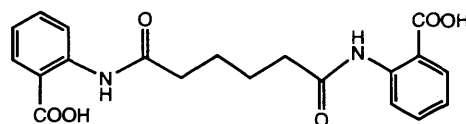
6 (IC₅₀ 7 μM)
Viropharma Inc.¹⁶²



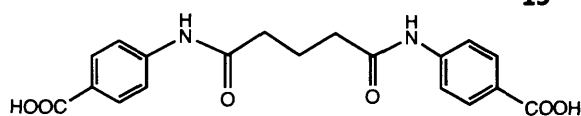
QU663¹³⁷



14

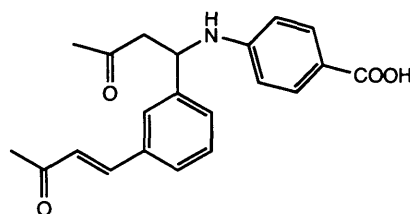


15



16

Chemical structure of Diaryl dicarboxylic acid derivatives (**14-16**) HCV helicase inhibitors²⁵⁵



17

Figure 3.7 Chemical structure of compounds used in searching for binding regions, compound **17** was suggested in a previous study²⁵⁶

Compound **17** was previously²⁵⁶ suggested by the aid of Ligbuilder²¹⁴, a de novo drug design program. **17** was designed to fit in the area centred on Arg 393 clamp, which is known to play an important role in the unwinding activity of helicase enzyme via electrostatically tethering phosphate groups of the polynucleotide (Figure 3.8).

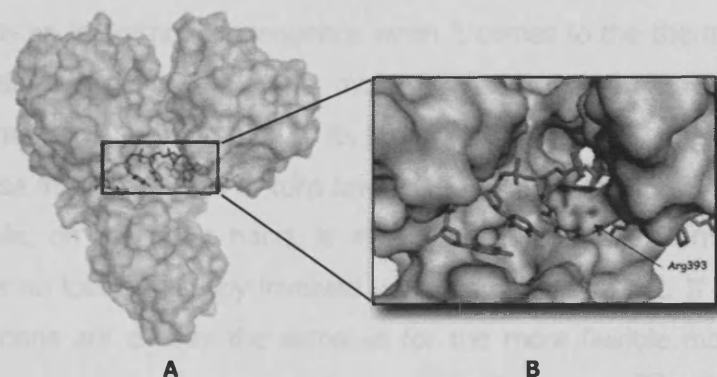


Figure 3.8 A) Crystal structure of HCV NS3 helicase protein complexed with a polynucleotide strand. B) Arg 393 clamp electrostatically holding the negatively charged phosphate group.

3.3.2 Inhibitor Structure Design

Unfortunately, previous attempts to prepare **17** were not successful. Thus, it was proposed to enhance the properties of this compound through simplification and rigidification of its structure. In particular, since it was noticed that while **17** exhibits promising interactions with the RNA binding pocket of the helicase enzyme, this compound has little fidelity to the same site; in other words different conformations binds different regions. This may be because of the relative flexibility of this compound attributed to the number of single rotatable bonds. Bond rotation can lead to a large number of conformations or shapes, which in turn bind to different sites in the apparently wide RNA binding pocket.

• Rigidification Strategy in Drug Design

Generally, rigidification is used to increase the activity of a drug or to reduce its side effects. This approach can work because the more flexible a

structure is, the more likely it will interact with more than one receptor and produce other biological responses (side effects). Usually, this involves locking the molecule into a more rigid conformation so that it cannot take up other shapes or conformations. Thereby, the molecule is ready to fit its target receptor site without the need to adjust to the correct conformation. There is an important consequence when it comes to the thermodynamics of binding as well. A flexible molecule has to adopt a single active conformation in order to bind to its target (more ordered). This results in a decrease in entropy and in turn lowers the binding affinity. A totally rigid molecule, on the other hand, is already in its active conformation, and there is no loss of entropy involved in binding to the target. If the binding interactions are exactly the same as for the more flexible molecule, the rigid molecule will have the better overall binding affinity²⁵⁷.

- **Rigidification of 17**

Since incorporating the skeleton of a flexible compound into a ring is the usual way of locking a conformation, this strategy was applied on structure **17**. Three possible ring systems, **A**, **B** and **C** were suggested to make **17** more rigid as shown in (Figure 3.9). The suggested ring systems represented a core scaffold liable to further functionalisation excluding the Michael acceptor and keeping the major features; the benzoate group and the Hydrogen-bond forming atoms. In a preliminary docking simulation study, a particularly favourable binding mode was exhibited by the bicyclic ring system **A** as evaluated by both score and visual inspection, in addition to a remarkably improved selectivity to the same site of the RNA binding pocket. Furthermore, the modification of the chiral chromene ring in structure **A** into an aromatic bicycle ring (naphthalene) proved to give a better alignment of the carboxylate anchor moiety with the phosphate group of the polynucleotide so that it will be more likely to form a salt bridge with Arg 393 clamp in the active site. Therefore a scaffold

composed of an aromatic bicyclic ring attached to a benzoate group was proposed to provide better geometric fit into the binding pocket.

In this context a number of plausible focused databases of compounds were suggested to identify the most promising ring system with an optimised arrangement of substituents. Docking/scoring simulation experiments were used to evaluate the binding affinities of some of the aromatic bicyclic ring analogues of structure **A** such as naphthalene, quinoline and quinazoline.

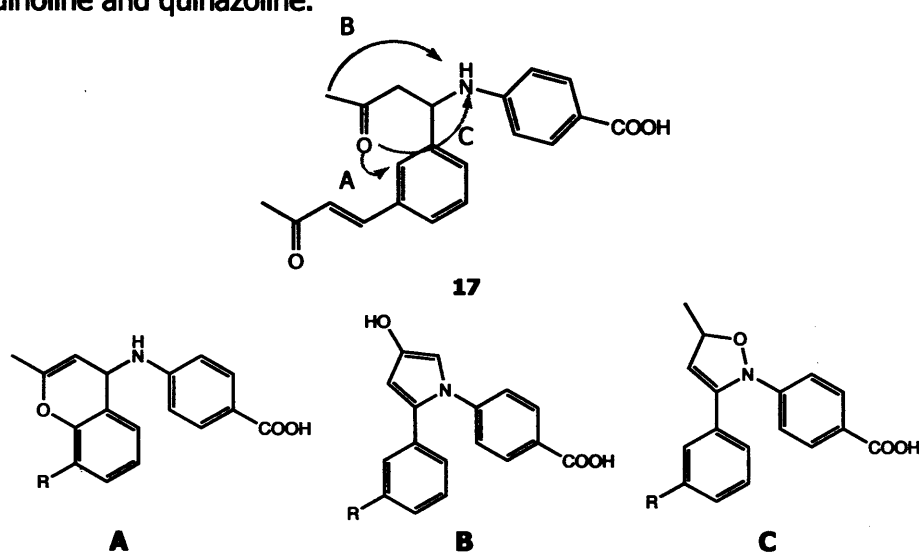


Figure 3.9 Rigidification of **17** lead to 3 main ring systems; **A**, **B** and **C**

The new ligand structures were built taking in consideration the charge adjustment of the carboxylate group then energy minimised, using the MMFF94x force field with 1000 iterations then saved as a small focused database. The active site was defined automatically as the area surrounding the co crystallised polynucleotide strand. Flexible ligand docking using SYBYL (FlexX)²⁵⁸ was carried out to search the conformation space and predict the orientations of each of the designed inhibitors within the previously defined pocket; this was followed by energy ranking of the different possible solutions obtained. The output database, which included

the top scoring 30 poses for each single ligand was visualised using MOE compute interactions option.²⁵⁹ The ranking of these conformations was done by the scoring function, which includes van der Waals and electrostatic energies among other parameters. For the visual inspection purposes, the Connolly surface of the enzyme pocket, which shows the solvent accessible area in the protein structure, was generated. (Figure 3.11).

• Naphthalene based scaffold

The chromene ring in structure **A** was replaced by a naphthalene nucleus with one substituent at different positions to construct a small focused library of naphthalene based structures (Figure 3.10) this database underwent docking simulation study, the results of which guided the following structural modification steps.

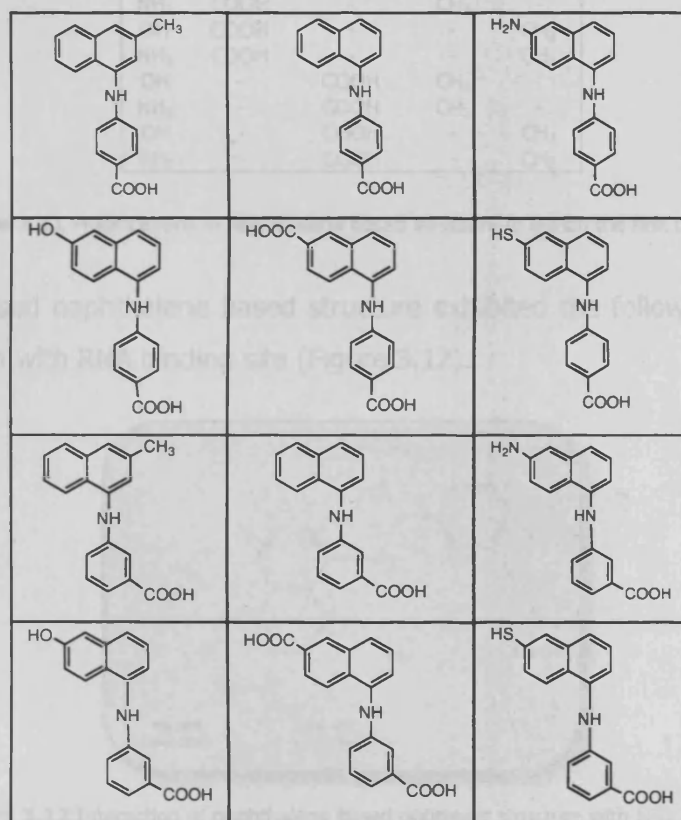


Figure 3.10 Naphthalene based focused library

Based on the results of the docking simulation of the above database and in order to optimise the position of the functional groups, the database was enriched with structures that have extended distance between the benzoate group and the bicyclic ring with one hydrogen-forming substituent at position 6 by studying the effect of inserting a methylene group between the two ring systems either in x or y position (Figure 3.11).

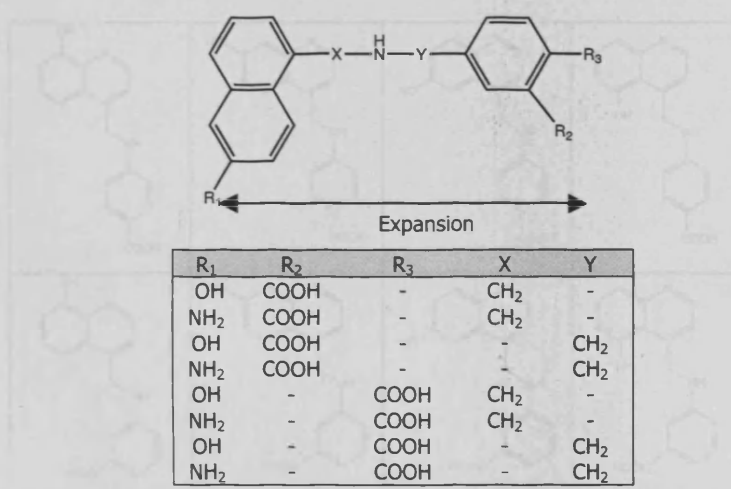


Figure 3.11 Modifications in naphthalene based structure to enrich the first database

An optimised naphthalene based structure exhibited the following binding interaction with RNA binding site (Figure 3.12).

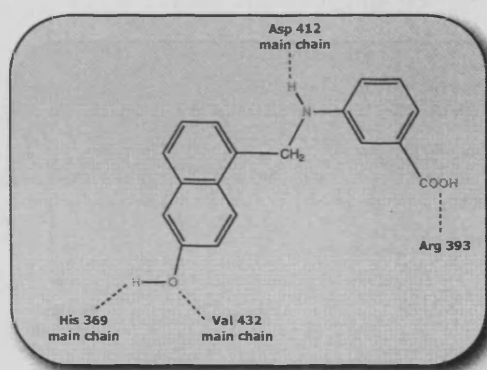


Figure 3.12 Interaction of naphthalene based optimised structure with NS3 helicase

- Quinoline based scaffold

In an analogous docking experiment, a quinoline based focused database was constructed and the binding interactions with the active site were studied (Figure 3.13).

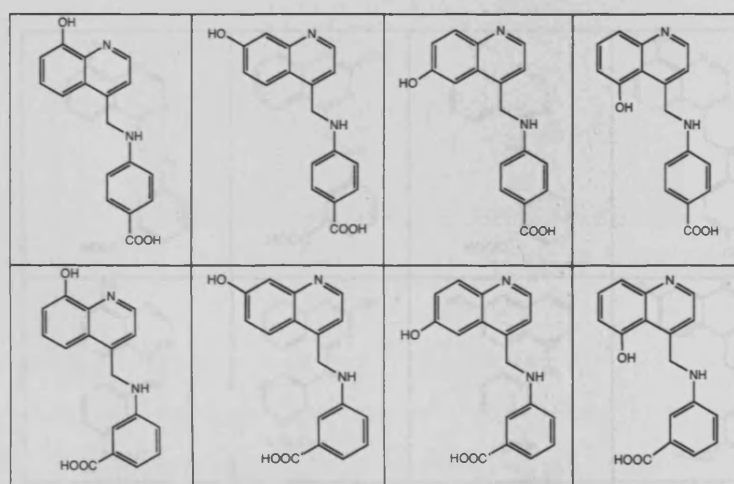


Figure 3.13 Quinoline based focused library

An optimised quinoline based structure had the following binding pattern with helicase enzyme as shown in (Figure 3.14).

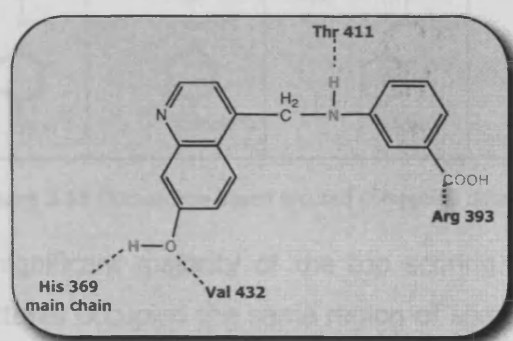


Figure 3.14 Interaction of quinoline based optimised structure with NS3 helicase

• Quinazoline based scaffold

In our search for more synthetically feasible compounds a quinazoline based database (Figure 3.15) was built and docked then enriched by more structurally modified compounds guided by the score and the visual inspection of the docking results of each docking run.

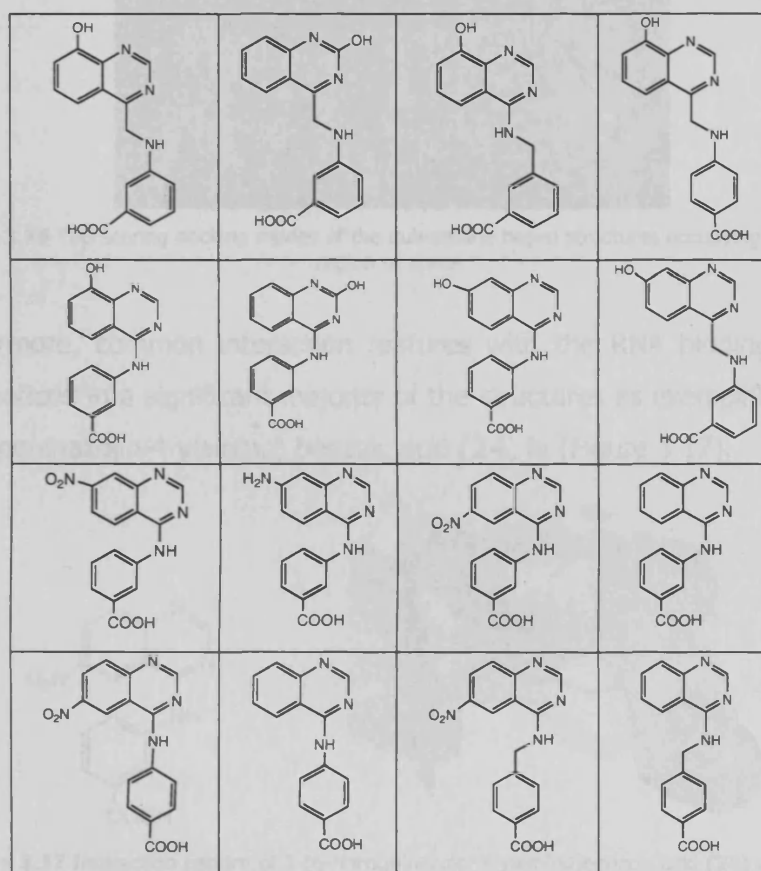


Figure 3.15 Quinazoline based focused compound database

Interestingly, a significant majority of the top scoring binding modes of most of the structures occupied the same region of space and in particular having some binding interactions with those crucial residues involved in tethering the phosphate backbone of the polynucleotide strand to the

helicase enzyme namely; Thr 411, Val 432 besides the Arg 393 clamp (Figure 3.16).

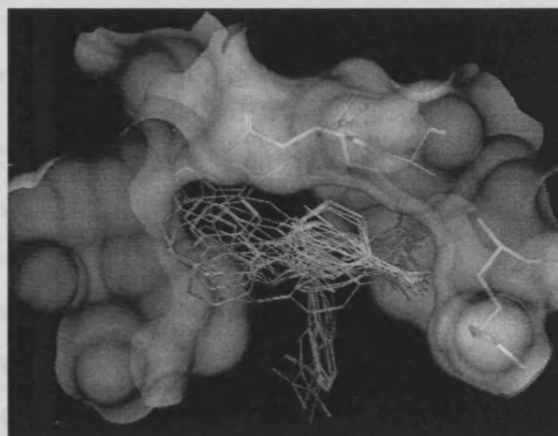


Figure 3.16 Top scoring docking modes of the quinazoline based structures occupying the same region of space.

Furthermore, common interaction features with the RNA binding pocket were noticed in a significant majority of the structures as exemplified by 3-(6-Nitroquinazolin-4-ylamino) benzoic acid (**24**) in (Figure 3.17).

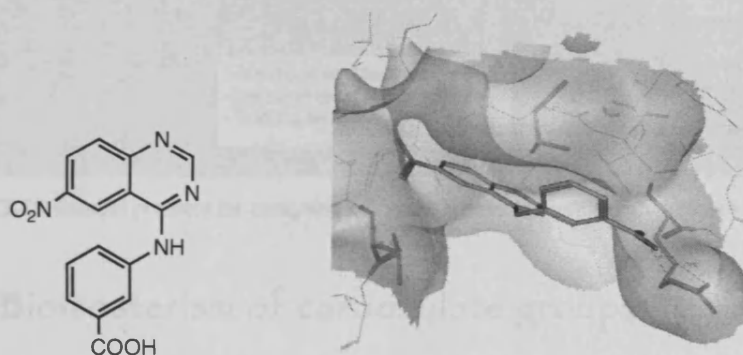


Figure 3.17 Interaction pattern of 3-(6-Nitroquinazolin-4-ylamino)benzoic acid (**24**) with NS3 helicase

Among these interactions is the negatively charged carboxylate group, which interacts with the positively charged guanidine group of Arg 393 residue by forming an ion pair at distances ranging from 2.5 to 3.1 Å. The N1 nitrogen of the quinazoline ring is involved in hydrogen bond interaction with the backbone NH of Val 432 at distances less than 3 Å. On

the other hand N3 of the heterocycle is interacting with the backbone NH of Asp acid 412. The central NH group forms hydrogen bond with Thr 411 residue. In addition to the hydrophobic interactions with Thr 411, Ala 413, Ser 370, Lys 371, Tyr 392, Val 432, Thr 448, Asp 412.

In summary, it was quite reasonable to prepare a series of quinazoline-based compounds following the approach outlined in Figure 3.18.

And even more encouraged by the fact that some quinazoline based compounds were reported to have antihepatitis C virus activity targeting either one of the viral non-structural proteins²⁶⁰ or one of the intracellular machinery components involved in HCV replication.²⁶¹

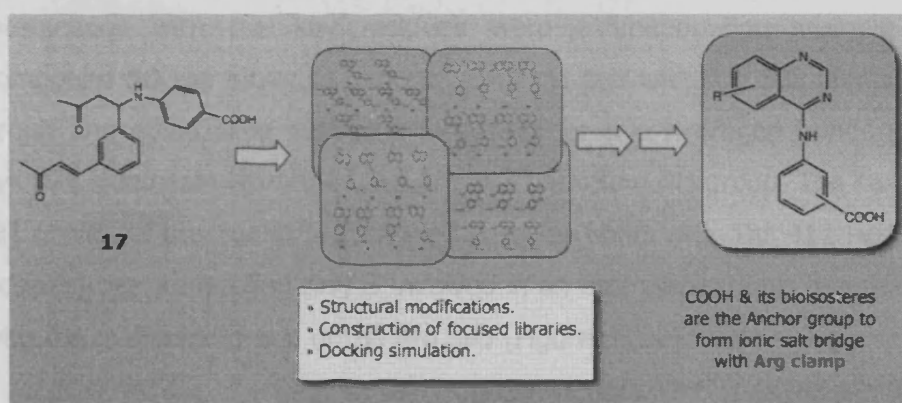


Figure 3.18 General protocol for designing quinazoline based potential HCV helicase inhibitors

3.3.3 Bioisosterism of carboxylate group

The carboxylate group in the above compounds is considered to be the major anchor moiety for its proposed interaction with the Arg 393 amino acid in the RNA binding site of the helicase enzyme. The tetrazole ring is known to be a bioisostere of the carboxylate function group principally in terms of its physicochemical properties related to acidity.²⁶² It was anticipated to prepare the tetrazole analogues of the carboxylic group

derivatives because the former is more stable and more lipophilic²⁶³ hence better cell membrane permeability could be anticipated (Figure 3.19).

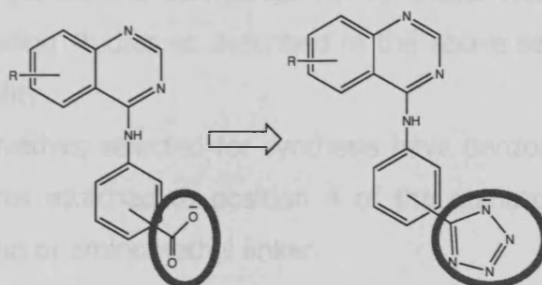


Figure 3.19 Bioisosteric replacement of COOH group by tetrazole ring

This assumption was even more supported by the docking results of the tetrazole analogues, where good alignment geometry and binding interactions with the key residues were exhibited. For example in compound **50** the tetrazole ring can form ion pair with Arg 393 guanidine group and quinazoline nitrogens form another two hydrogen bonds with Lys 371 side chain NH group and Arg 393 backbone NH group. The central NH group of the ligand is forming hydrogen bond with Thr 411 residue besides; the quinazoline ring is involved in an aromatic-cationic interaction with the guanidine group of the Arg 393 (Figures 3.20).

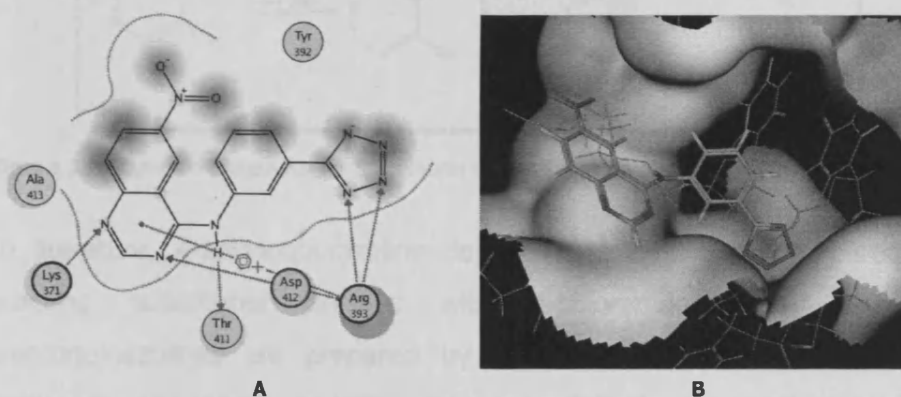


Figure 3.20 A) 2D diagram showing the predicted binding interactions of tetrazole analogue **50** with helicase. Hydrophobic residues (green), polar residues (pink); [basic (blue ring), acidic (red ring)], H-bonds [(arrow); residue side chain (green), residue backbone (blue)] annotated proximity contour and solvent surface area exposure properties are drawn in the background of the diagram. B) 3D depiction, polynucleotide (green)

3.3.4 Synthesis of Quinazoline based series

A selection of Quinazoline derivatives for synthesis was based on both molecular modelling studies as described in the above sections as well as synthetic feasibility.

Mainly, the derivatives selected for synthesis have benzoate group or one of its bioisosteres attached to position 4 of the quinazoline ring system through an amino or aminomethyl linker.

Synthetic routes of the selected compounds were designed based on established procedures.

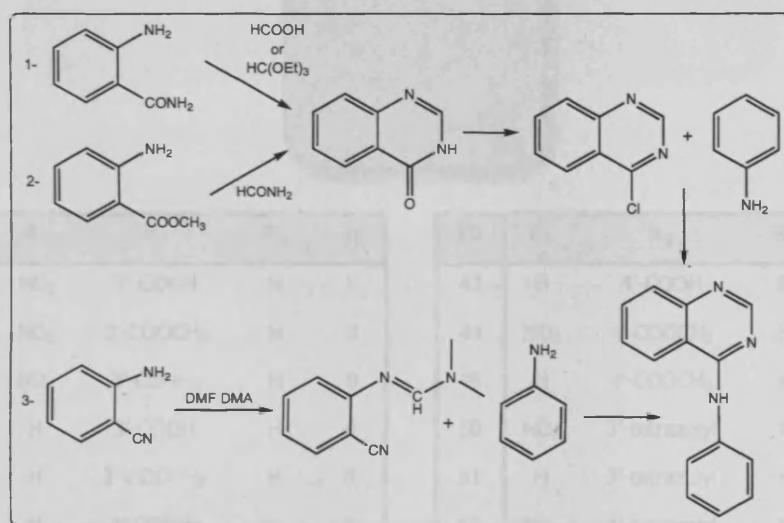
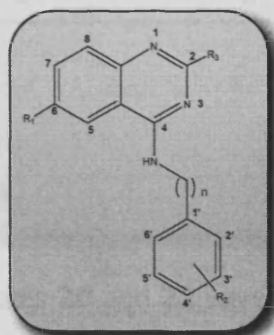


Figure 3.21 common procedures for preparation of 4-anilinoquinazolines (1 and 2) vs Tosu one (3)

In literature, 4-anilinoquinazoline derivatives are generally prepared by reacting substituted anilines with 4-chloroquinazolines. The 4-chloroquinazolines are prepared by chlorination of the corresponding quinazolinones, which, in turn, are prepared by heating anthranilamides with formic acid²⁶⁴ or triethyl orthoformate²⁶⁵ or by heating anthranilates with formamide²⁶⁶. In this work the more efficient synthesis pathway

reported by *Tsou et al*²⁶⁷ was adopted in which, ring cyclisation and incorporation of the 4-anilino group occurs in a single step (Figure 3.21).

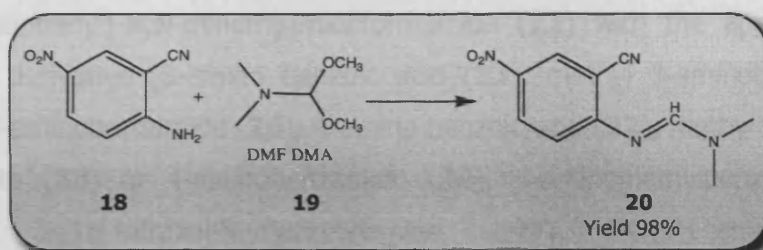
The ester, amide and tetrazole derivatives were prepared together with their free acidic counterparts in order to mask the high polarity of the latter and hence attain better cell membrane permeability. Also, it was set out to synthesise more analogues to probe the vicinal chemical space of this scaffold (Figure 3.22).



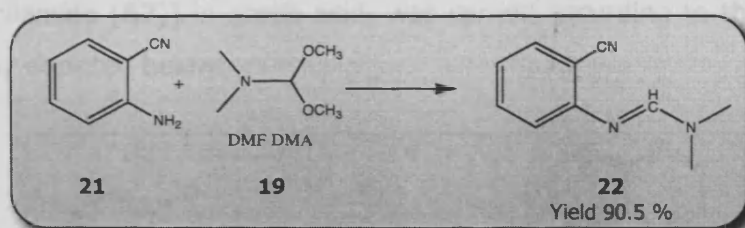
ID	R ₁	R ₂	R ₃	n	ID	R ₁	R ₂	R ₃	n
24	NO ₂	3'-COOH	H	0	43	H	4'-COOH	H	1
25	NO ₂	3'-COOCH ₃	H	0	44	NO ₂	4'-COOCH ₃	H	1
26	NO ₂	3'-CONH ₂	H	0	45	H	4'-COOCH ₃	H	1
29	H	3'-COOH	H	0	50	NO ₂	3'-tetrazolyl	H	0
30	H	3'-COOCH ₃	H	0	51	H	3'-tetrazolyl	H	0
31	H	3'-CONH ₂	H	0	52	NO ₂	4'-tetrazolyl	H	0
35	NO ₂	4'-COOH	H	0	53	H	4'-tetrazolyl	H	0
36	NO ₂	4'-COOCH ₃	H	0	55	NO ₂	4'-CN	H	0
37	NO ₂	4'-CONH ₂	H	0	58	NO ₂	2'-COOH	H	0
38	H	4'-COOH	H	0	59	NO ₂	2'-CONH ₂	H	0
39	H	4'-COOCH ₃	H	0	60	H	2'-COOH	H	0
40	H	4'-CONH ₂	H	0	61	H	2'-CONH ₂	H	0
42	NO ₂	4'-COOH	H	1	65	H	4'-COOCH ₃	CH ₃	0

Figure 3.22 structure of the synthesised 4-anilinoquinazoline derivatives

N'-(2-Cyano-4-nitrophenyl)-N,N-dimethylimidoformamide



N'-(2-Cyanophenyl)-N,N-dimethylimidoformamide



The formamidine intermediates **20** and **22** were prepared by condensation of 5-nitroanthranilonitrile (**18**) or anthranilonitrile (**21**) respectively, and dimethylformamide dimethyl acetal (DMF DMA) (**23**) following a reported procedure by Tsou²⁶⁷. Compounds **20** and **22** were obtained in a very good yield. It may be worth mentioning that **22** was previously prepared from different precursors namely; 2-amino-benzamide and DMF/ POCl_3 ²⁶⁸. A proposed mechanism for this kind of reaction is depicted in Figure 3.23.

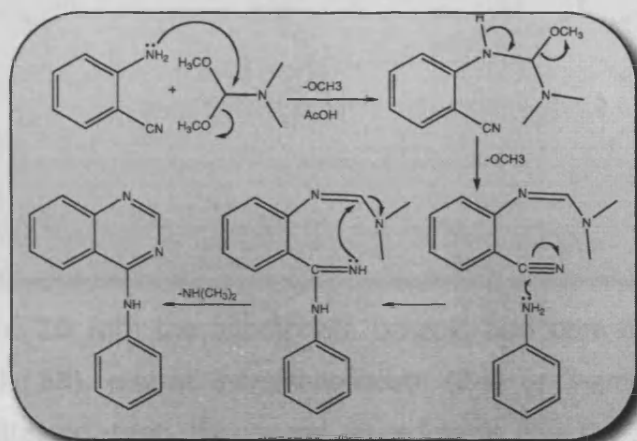
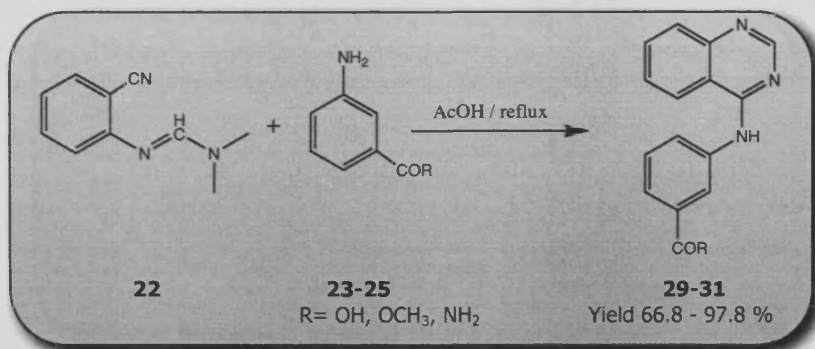


Figure 3.23 General mechanism for the preparation of 4-anilinoquinazolines via formamidine intermediates

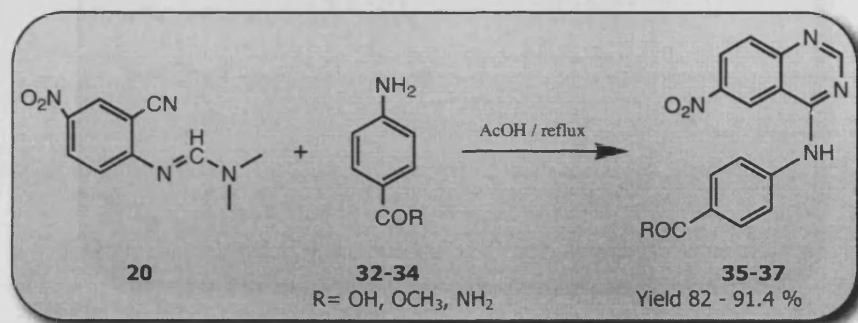
anilinoquinazoline derivatives (**26-28**). The proposed reaction mechanism is illustrated in Figure 3.23.

3-(Quinazolin-4-ylamino)benzoic acid derivatives



The 3'-substituted quinazolines (**29-31**) were obtained by following the general pathway using the formamidine intermediate **22** with the appropriate benzoic acid derivative [3-amino benzoic acid (**23**), methyl 3-aminobenzoate (**24**) or 3-aminobenzamide (**25**)]. 3-(quinazolin-4-ylamino)benzoic acid (**29**) was prepared previously following a different pathway that involved cyclisation of quinazolinone ring, chlorination and coupling with 3-aminobenzoic acid.²⁶⁹

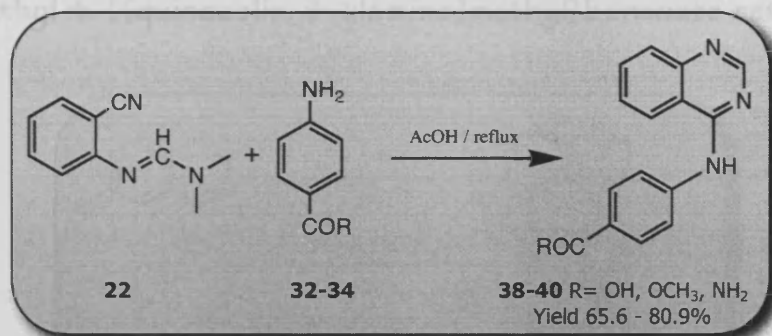
4-(6-Nitroquinazolin-4-ylamino)benzoic acid derivatives



The 4'-substituted benzoic acid derivatives (**35-37**) were prepared from **20** and the appropriate benzoic acid derivative [4-amino benzoic acid (**32**), methyl 4-aminobenzoate (**33**) or 4-aminobenzamide (**34**)] following

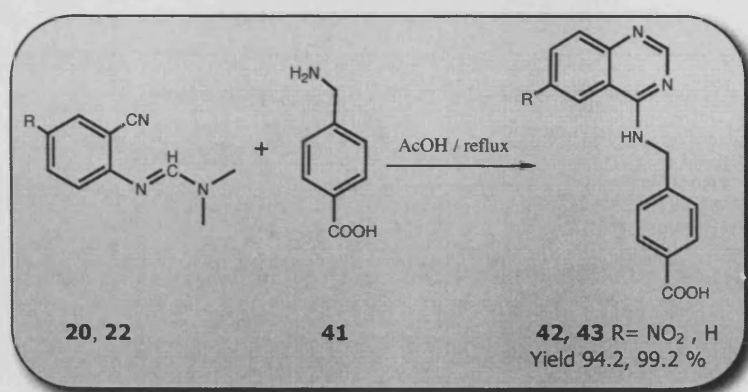
the general pathway. **35** and **37** were also reported previously through different pathway.^{270,271}

4-(Quinazolin-4-ylamino)benzoic acid derivatives



Similarly, a mixture of **22** and the appropriate benzoic acid derivative [4-amino benzoic acid (**32**), methyl 4-aminobenzoate (**33**) or 4-aminobenzamide (**34**)] to afford (**38 -40**). Compound **38** was reported before using a different pathway.²⁶⁹

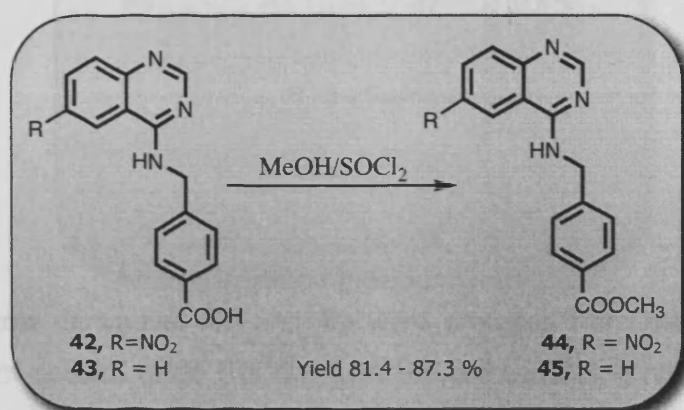
4-((6-Nitroquinazolin-4-ylamino)methyl)benzoic acid derivatives



Compounds **42** and **43** were prepared from a mixture of either N'-(2-cyano-4-nitrophenyl)-N,N-dimethylimidomethanimine (**20**) or N'-(2-cyano phenyl)-N,N-dimethylimidomethanimine (**22**) with 4-aminomethyl benzoic acid (**41**). Both products were obtained in a very good yield. **43** was

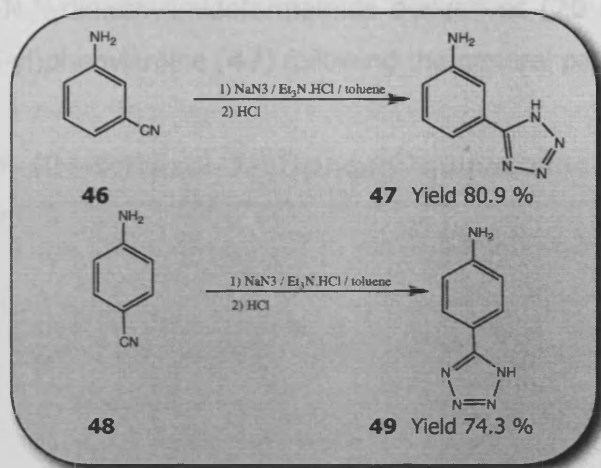
reported before via ring cyclisation into quinazolinone, chlorination and substitution with amino benzoic acid.²⁷²

Methyl 4-((quinazolin-4-ylamino)methyl)benzoate esters



The esters of **42** and **43** were prepared using dry methanol and thionyl chloride mixture, which produces methyl chloride, hydrochloric and sulphuric acids *in situ*.

3 and 4-(1H-tetrazol-5-yl)benzenamines



The tetrazole derivatives **47** and **49** were prepared from the respective reactions of sodium azide (Na N₃) with the corresponding nitriles (**46** or **48**) in the presence of triethylamine hydrochloride ²⁷³. A plausible mechanism for this reaction is the formation of an intermediate complex between triethylamine hydrochloride and sodium azide [Et₃N.HN₃], which in turn reacts with the triple bond of the nitrile group ²⁷³ (Figure 3.24).

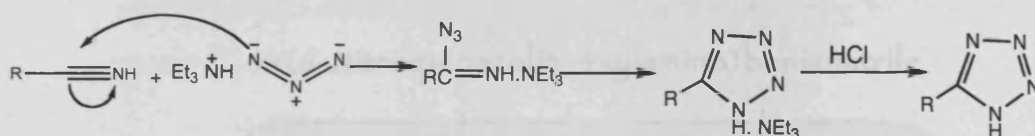
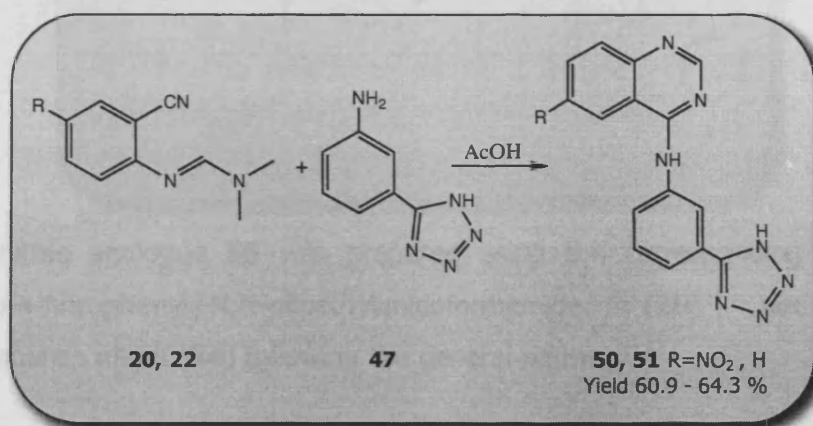


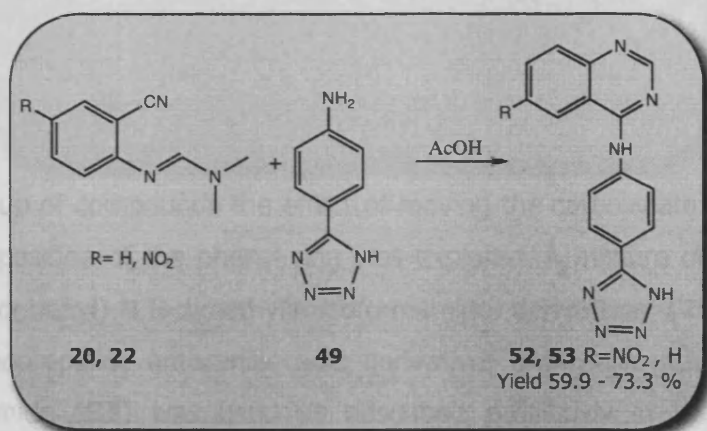
Figure 3.24 Mechanism of tetrazole ring formation

4-Amino(3-(1H-tetrazol-5-yl)phenyl)quinazoline derivatives



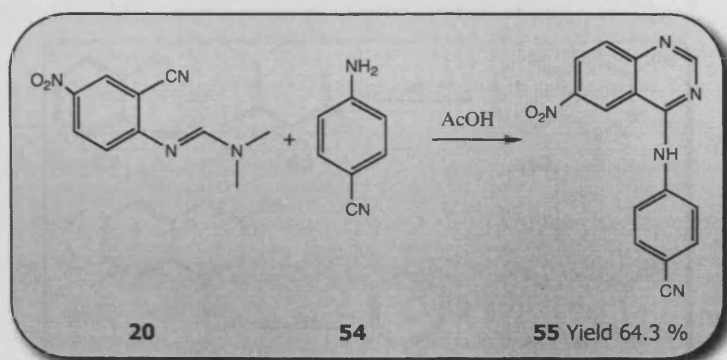
The tetrazole analogues **50** and **51** were prepared using either of *N'*-(2-cyanophenyl)-*N,N*-dimethylimidoforamide derivatives (**20** or **22**) and 3-(1*H*-tetrazol-5-yl)phenylamine (**47**) following the general pathway.

4-Amino(4-(1*H*-tetrazol-5-yl)phenyl)quinazoline derivatives



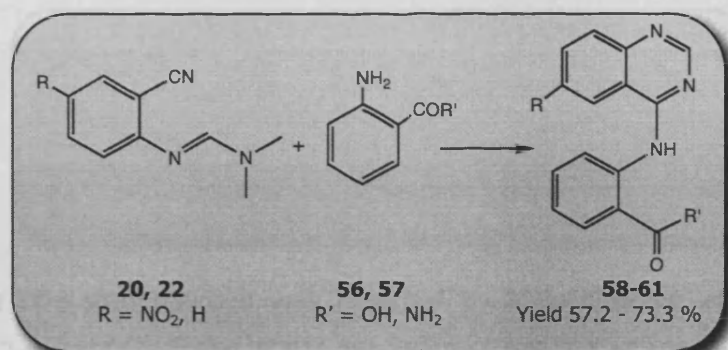
Similarly, tetrazole analogues **52** and **53** were prepared using the corresponding *N'*-(2-cyanophenyl)-*N,N*-dimethylimido formamide derivative (**20** or **22**) and 4-(1*H*-tetrazol-5-yl)phenylamine (**49**).

4-(6-Nitroquinazolin-4-ylamino)benzonitrile



The nitrile analogue **55** was prepared using the corresponding *N'*-(2-cyano-4-nitrophenyl)-*N,N*-dimethylimidoforamide (**20**) and 4-aminobenzonitrile (**54**) following the general pathway.

2-(Quinazolin-4-ylamino)benzoic acid derivatives



In this group of compounds the effect of moving the carboxylate moiety to the ortho position of the phenyl ring was explored. A mixture of either of *N'*-(2-cyanophenyl)-*N,N*-dimethylimidoforamides (**20** or **22**) and the appropriate anthranilic acid derivative; anthranilic acid (**56**) or anthranilamide (**57**) was used as described previously in the general pathway. Compounds **60**²⁷⁴ and **61**²⁷⁵ were reported previously.

Moreover, it was interesting to probe the activity of the 2-methylquinazoline derivative (**65**), which was prepared according to the following scheme (Figure 3.25).

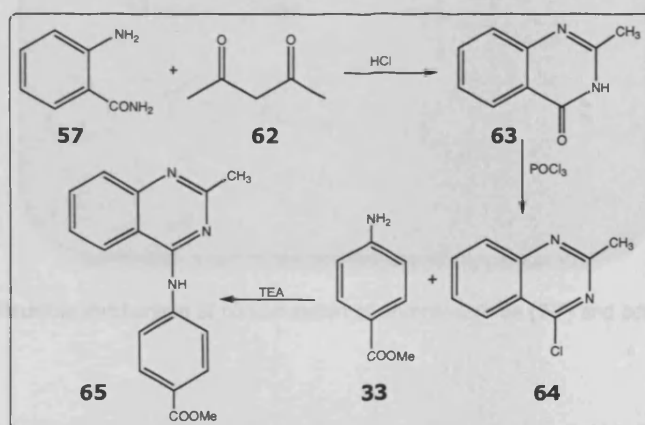
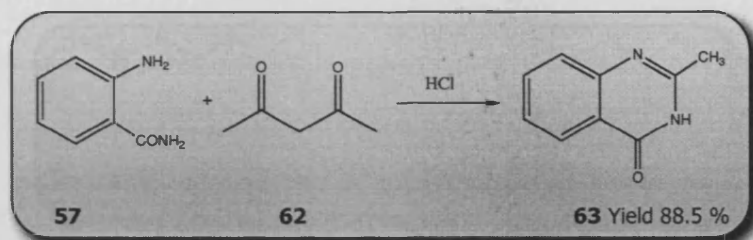


Figure 3.25 Synthesis scheme of Methyl 4-(2-methylquinazolin-4-ylamino)benzoate

2-Methylquinazolin-4(3H)-one



A simple one step reaction was reported by Maloshitskaya *et al*²⁷⁶ to prepare 2-methylquinazolin-4(3H)-one (**63**) as a result of condensation of anthranilamide (**57**) and acetylacetone (**62**) in the presence of catalytic amount of concentrated hydrochloric acid due to the loss of an acetone molecule, however no spectral data were reported. The same compound **63** was prepared by a different pathway, the spectral data of which was used to compare our results.²⁷⁷ A plausible mechanism of this condensation reaction is shown in Figure 3.26.

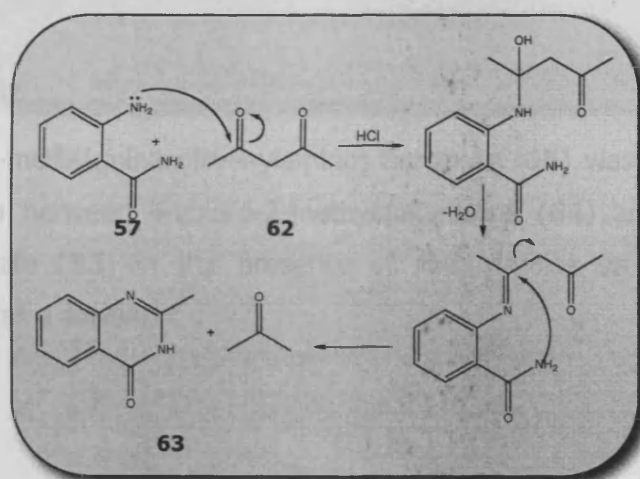
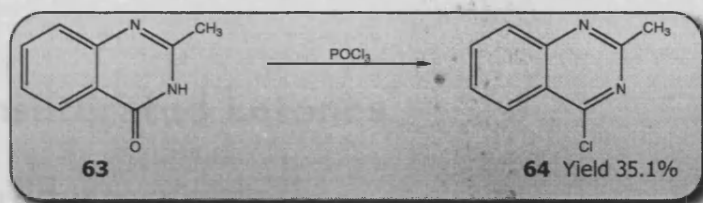


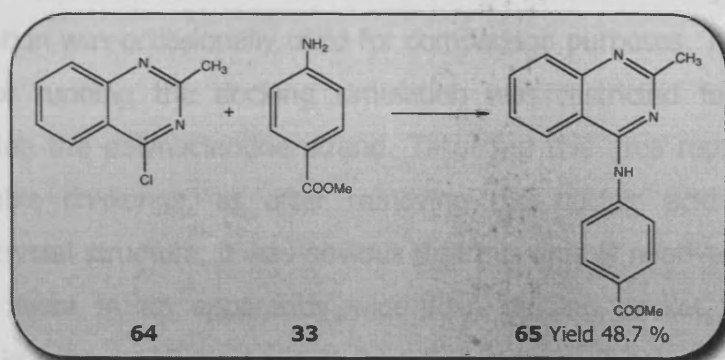
Figure 3.26 Plausible mechanism of condensation of anthranilamide (**57**) and acetylacetone (**62**)

4-Chloro-2-methylquinazoline



The corresponding 4-Chloro-2-methylquinazoline (**64**) was prepared by the reaction of **63** with phosphorus oxychloride (POCl_3)²⁷⁸.

Methyl 4-(2-methylquinazolin-4-ylamino)benzoate



Methyl 4-(2-methylquinazolin-4-ylamino) benzoate (**65**) was prepared by the reaction between 4-chloro-2-methylquinazoline (**64**) and methyl 4-aminobenzoate (**33**) in the presence of triethylamine as a base and isopropanol as a solvent.

Chapter 3

Part 4

α , β Unsaturated ketones

3.4.1 Introductory remarks

In this part the 3D X-ray crystal structure of NS3 helicase enzyme co-crystallized with a deoxyuridine octamer (PDB entry 1A1V)¹¹⁶ with no water molecules was used as an input file for molecular modelling studies. Also the crystal structure of the PDB file 8OHM¹⁴⁴ showing the open conformation was occasionally used for comparison purposes. The enzyme region for running the docking simulation was restricted to the area surrounding the polynucleotide strand. Targeting this area represented a considerable challenge, as after removing the nucleic acid from the enzyme crystal structure, it was obvious that this area is relatively exposed to the solvent in an apparently wide RNA binding pocket, making it implausible for a small molecule to bind tightly without being replaced by water or the nucleic acid itself.

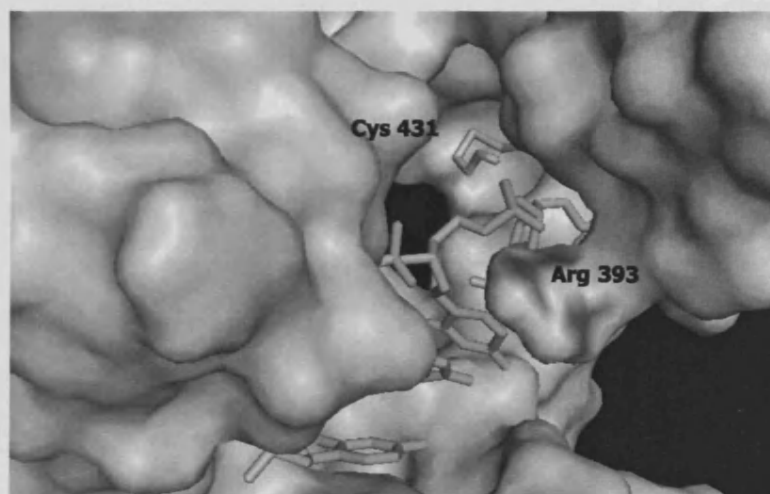


Figure 3.27 NS3 HCV helicase (1A1V PDB file) RNA binding site showing Cys 431 crystal artefact (adduct of β -mercaptoethanol), Arg 393 clamp (blue) and nucleic acid strand (green)

It is not uncommon to see adventitious binding of components of the crystallisation buffer to certain amino acid residues in a protein. Typically these observations are ignored, but these 'crystallisation artefacts' recorded in the Protein Data Bank (PDB) files may provide a good source of valuable information on serendipitous binding sites²⁷⁹.

An example of these crystallisation artefacts was noticed in the 1A1V PDB file where Cys 431 residue was implicated in the formation of a disulfide bond with a molecule of β -mercaptoethanol; a component of the crystallisation buffer (Figure 3.27). This information gave us a clue of both the exposure and reactivity nature of such residue. The location of Cys 431 residue is not far from the substantial Arg 393 clamp residue (15.3 Å in the open 8OHM conformation and 16.4 Å in the closed 1A1V one). As mentioned earlier the Arg 393 clamp plays a key role in the unwinding activity as proved by the detrimental effect of R393A mutation as well as being conserved in all HCV isolates²⁴⁷.

From the crystal structure, it is perceived that Cys 341 residue is not directly involved in the binding of the nucleic acid strand since the latter was not inhibited in spite of its blockage by β -mercaptoethanol molecule. However, its proximity from the RNA binding site as well as its reactivity made it a good candidate for designing a molecule that could react covalently with its thiol group. In fact these observations led to the hypothesis that a molecule that would interact with Arg 393 and at the same time form a covalent bond with Cys 431 through a chemically reactive group may have the potential to inhibit the helicase enzyme activity. In this context a molecule having an α, β unsaturated ketone (Michael's acceptor) group, which also would form hydrogen bond with the nearby Arg 481 and a carboxylate group to form an ionic salt bridge with Arg 393 clamp was suggested. By this geometric arrangement we would expect a molecule capable of forming a bridge between Arg 481, Cys 431

and most important with Arg 393, thereby blocking the way of viral RNA passage through the helicase enzyme. Such a covalent binder would not be easy to be displaced by water or the incoming RNA (Figure 3.28).

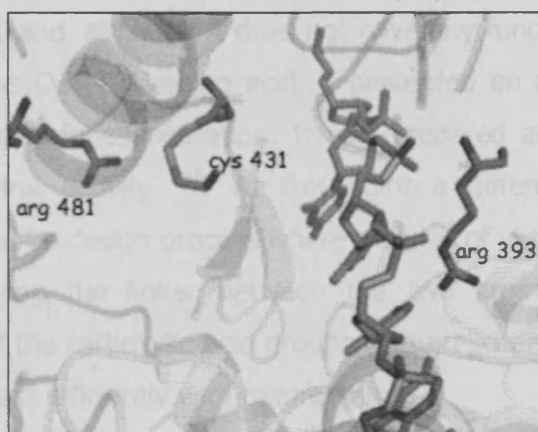


Figure 3.28 The arrangement of the three key amino acids relative to the nucleic acid (orange)

3.4.2 Previous work overview

Initially, the *de novo* software package LigBuilder²¹⁴ was used.²⁵⁶ Programs of this type normally require the user to define an initial “seed” in a binding site, after which the computer builds a series of molecules by adding to the growing structures the most suitable fragments taken from a given library in the software²²². The major drawback of this approach is that very often the proposed structures are highly complex and hence not synthetically accessible. Indeed this was the problem faced upon examining the results obtained. It was evident that the software attempted to fill the space in the large pocket generating a series of structures with relative complex chirality (Figure 3.29, Structure **i**). To overcome this problem, the fragment library was reduced to include only residues that, when combined, were less likely to generate stereocentres. Also, a virtual wall of dummy atoms (atoms with no physical property) was created to reduce the site size and to force the software to design a molecule that

would interact with the desired residues. The compounds thus generated were considerably simpler. Among them structure **ii** was able to interact with Arg 393 and also with another arginine located on the opposite side of the putative binding site, namely Arg 481. Furthermore **ii** was located close to Cys 431 and, although it does not have any functional group able to react with the Cys 431 amino acid, it presented an appropriate basic scaffold for further functionalisation. It was prepared and did not show promising antiviral activity²⁵⁶. At this point a different approach to optimise the *in silico* design process where a series of virtual libraries were generated, varying the linker between the two aromatic rings whilst replacing one of the carboxylic acid groups with a Michael acceptor, which are known to react efficiently with natural thiols.

These libraries were then docked into the binding site then scored according to their ability to interact with the two arginines (481 and 393) and to place the vinyl ketone in close proximity to Cys 431 thiol. The most promising structure (Figure 3.29, **17**) presented the reactive centre of the Michael's acceptor close to the sulphur atom of the Cys 431 with a side chain that could increase the interaction with the enzyme. A molecular dynamic simulation was performed on the ligand/protein complex to show the relative stability of the binding interactions of this compound.^{256, 281}

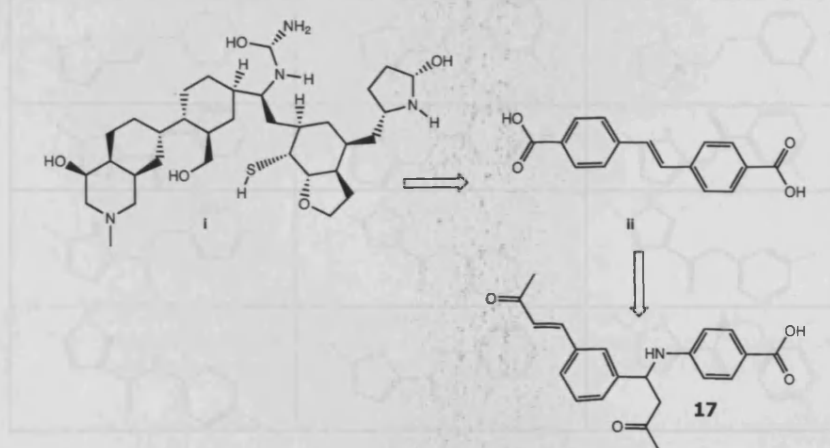


Figure 3.29. Evolution in the design of a novel HCV helicase inhibitor

3.4.3 Inhibitor Structure Design

As previously mentioned, compound **17** was designed with the help of ligbuilder²⁵⁸. Nevertheless, the stereocentre and relative structure complexity presented a synthetic accessibility challenge. For that a focused library of compounds was built to modify the core region of **17** while keeping the carboxylate and Michael acceptor features, which are believed to interact with the three key amino acids in the following pattern; Arg 393 (salt bridge with the carboxylate group), Cys 431 (electrophilic conjugate addition of the α, β unsaturated ketone) and Arg 481 (hydrogen bond with the carbonyl group of the Michael acceptor) (Figure 3.30).

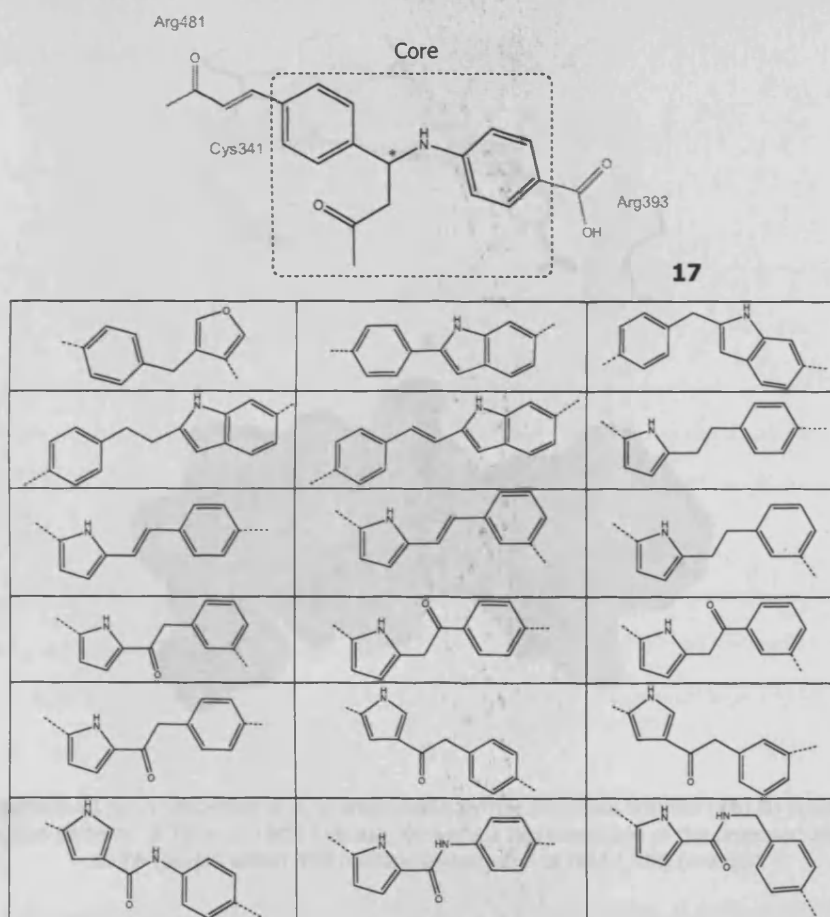


Figure 3.30 Replacing the core of **17** with different ring and linker combinations

Adopting a guided cycle of focused library design and docking simulation runs, structure **76** (Figure 3.31) emerged to have promising interactions in terms of orientation and shape complementarity within the helicase enzyme as examined by both scoring and visual inspection of FlexX²⁶¹ docking output database results²⁸¹.

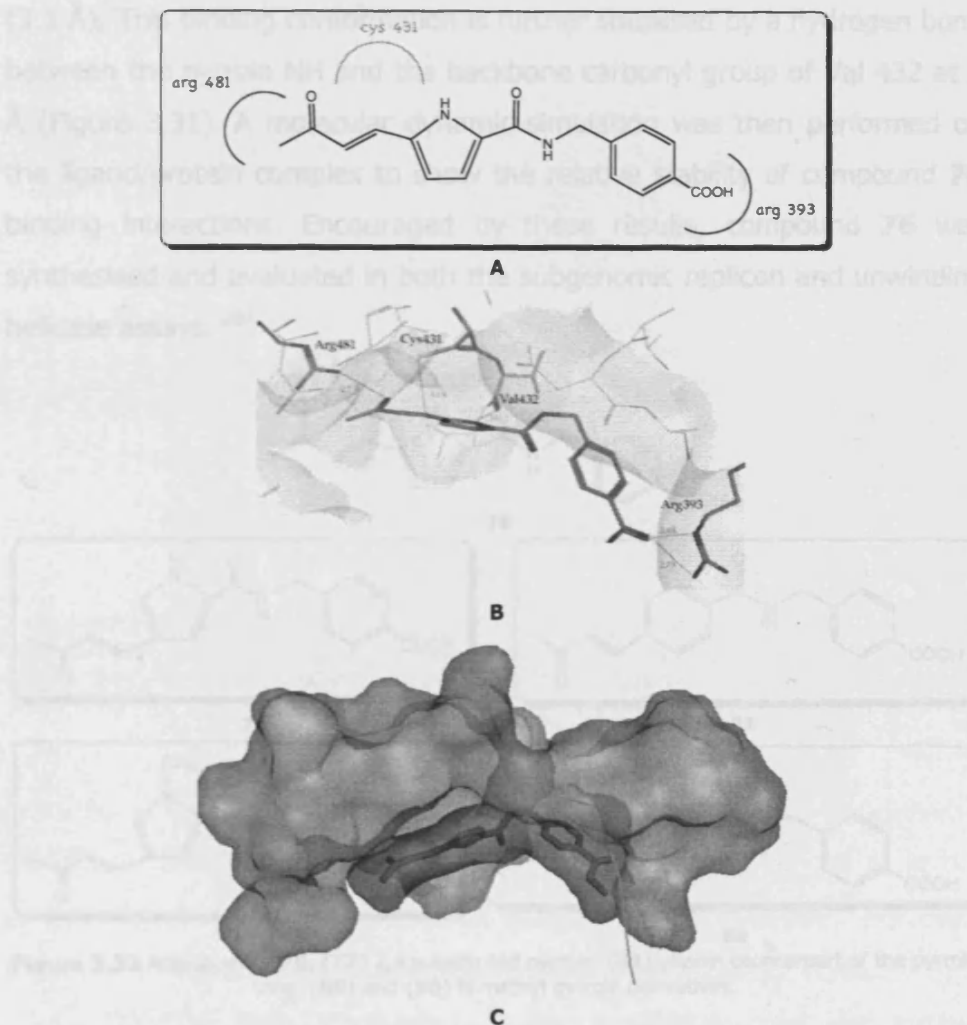


Figure 3.31 A) 2D depiction of α, β unsaturated pyrrole benzoate scaffold (**76**) B) proposed interaction patterns of **76** within NS3 helicase, C) surface representation of the proposed alignment of **76** (purple within NS3 helicase (green) and ss DNA (dUB) (orange)

Compound **76** provided the required scaffold for the plausible key interactions; hydrogen bond between the carboxylate group and guanidine group of Arg 393 at 2.7 and 2.8 Å, hydrogen bond between the carbonyl group of the vinyl ketone and Arg 481 at 2.7 Å, proximity between the reactive point of the vinyl ketone moiety and the sulfur atom of Cys 431 (3.1 Å), This binding conformation is further stabilised by a hydrogen bond between the pyrrole NH and the backbone carbonyl group of Val 432 at 3 Å (Figure 3.31). A molecular dynamic simulation was then performed on the ligand/protein complex to show the relative stability of compound **76** binding interactions. Encouraged by these results, compound **76** was synthesised and evaluated in both the subgenomic replicon and unwinding helicase assays.²⁸¹

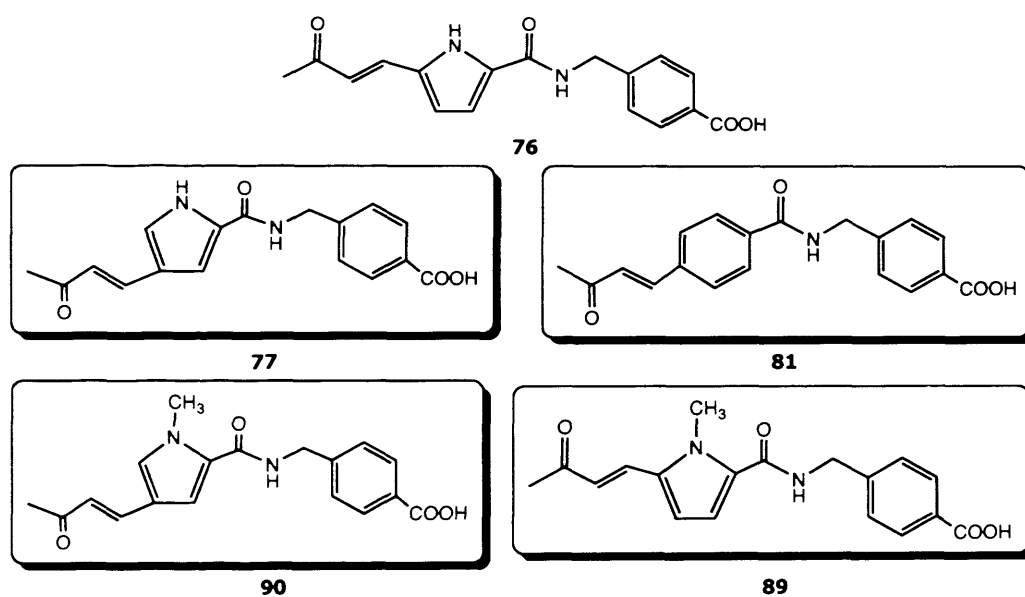


Figure 3.32 Analogues of **76**, (**77**) 2,4 substituted pyrrole, (**81**) phenyl counterpart of the pyrrole ring, (**89**) and (**90**) N-methyl pyrrole derivatives.

Alongside with **76**, four other analogues were synthesised to probe the effect of such structure modifications on the activity. The first one is the 2,4-substituted pyrrole analogue (**77**), the second is the phenyl analogue of the pyrrole ring (**81**) and the N-methyl pyrrole derivatives (**89** and **90**)

(Figure 3.32). The selection of these analogues to be prepared was based on the modeling studies and synthetic feasibility. For example, the phenyl analogue exhibited very similar interaction with Arg 393 and Arg 481, besides the proximity of the electrophilic Michael's acceptor warhead to Cys 431 residue. It was also observed that a hydrogen bond could form between the amide carbonyl group and the backbone NH of Asp 412. In addition to the hydrophobic interactions with Leu 451, Pro 452, Val 452, Thr 411 and Ala 413 as depicted in Figure 3.33.

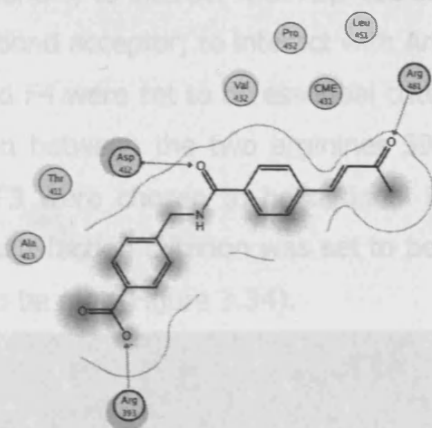


Figure 3.33 2D diagram showing the predicted binding interactions of the phenyl analogue **81** with NS3 helicase enzyme. Hydrophobic residues (green), polar residues (pink); [basic (blue ring), acidic (red ring)], H-bonds [(arrow); residue side chain (green), residue backbone (blue)] annotated proximity contour and solvent surface area exposure properties are drawn in the background of the diagram

Actually, compound **76** exhibited promising activity in the subgenomic replicon antiviral assay ($EC_{50} = 9.2 \mu\text{M}$), ($CC_{50} = 30 \mu\text{M}$) as well as in the unwinding activity assay of helicase enzyme ($IC_{50} = 0.26 \mu\text{M}$). However, the other four analogues showed less activity in both assays. (Chapter 5) These findings were considered to be a starting point for further optimisation of structure/activity/toxicity profiles²⁸¹.

3.4.4 Pharmacophore based virtual screening

In an attempt to identify compounds of similar architecture to **76**, virtual screening of the Enamine Database²⁸² was performed using the available structural data of NS3 helicase and the proposed scaffold orientation of compound **76**. Using MOE²⁶⁰ Pharmacophore Elucidation module, a 3D pharmacophore query was developed having four main features (F);

- F1, Anionic; to interact with Arg 393 residue.
- F2, F3; H-bond donors; to interact with Asp 412 and Val 432 residues.
- F4; Anionic or H-bond acceptor; to interact with Arg 481 residue.

Both features F1 and F4 were set to be essential criteria being required for the bridge formation between the two arginines 393 and 481, while the other two F2 and F3 were chosen to be optional in order to widen the search scope. The satisfaction criterion was set to be partial match with at least two features to be met (Figure 3.34).

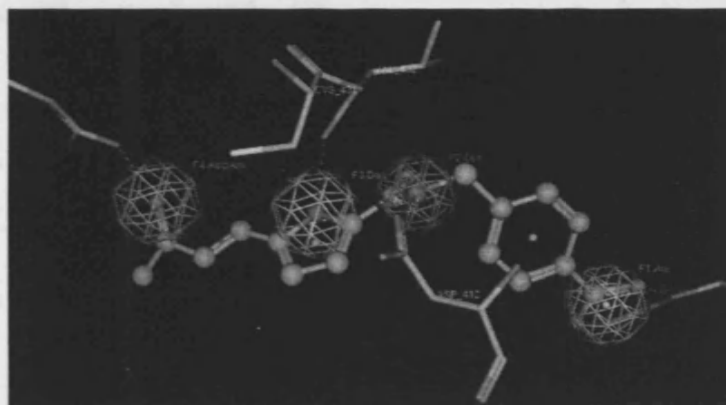


Figure 3.34 3D Pharmacophore Query based on structural data of NS3 helicase and the proposed binding mode of **72**, showing 4 features, **F1** (grey), **F2** (purple), **F3** (yellow) and **F4** (cyan)

Drug likeness filters were applied to remove compounds that fail Lipinski's rule of five²²⁰ from the Enamine Database. FlexX docking simulation was used to virtually screen the database using the 3D Pharmacophore query as a pre-filter. The output database retained the best 10 conformations for each virtual hit. Analysis of the results was based on both the score and

visual inspection of interactions with the active site residues, with a higher priority given to the chemically accessible smaller compounds, which allow further functionalisation in subsequent stages. Table 3.1 presents the top scoring sixteen virtual hits.

ID	Structure	ID	Structure
A1		A9	
A2		A10	
A3		A11	
A4		A12	
A5		A13	
A6		A14	
A7		A15	
A8		A16	

Table 3.1 chemical structures of top scoring compounds from the pharmacophore based virtual screening of enamine database.

Of particular interest was compound **A1** whose predicted binding interactions with the RNA binding site of the helicase enzyme is shown in (Figure 3.35), together with its facile preparation via one step reaction as reported by Mani *et al*²⁸³ (Figure 3.36). Actually, this compound has a methyl ketone substituent, which can provide a precursor for further functionalisation into a Michael's acceptor group so that the effect of the presence of such group on the activity of the helicase enzyme could be probed.

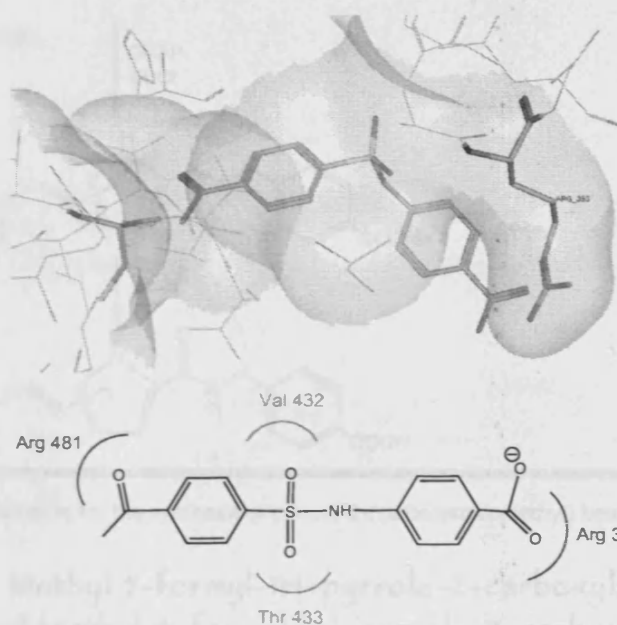


Figure 3.35 predicted binding interactions of **A1** with NS3 Helicase

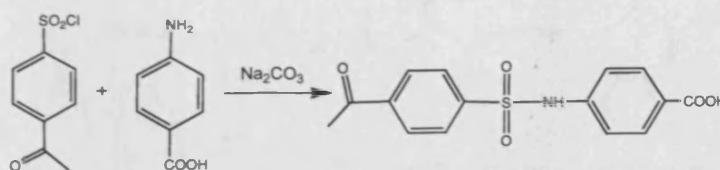


Figure 3.36 Reported one-step synthesis of **A1**

3.4.5 Synthesis of Pyrrole derivatives

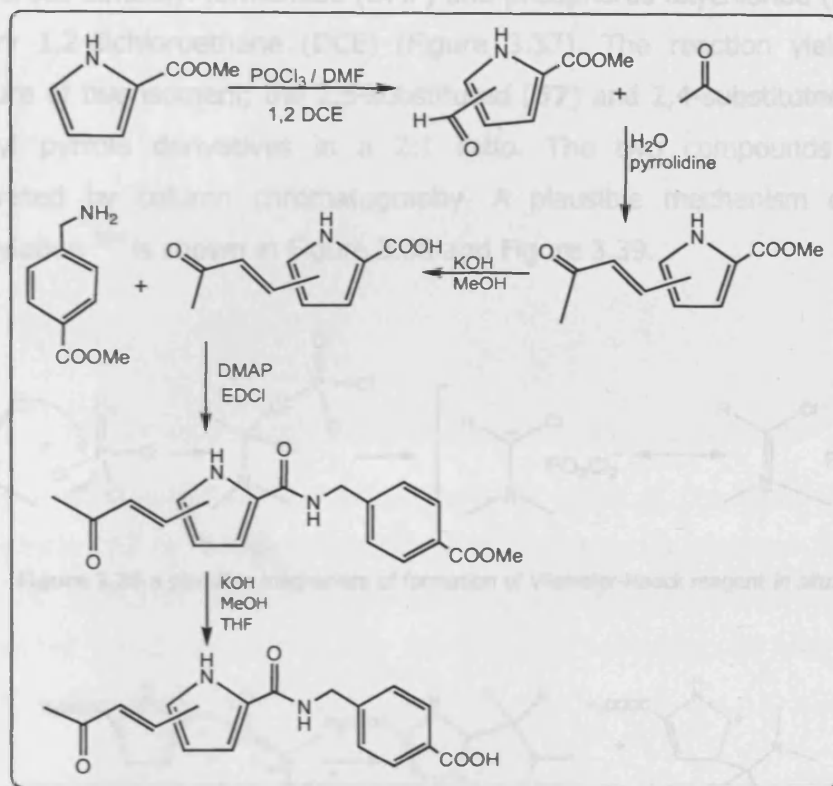
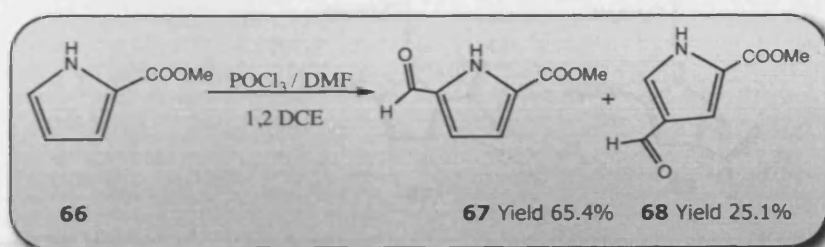


Figure 3.37 Scheme for the synthesis of pyrrole-2-(carboxamidomethyl) benzoic acid derivatives

Methyl 5-formyl-1H-pyrrole-2-carboxylate
and Methyl 4-formyl-1H-pyrrole-2-carboxylate



Formylation of methyl pyrrole-2-carboxylate (**66**) was carried out using Vilsmeier-Haack reaction which allows formylation of electron-rich arene.

According to Wallace procedure, Vilsmeier reagent is formed *in situ* from anhydrous dimethyl formamide (DMF) and phosphorus oxychloride (POCl_3) in dry 1,2-dichloroethane (DCE) (Figure 3.37). The reaction yielded a mixture of two isomers; the 2,5-substituted (**67**) and 2,4-substituted (**68**) formyl pyrrole derivatives in a 2:1 ratio. The two compounds were separated by column chromatography. A plausible mechanism of the formylation²⁸⁴ is shown in Figure 3.38 and Figure 3.39.

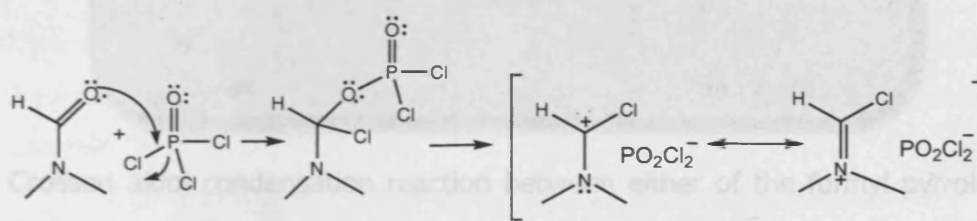


Figure 3.38 a plausible mechanism of formation of Vilsmeier-Haack reagent *in situ*.

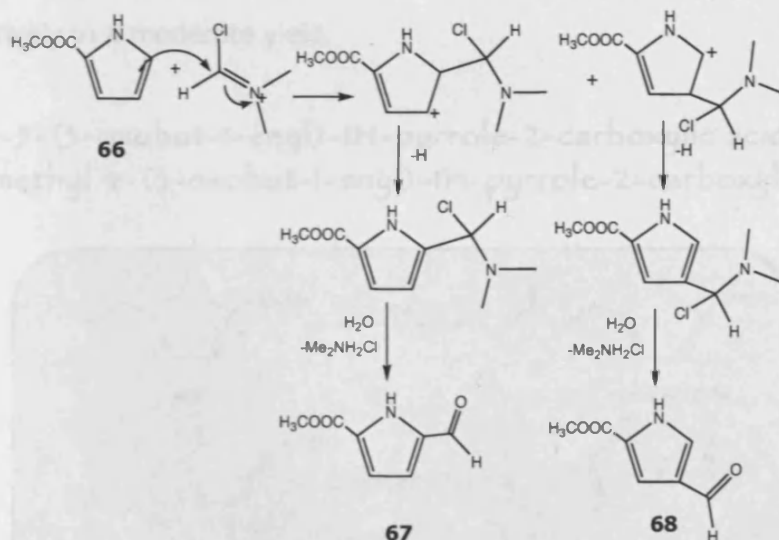
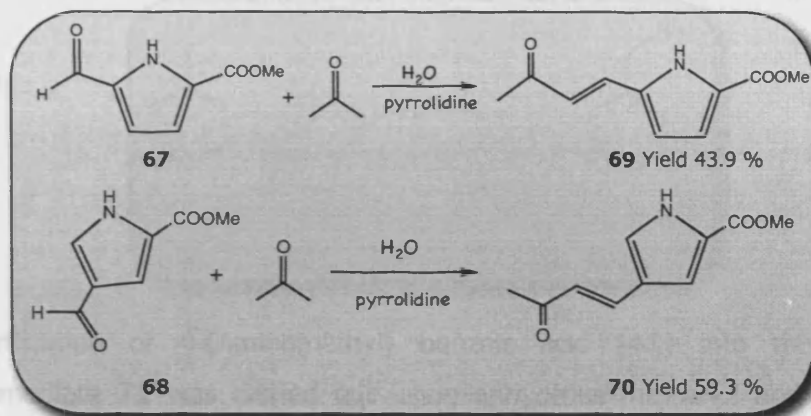


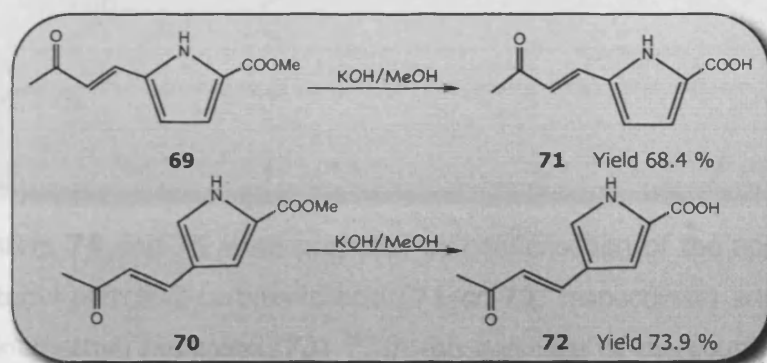
Figure 3.39 Mechanism of Vilsmeier-Haack reaction involving electrophilic aromatic substitution and hydrolysis

(E)-Methyl 5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylate and
(E)-methyl 4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylate



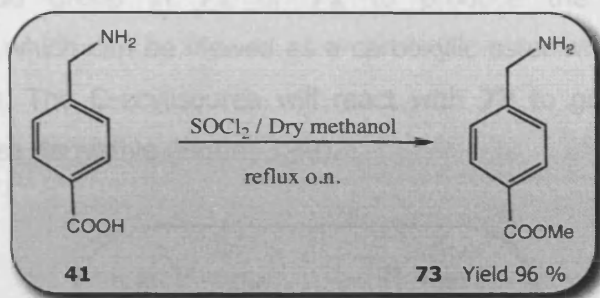
Crossed aldol condensation reaction between either of the formyl pyrrole derivatives (**67** or **68**) and acetone was conducted in aqueous medium and the presence of pyrrolidine (30 mol%) according to a reported procedure by Chimini *et al*²⁸⁵. The reaction afforded **69** and **70**, respectively in a moderate yield.

(E)-5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylic acid and
(E)-methyl 4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylic acid



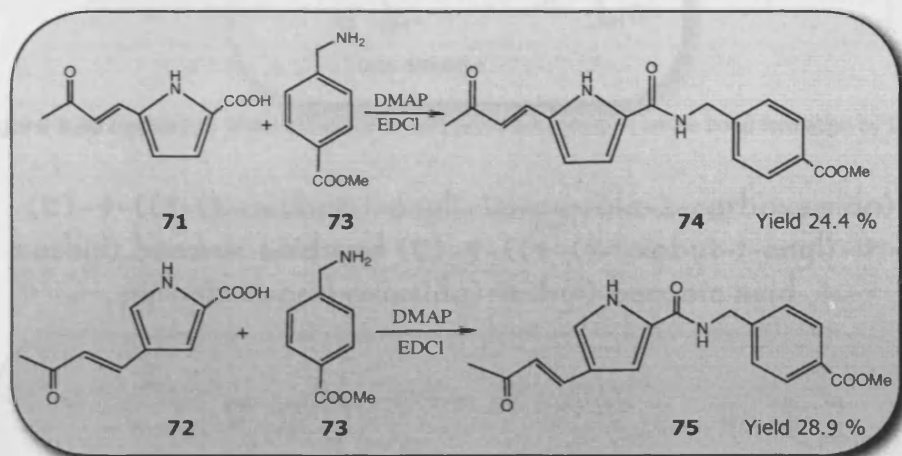
Hydrolysis of Methyl pyrrole-2-carboxylate esters **69** and **70** using KOH yielded the free carboxylic acids analogues **71** and **72**, respectively.

Methyl 4-(aminomethyl)benzoate



Esterification of 4-(Aminomethyl) benzoic acid (**41**) into the ester intermediate **71** was carried out using anhydrous methanol and thionyl chloride ²⁸⁵.

(*E*)-methyl 4-((5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoate and (*E*)-methyl 4-((4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoate



The esters **74** and **75** were prepared by condensation of the appropriate oxobutenyl pyrrole-2-carboxylic acid (**71** or **72**, respectively) and methyl 4-(aminomethyl) benzoate (**73**) ²⁸⁶ in the presence of the coupling agent 1-ethyl 3-(dimethyl aminopropyl) carbodiimide hydrochloride (EDCI) and dimethylaminopyridine (DMAP) according to the parallel synthesis protocol of Boger. ²⁸⁷

The carbodiimide (EDCI) is a dehydrating agent, which reacts with the carboxylic acid group in **71** or **72** to produce the O-acylisourea intermediate, which can be viewed as a carboxylic ester with an activated leaving group. The O-acylisourea will react with **73** to give the desired amide and urea derivative (Figure 3.40).

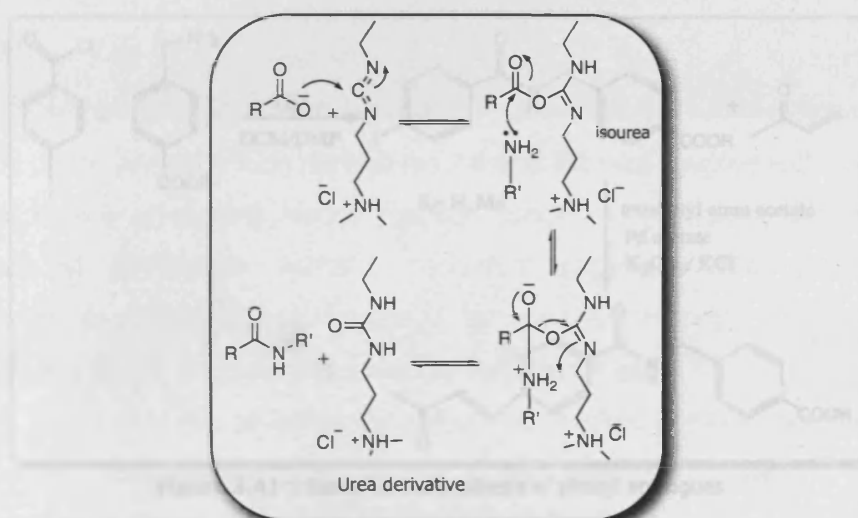
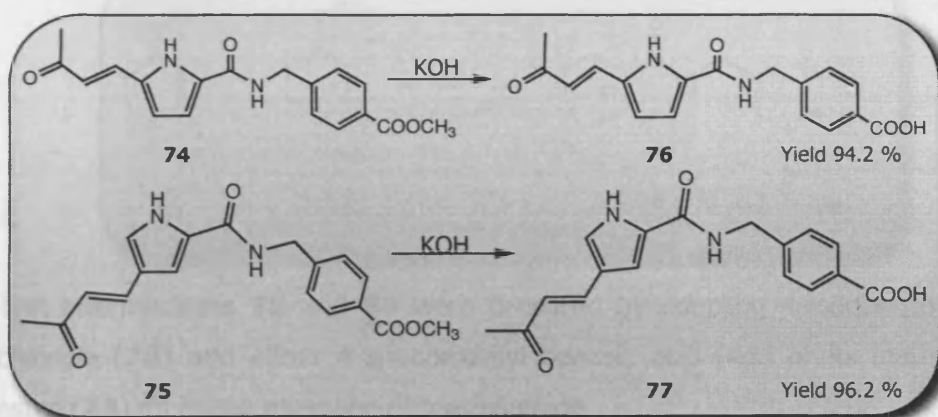


Figure 3.40 Mechanism of the activation of carboxylic acid group in amide bond formation by EDCI

(*E*)-4-((5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoic acid and (*E*)-4-((4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoic acid



Hydrolysis of the ester group in methyl pyrrole benzoate derivative (**74** and **75**) was carried out using aqueous KOH solution to afford the corresponding carboxylic acid derivatives **76** and **77**, respectively.

3.4.6 Preparation of Phenyl analogues

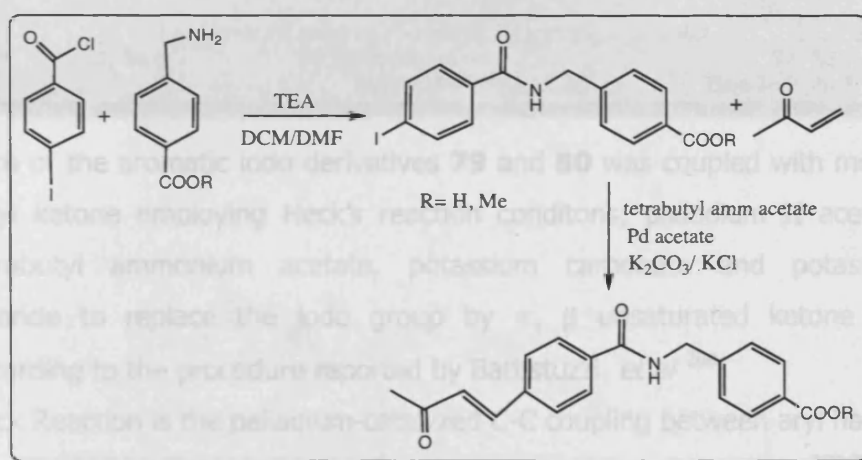
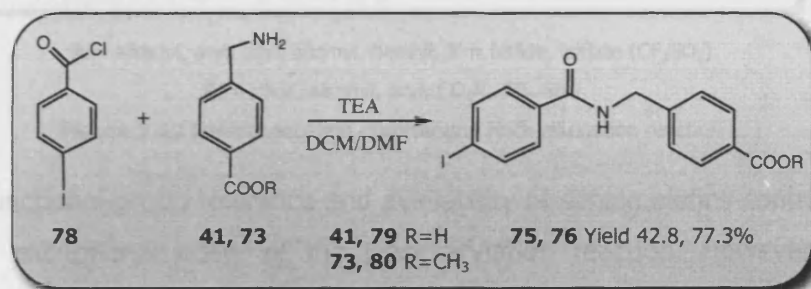


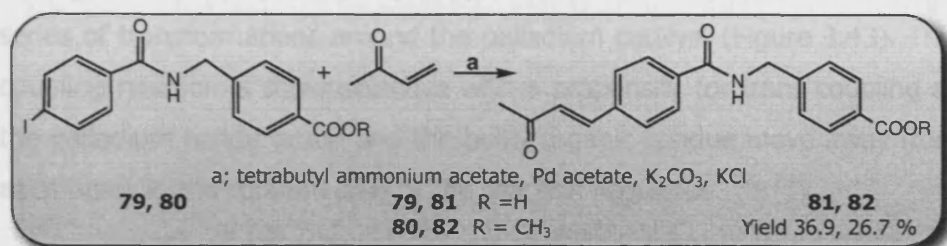
Figure 3.41 Scheme for the synthesis of phenyl analogues

4-((4-Iodobenzamido)methyl)benzoic acid and Methyl 4-((4-Iodobenzamido)methyl)benzoate



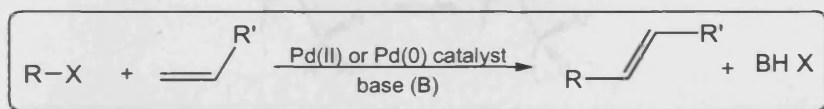
The intermediates **79** and **80** were prepared by coupling 4-iodobenzoyl chloride (**78**) and either 4-aminomethyl benzoic acid (**41**) or its methyl ester (**73**)²⁸⁶ in the presence of triethylamine.

(E)-4-((4-(3-Oxobut-1-enyl)benzamido) methyl)benzoic acid
and (E)-methyl 4-((4-(3-oxobut-1-enyl)benzamido)
methyl)benzoate



Each of the aromatic iodo derivatives **79** and **80** was coupled with methyl vinyl ketone employing Heck's reaction conditions; palladium II acetate, tetrabutyl ammonium acetate, potassium carbonate and potassium chloride to replace the iodo group by α , β unsaturated ketone one according to the procedure reported by Battistuzzi *et al*²⁸⁸.

Heck Reaction is the palladium-catalyzed C-C coupling between aryl halides or vinyl halides and activated alkenes in the presence of a base²⁸⁹⁻²⁹² as illustrated in Figure 3.42.



R = alkenyl, aryl, allyl, alkynyl, benzyl, X = halide, triflate (CF₃SO₃)

R' = alkyl, alkenyl, aryl, CO₂R, OR, SiR₃

Figure 3.42 General equation representing Heck alkenation reaction

High functional-group tolerance and availability of simple olefins contribute to the exceptional utility of the Heck arylation reaction. However the mechanism of this reaction is a topic of debate since its discovery.^{293, 294}

The most common reaction type would involve an electron-poor olefin (R' = electron-withdrawing group) and an electron rich halide (R = Ar or vinyl).

Due to the increasing aryl-X bond strengths of the different halides, the reactivity of aryl halides decreases in the order $I > Br > Cl$ ^{292,294}.

The generally accepted catalytic cycle for the Heck reaction involves a series of transformations around the palladium catalyst (Figure 3.43). This coupling reaction is stereoselective with a propensity for trans coupling as the palladium halide group and the bulky organic residue move away from each other in the rotation step of the reaction sequence. ^{290, 294}

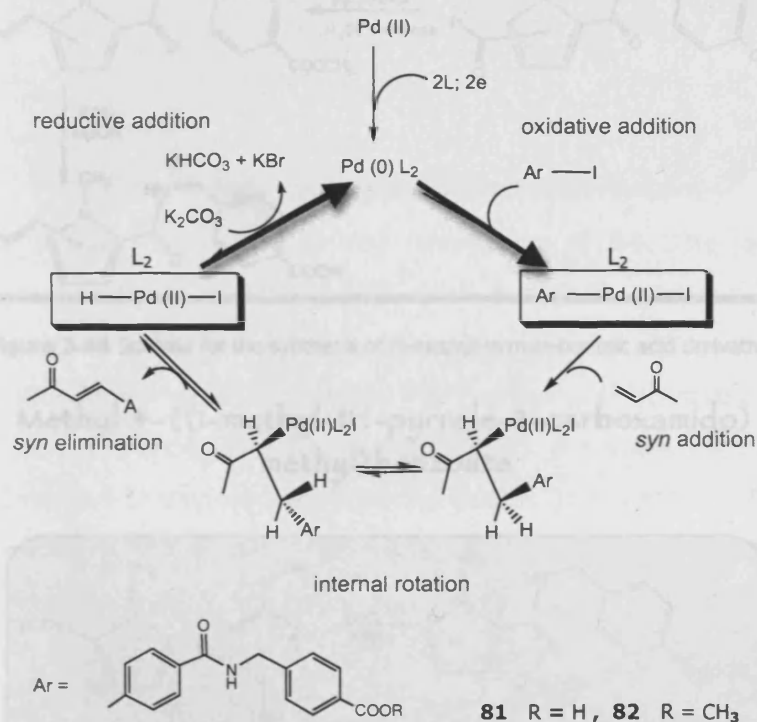


Figure 3.43 Mechanistic hypothesis of the catalytic cycle of Heck olefination

1-Ethyl 3-(3-aminobenzoyl)carboxamide hydrochloride (EDCI) was used as coupling agent between 1-methyl-1H-pyrrole-2-carboxylic acid (**83**) and methyl 4-(aminomethyl)benzoate (**73**) ⁷⁹ in the presence of dimethylamino pyridine (DMAP) to give **84** in a relatively good yield.

3.4.7 Preparation of N-MethylPyrrole derivatives

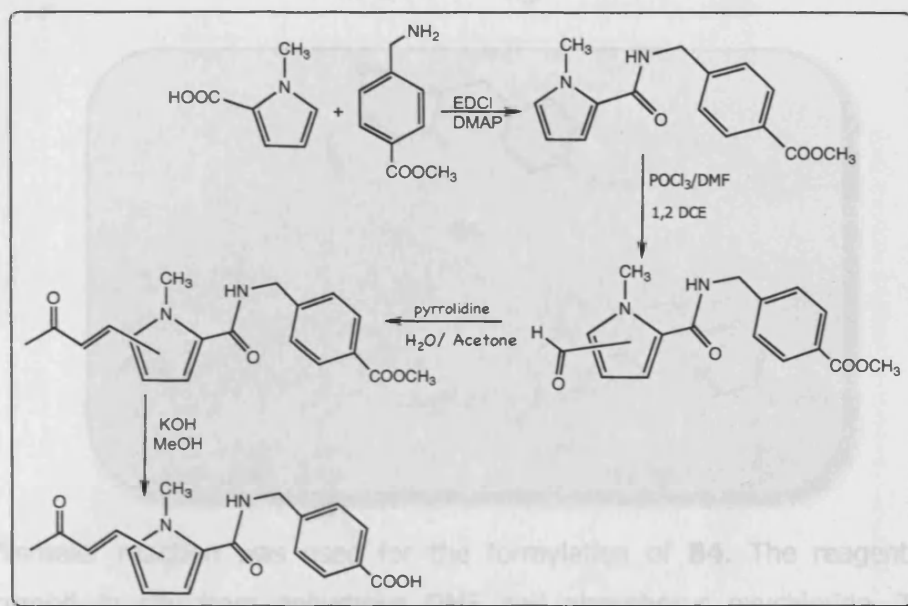
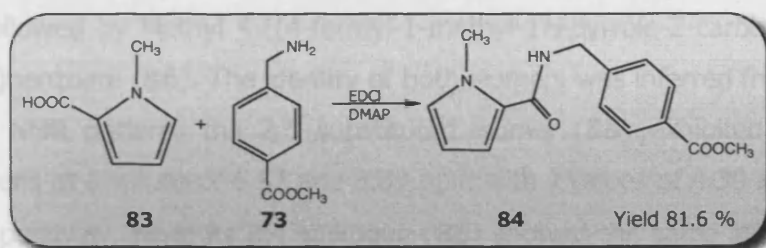


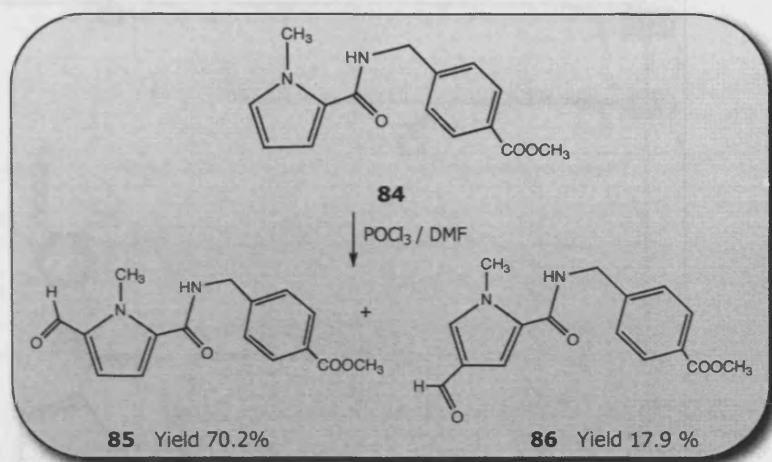
Figure 3.44 Scheme for the synthesis of N-methyl-pyrrole-benzoic acid derivatives

Methyl 4-((1-methyl-1H-pyrrole-2-carboxamido) methyl)benzoate



1-Ethyl 3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI) was used as coupling agent between 1-methyl-1H-pyrrole-2-carboxylic acid (**83**) and methyl 4-(aminomethyl)benzoate (**73**)²⁸⁶ in the presence of dimethylamino pyridine (DMAP) to give **84** in a relatively good yield.

Methyl 4-((5-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate & Methyl 4-((4-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate



Vilsmeier reaction was used for the formylation of **84**. The reagent is formed *in situ* from anhydrous DMF and phosphorus oxychloride. The reaction yielded a mixture of two isomers; the 5-formyl (**85**) and 4-formyl (**86**) pyrrole derivatives in a 4:1 ratio. The two compounds were separated by column chromatography where the first to elute was Methyl 4-((5-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl) benzoate (**85**) and then followed by Methyl 4-((4-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate (**86**). The identity of both isomers was inferred from their proton NMR pattern; the 2,5 substituted isomer (**85**) exhibited pyrrole hydrogens at δ values of 6.57 and 6.89 ppm with *J* values of 4.30 and 4.25 Hz, respectively, while its 2,4 analogue (**86**) showed the same at δ values of 7.06 and 7.38 ppm with *J* values of 1.65 and 1.70 Hz, respectively. (Figure 3.45 A and B, respectively).

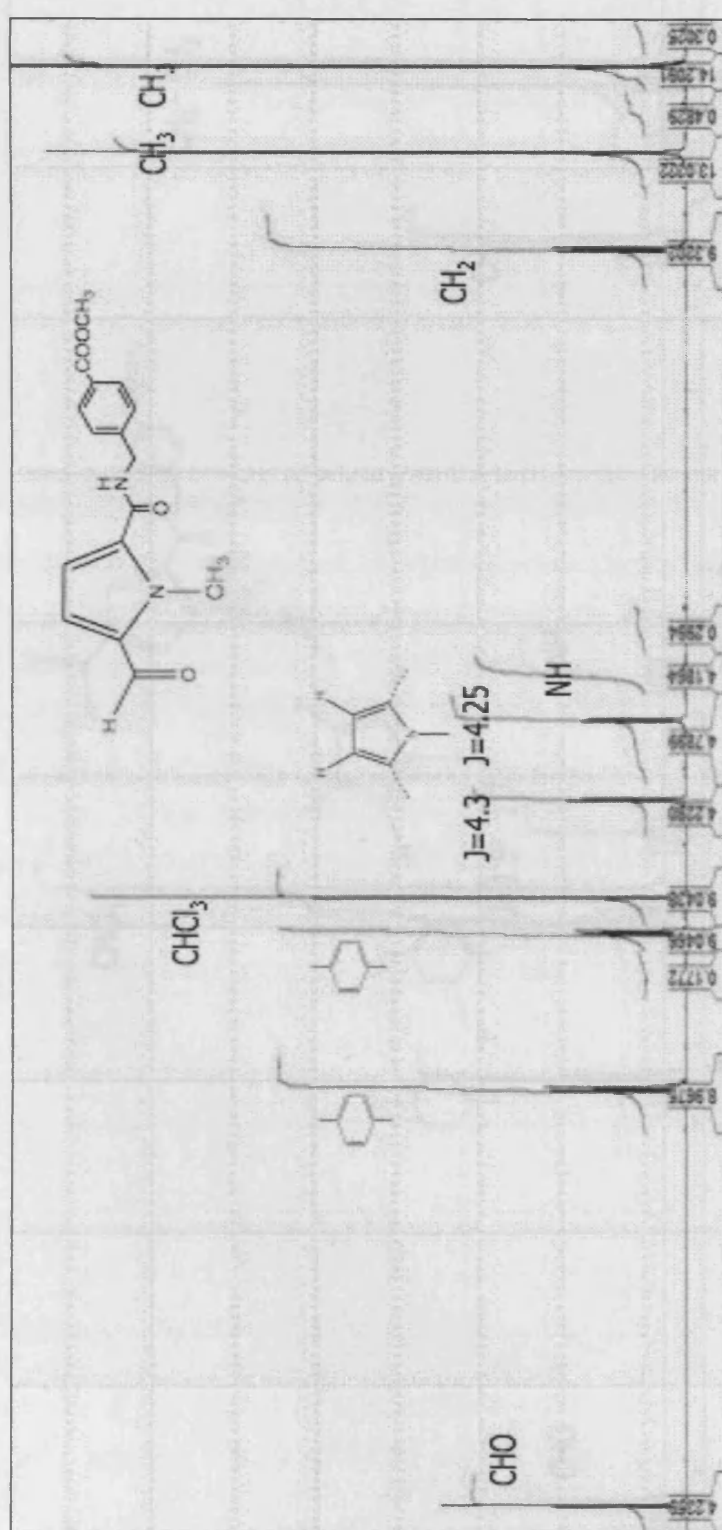


Figure 3.45 A) NMR chart of Methyl 4-((5-formyl-1-methyl-1H-pyrrole-2-carboxamido)methyl) benzoate (85)

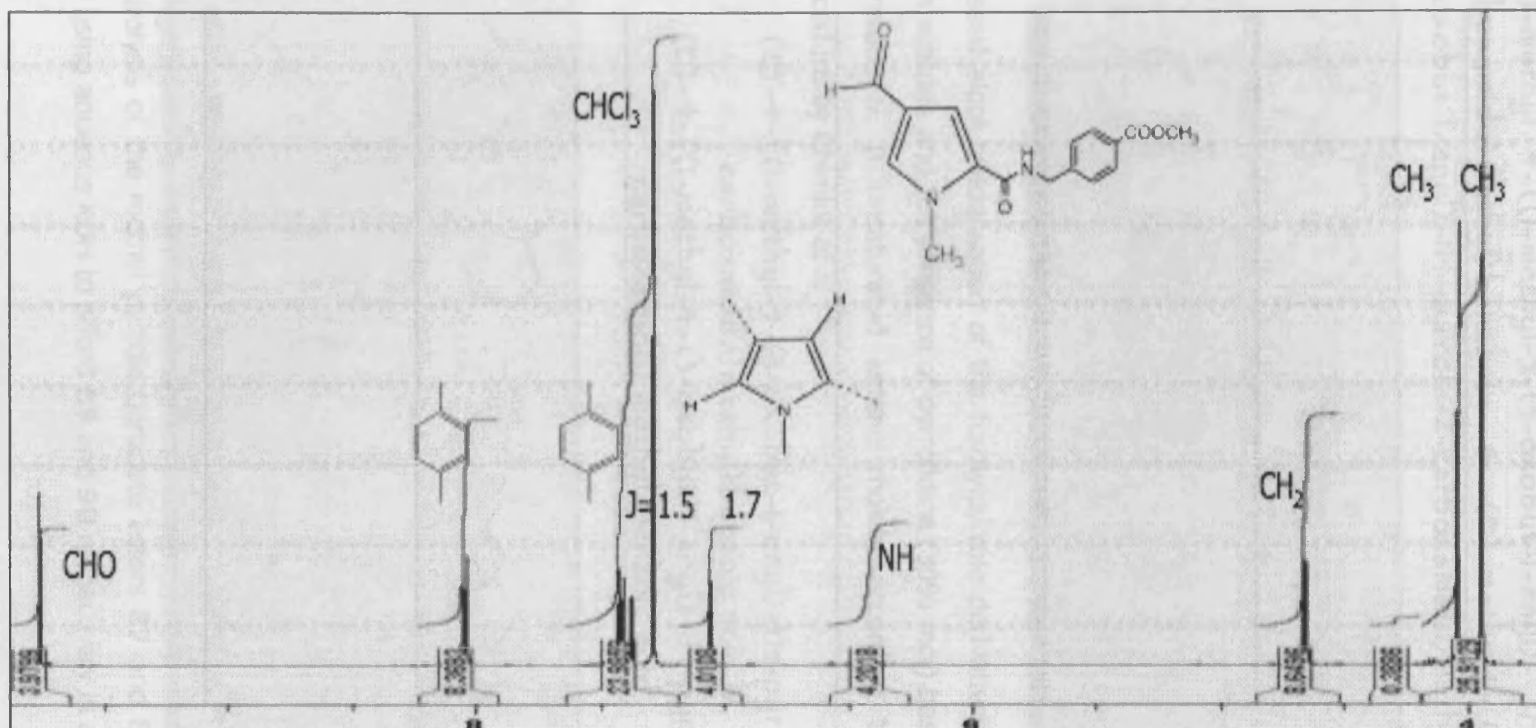
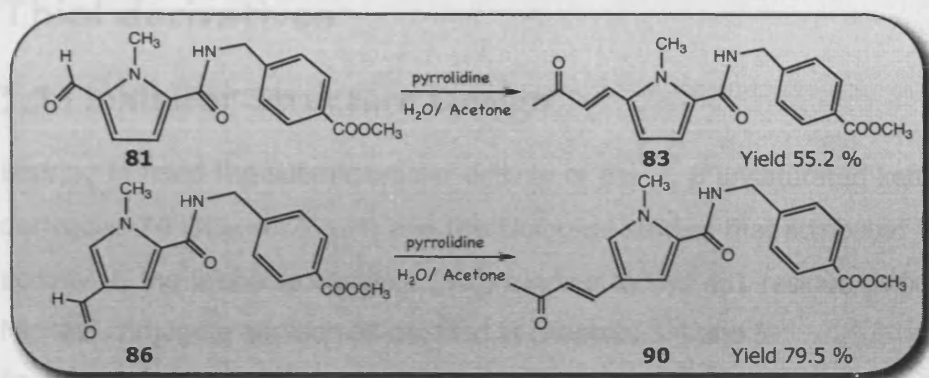


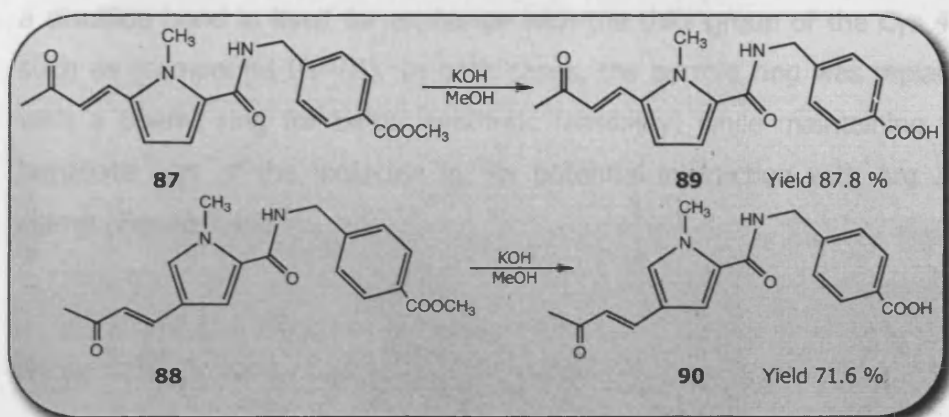
Figure 3.45 B) NMR chart of Methyl 4-((4-formyl-1-methyl-1H-pyrrole-2-carboxamido) methyl)benzoate (86)

(*E*)-methyl 4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoate & (*E*)-methyl 4-((1-methyl-4-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoate



Crossed aldol condensation of the formylpyrrole derivatives (**85** and **86**) with acetone and the presence of pyrrolidine (30% mol) base was used to introduce α , β unsaturated ketone moiety according to a reported procedure by Chimini *et al*²⁸⁵.

(*E*)-4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoic acid &
 (*E*)-4-((1-methyl-4-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoic acid



Hydrolysis of the methyl pyrrole benzoate esters **87** and **88** was carried out using aqueous KOH to afford **89** and **90** respectively in a good yield.

Chapter 3

Part 5

Thiol derivatives

3.5.1 Inhibitor Structure Design

Bearing in mind the submicromolar activity of the α , β unsaturated ketone derivative **74** ($IC_{50}=0.26 \mu\text{M}$) and the biological studies that attributed this activity to the probable covalent (S-C) binding to Cys 431 residue through Michael conjugate addition as detailed in chapters 3.4 and 5.

In this part, a different approach is adopted to improve the activity/toxicity profile of **74**. This was proposed by the replacement of the Michael acceptor group by another one capable of forming a disulfide (S-S) bond with Cys 431 thiol and hence inactivates the helicase enzyme²⁹⁵.

Cys 431 residue has particular solvent exposure and reactivity properties as indicated by the adduct formation with β -mercaptoethanol molecule in the crystal structure 1A1V of NS3 helicase (chapter 3.4).

In this context, the suggested compound would have either a free thiol group for the covalent S-S bond formation such as (compound I and II) or a disulfide bond in itself for exchange with the thiol group of the Cys 431 such as (compound III-VII). In both cases, the pyrrole ring was replaced with a phenyl ring for better synthetic feasibility, while maintaining the benzoate part of the molecule for its potential interaction with Arg 393 clamp (Figure 3.46).

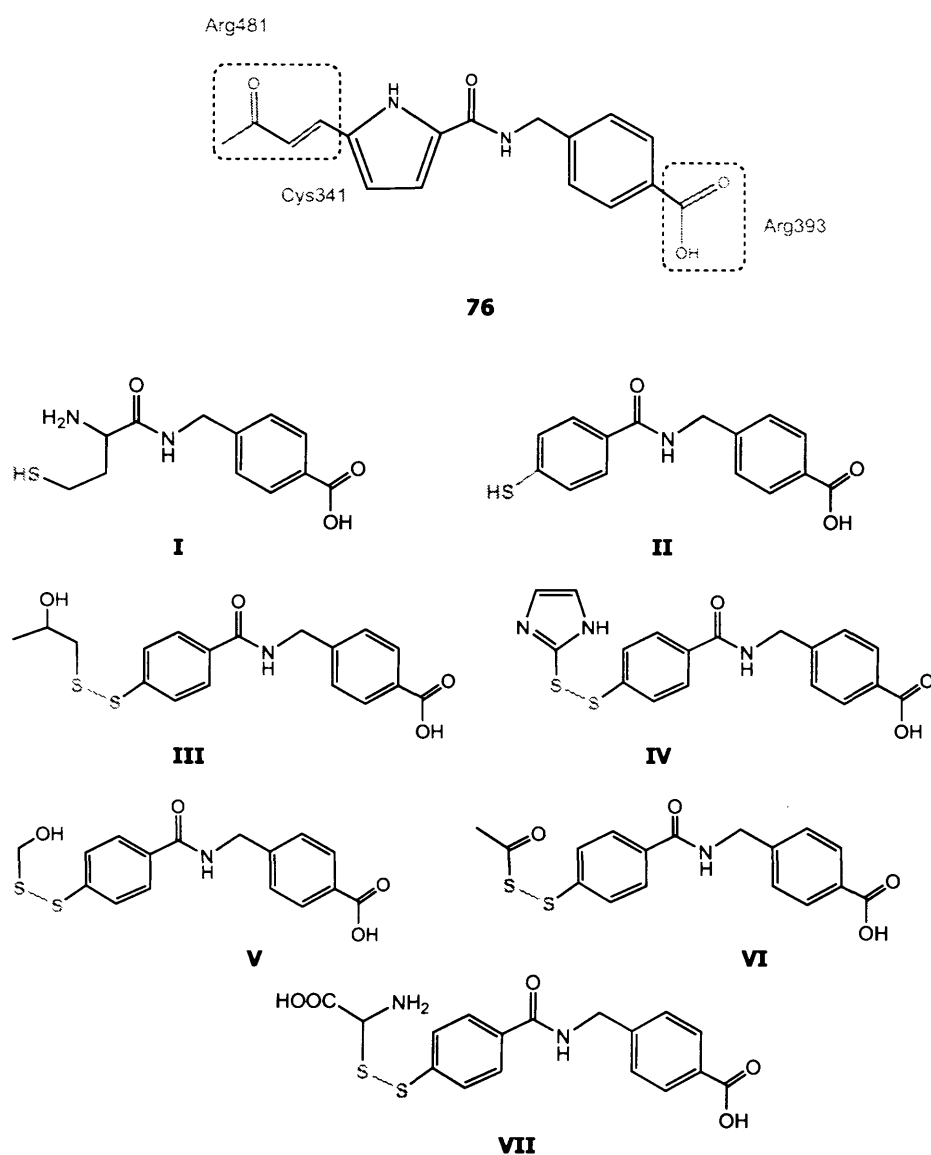


Figure 3.46 Suggested structural modification of compound **76** aiming to target Cys 431 through S-S bond formation

Compound **I** features an aliphatic thiol group without the central aromatic ring to probe its importance for the anticipated activity. Compound **II** exhibited similar binding mode to the vinyl ketone analogue as depicted in (Figure 3.47). This compound retained the central aromatic ring but with an attached thiol group instead of the vinyl ketone moiety, also it is perceivable that the putative interaction established between the ketone

oxygen of the vinyl group of compound **76** and Arg 481 residue would be missing but this will give some insight on what role this interaction plays towards the activity, moreover a better toxicity profile could be anticipated. Also it spans the distance between Arg 393 and Cys 431 efficiently with interesting interaction pattern with both amino acids.

On the other hand, compounds (**III-VI**) with different leaving groups that were examined for the plausible interactions within the putative binding site centred around Arg 393, nevertheless their interaction was not satisfactory enough to be selected for the next synthesis stage. Compound **VII** has cysteine amino acid in place of the vinyl ketone moiety, this particular amino acid has the potential not only to form a disulfide bond with Cys 431 thiol (only 4.3 Å away) but also to form a hydrogen bond with Arg 481 via its carboxylic group and interact with Thr 295 through its amino group, hence overall improvement of the enzyme/compound interactions was anticipated. Furthermore, cysteine release following the disulphide exchange with Cys 431 thiol would not pose a metabolic concern being a natural amino acid. Based on the promising docking results, and as a starting point compound **II** was selected for the subsequent synthesis step.

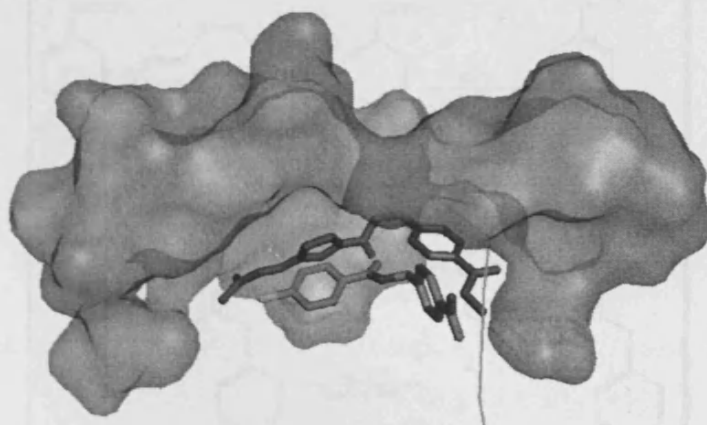


Figure 3.47 The alignment of structure **II** (coloured by element type) with compound **76** (purple) within the putative binding site (green), the nucleic acid backbone is shown as orange line

3.5.2 Synthesis of Thiol derivatives

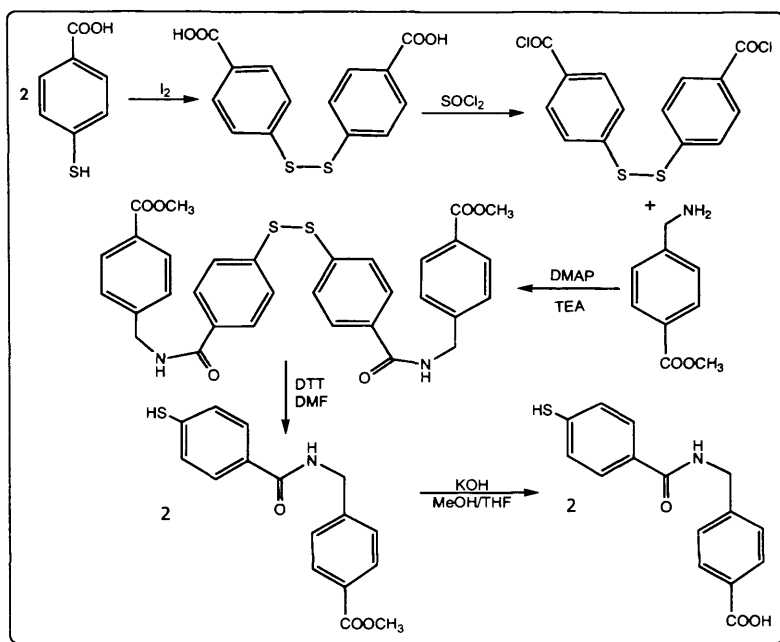


Figure 3.48 Scheme 1 for the preparation of the thiol derivative (compound II)

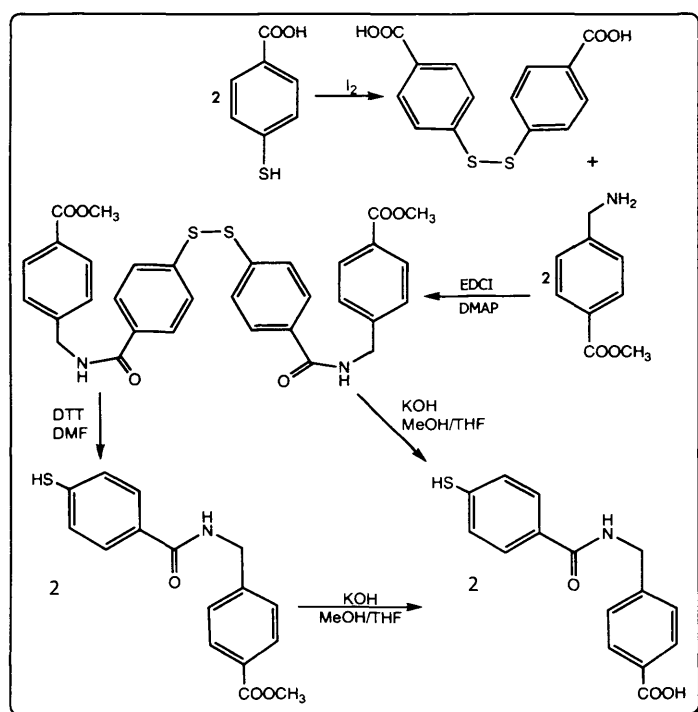
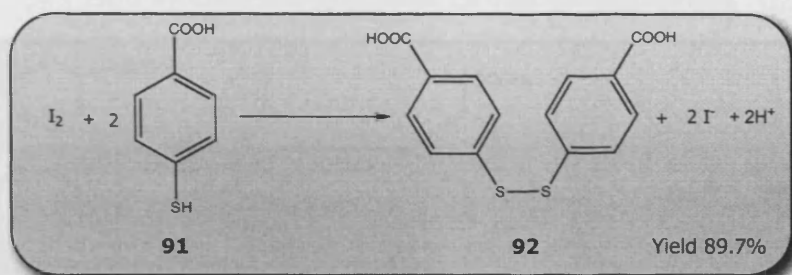


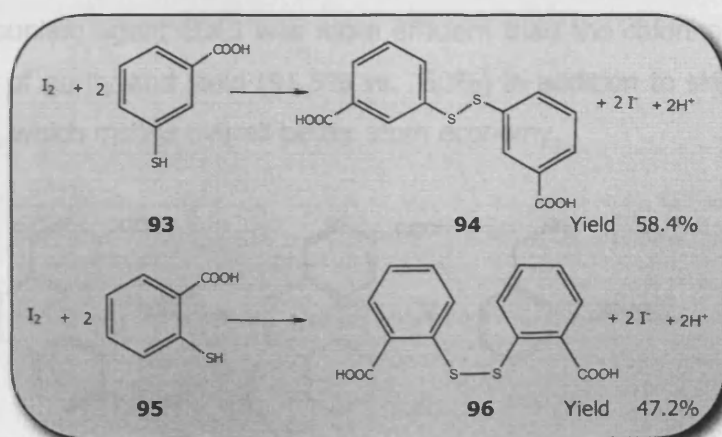
Figure 3.49 Scheme 2 for the preparation of the thiol derivative (compound II)

4,4' Dithiobisbenzoic acid



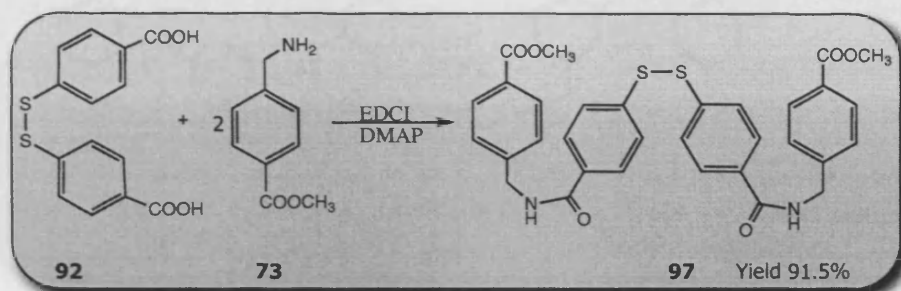
Self protection of the thiol group was done by oxidation of 4-mercaptobenzoic acid (**91**) with iodine into the disulfide dimer (**92**)²⁹⁶ in a quantitative yield. The reaction is a reduction-oxidation type where iodine acts as an oxidising agent reduced by the mercapto benzoic acid (**91**). Triethylamine was added as acid scavenger of the produced hydroiodic acid.

3,3' Dithiobisbenzoic acid & 2,2' Dithiobisbenzoic acid



Same procedure was employed to prepare both the meta and ortho isomers of **92** namely; 3,3' dithiobisbenzoic acid (**94**)²⁹⁷ and 2,2' dithiobisbenzoic acid (**96**) using 3-mercapto benzoic acid (**93**) and 2-mercapto benzoic acid (**95**), respectively.

Bis 4-((4-(methyl benzamido)methyl)benzoate) disulfide



The coupling reaction between 4,4'-disulfanediyldibenzoic acid (**92**) and methyl 4-(aminomethyl)benzoate (**73**) using 1-ethyl 3-(3-dimethyl aminopropyl)carbodiimide hydrochloride (EDCI) and dimethylamino pyridine (DMAP) was an alternative way of preparing compound **97**, which was initially prepared from **92** via chlorination with thionyl chloride into the corresponding acid chloride then reaction with methyl 4-(aminomethyl)benzoate (**73**) as depicted in Figure 3.50. However, the use of the coupling agent EDCI was more efficient than the chlorination route in terms of purity and yield (91.5% vs. 50%) in addition to skipping one step out, which means overall better atom economy.

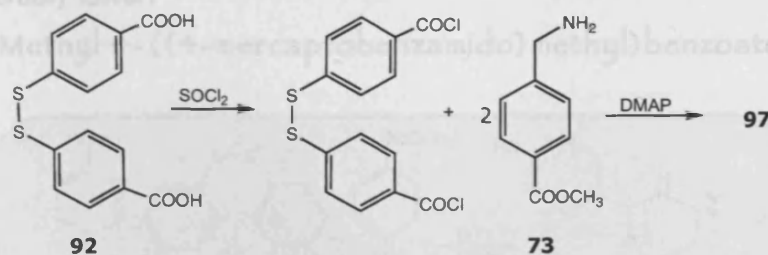
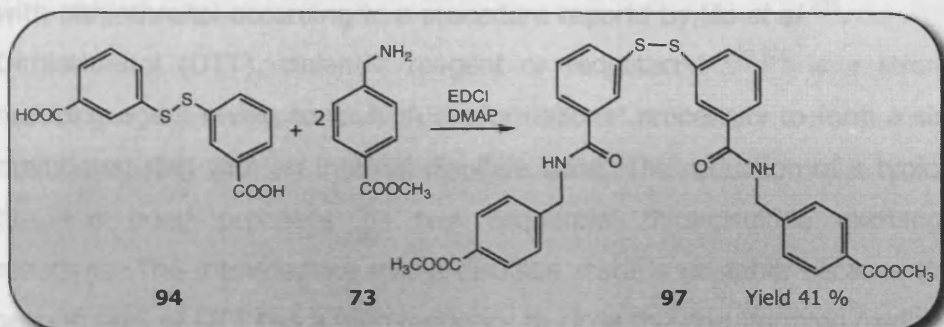
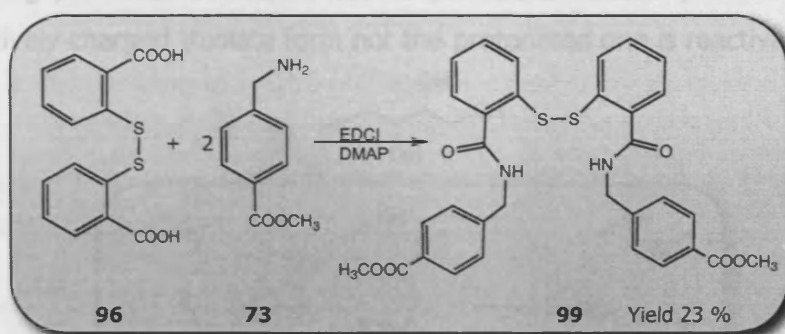


Figure 3.50 Scheme showing the alternative route for the preparation of **97**

Bis 3-((4-(methyl benzamido)methyl)benzoate) disulfide

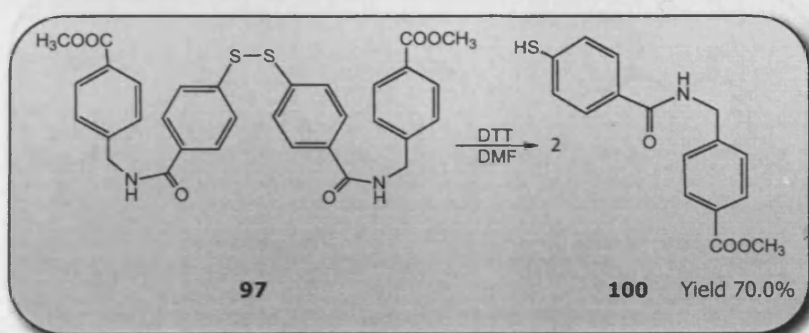


Bis 2-((4-(methyl benzamido)methyl)benzoate) disulfide



The same procedure was used to prepare the meta (**97**) and ortho (**99**) isomers of **97** using the coupling reagent (EDCI). However, the yield was considerably lower.

Methyl 4-((4-mercaptobenzamido)methyl)benzoate



The oxidised disulphide dimer **97** was reduced to the thiol analogue **100** with dithiothreitol according to a procedure reported by Hu *et al*²⁹⁸. Dithiothreitol (DTT), Cleland's reagent or reductacryl^{299,300} is a strong reducing agent owing to its high conformational propensity to form a six-membered ring with an internal disulfide bond. The reduction of a typical disulfide bond proceeds by two sequential thiol-disulfide exchange reactions. The intermediate mixed-disulfide state is unstable because the second thiol of DTT has a high tendency to close the ring, forming oxidized DTT and leaving behind a reduced disulfide bond (Figure 3.51). The reducing power of DTT is limited to pH values above 7, since only the negatively charged thiolate form not the protonated one is reactive.

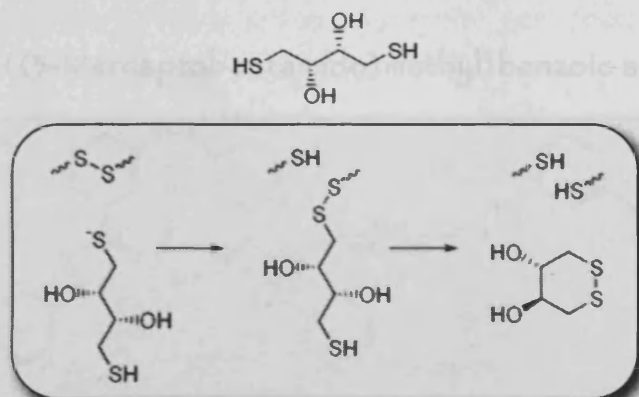
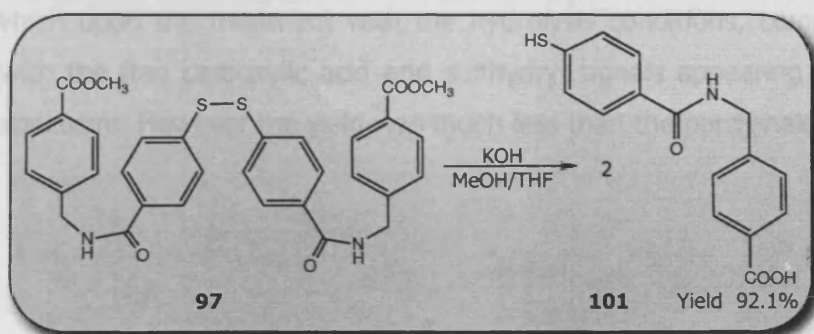


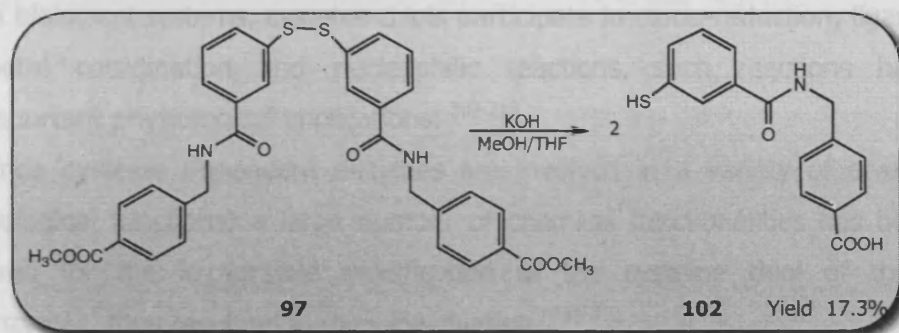
Figure 3.51 Structure and mechanism of the reduction of the disulphide bond using DTT

4-((4-Mercaptobenzamido)methyl)benzoic acid



It was devised to start with the hydrolysis of the ester group of **97** and then subsequently carry out the reduction step in order to avoid the side products resulting from the effect of the hydrolysis conditions on the free thiol group. Hydrolysis conditions were applied using aqueous potassium hydroxide. However it was striking surprise that compound **101** was obtained where hydrolysis of the ester as well as reduction of the disulphide S-S bond took place in one and the same step. Both NMR spectra as well as elemental analysis data confirmed the identity of the product. It is obvious that there is not enough data to accurately explain the mechanism of this reaction, which could be a starting point for further mechanistic study.

4-((3-Mercaptobenzamido)methyl)benzoic acid



Same observation was obtained with the meta disulfide dimer analogue **98**, which upon the treatment with the hydrolysis conditions, compound **102** with the free carboxylic acid and sulfhydryl signals appearing in the NMR spectrum. However the yield was much less than the para analogue.

Chapter 3

Part 6

1,2,4-Thiadiazole derivatives

3.6.1 Cysteine; A special amino acid

The multiple oxidation states of sulfur (6, 4, 2, 0 and -2) have been explained by its ability to accommodate extra electrons in its valence shell by using the available d-orbitals. While the four fundamental classes of reactions on the carbon atom are substitution, addition, elimination and rearrangement, those on the sulfur atom also include oxidation and reduction reactions. Cysteine is the only amino acid possessing a thiol moiety among the 20 natural amino acids.³⁰¹⁻³⁰⁶

In biological systems, cysteine thiols participate in oxido-reduction, ligand-metal coordination and nucleophilic reactions, such reactions have important physiological implications.³⁰⁷⁻³¹³

Since cysteine dependent enzymes are involved in a variety of diverse biological functions; a large number of chemical functionalities has been used for the irreversible modification of the cysteine thiol of these enzymes, thus resulting in their inactivation.³¹⁴⁻³¹⁹

Recently, AG7088, a Michael acceptor and a cysteine protease inhibitor is being developed for the treatment of common cold and has advanced to phase II/III clinical trials (Figure 3.52).³⁰²

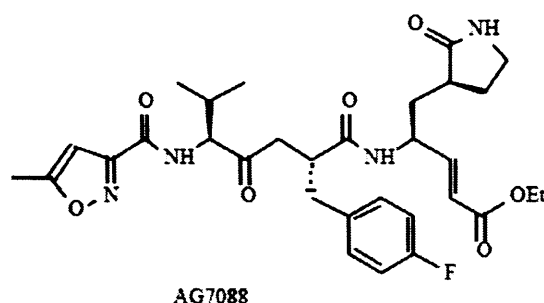


Figure 3.52 AG7088 an example of Michael acceptor inhibitor of a cysteine protease

3.6.2 [1,2,4-Thiadiazoles (THDs)]

1,2,4-Thiadiazole is a distinctive class of small heterocyclic thiol trapping agents that serve as an interesting pharmacophore in the design of inhibitors targeting the cysteine residues of proteins.³²⁰⁻³³⁰

Although currently the only commercial 1,2,4 thiadiazole drug is the antibiotic cefozopran³²¹ there are a number of synthetic products related to this system with a broad range of biological activities for example, SCH-202676 was identified as a promising allosteric modulator of G-protein coupled receptors³²² and KC 12291 with cardioprotective action³²³. Recently, the small heterocyclic thiadiazolidinones (TDZD-8) were described as the first non-ATP competitive glycogen synthase kinase 3 β inhibitors³²⁴ (Figure 3.53).

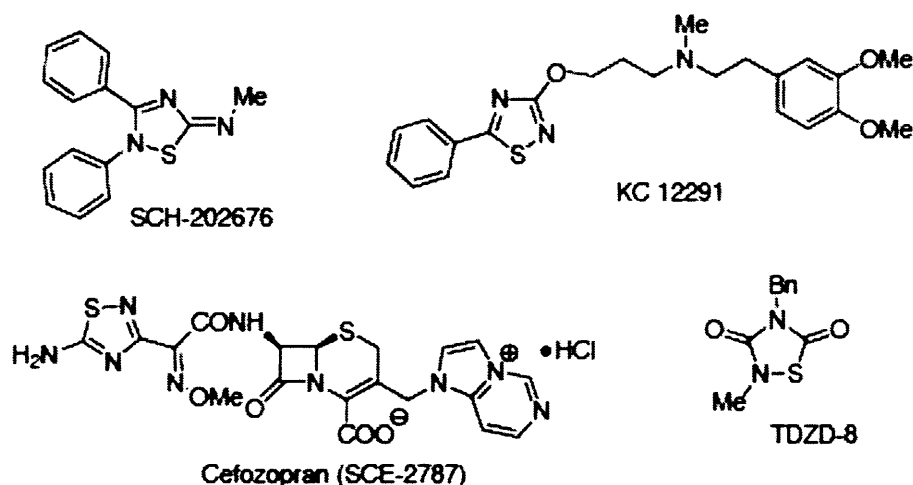


Figure 3.53 Some relevant structures with the 1,2,4-thiadiazole moiety

Basically, there are three stable chemical classes of 1,2,4-thiadiazoles (THDs); the monocyclic 1,2,4-thiadiazole (**I**), the bicyclic imidazo[1,2-d][1,2,4]thiadiazole (**II**), and the tricyclic benzo[4,5]imidazo[1,2-d][1,2,4]thiadiazole (**III**) (Figure 3.54). All can react with enzyme cysteine residue to form a disulfide adduct and thus inhibiting the enzyme.³³¹⁻³³⁴

One particular advantage of thiadiazole derivatives **I-III** is the lack of reactivity (S-N bond cleavage) with other nucleophiles such as amines and alcohols, which makes this heterocycle an excellent pharmacophore to be incorporated into lead molecules for targeting cysteine residue.³³⁵⁻³³⁹

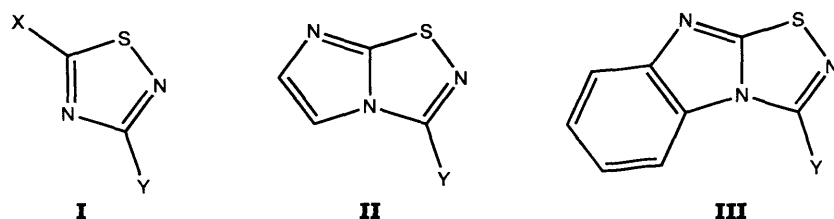


Figure 3.54 Monocyclic **I**, bicyclic **II** and tricyclic **III** 1,2,4-thiadiazoles

Ideally, any inhibitor must not only be enzyme specific but also active site directed. 1,2,4-Thiadiazoles meet this requirement as they can serve as both thiol trapping agents and at the same time the C3 and/or C5 substituents (Y and/or X) as a recognition arm to ensure the THD warhead is active site directed while the other substituent is reactivity tuner of the THD ring opening³³¹. However, in the case of the fused ring in the bicyclic (**II**) and tricyclic (**III**) thiadiazoles, the C3 substituent Y can be used to tune both affinity and reactivity of the N-S bond towards the incoming thiol nucleophile.³³²⁻³³⁴

3.6.3 Mechanism of Cysteine enzymes Inhibition

The mechanism of this adduct formation is the enzyme cysteine thiol attacks the sulfur atom in the heterocyclic ring to form a disulfide bond with concomitant ring opening. Depending on the substituents, the reaction can be reversible or irreversible³⁰² (Figure 3.55).

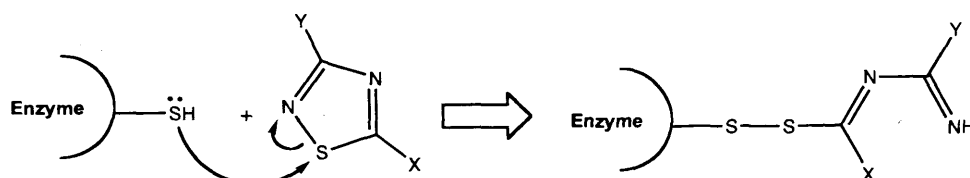


Figure 3.55 Depiction of S-S bond formation between cysteine thiol and [1,2,4]-thiadiazole ring. X and Y substituents tune both reactivity and directionality of the ring.

3.6.4 Proof of Principle

In biological systems cross-linking of proteins through disulfide bond formation increases their stability, and this is a common occurrence. Its disruption through the use of THD may be a viable strategy in the modification of enzyme cysteine residues³²⁰. Initially, there was ample evidence that the N-S bond of the 1,2,4-thiadiazole moiety is cleaved and form a disulfide bond with organic thiols³³⁷⁻³⁴¹. It was reasoned that a similar type of reaction might occur in a biological system. However, the proof of principle comes from the protein X-ray crystal structure of the papain-Apo1073 complex, which clearly reveals the formation of a disulfide bond between Cys 25 of papain (a cysteine protease from papaya fruit) and the sulfur atom of the thiadiazole moiety of Apo1073 inhibitor^{342,343} (Figure 3.56).

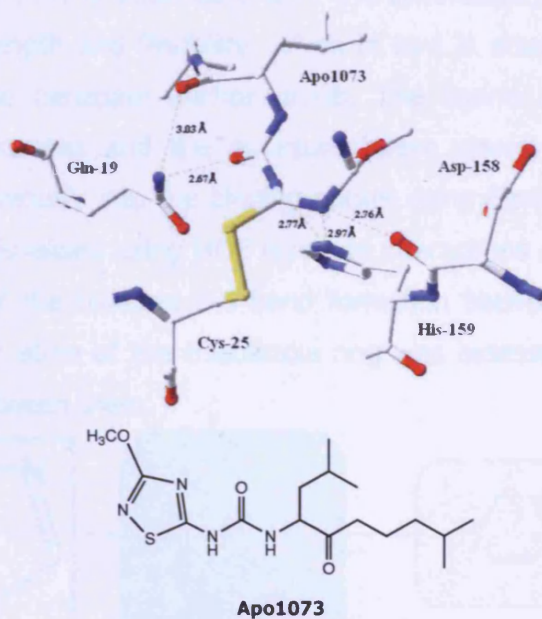


Figure 3.56 X-Ray crystal structure of the papain-Apo1073 adduct depicting the formation of S-S bond between Cys 25 and sulfur atom of Apo1073

3.6.5 Inhibitor Structure Design

In an attempt to find a pharmacophore group with better toxicity profile than the Michael acceptor in **76** (Chapter 3.4). The α , β unsaturated ketone moiety was replaced by the electrophilic 1,2,4 thiadiazole heterocycle that would maintain the interaction with Cys 431 while the benzoate part of the molecule was retained for the plausible interaction with the Arg 393 clamp (Figure 3.57).

A focused library of structures (Figure 3.58) with the general formula illustrated in Figure 3.57 was proposed. The THD ring was used with and without a methyl or methoxy substitution at position 3 of the ring. This was chosen partly from a synthetic point of view since the pathway of the intermediates 5-amino-3-methyl-1,2,4-thiadiazole (**105**) and 5-Chloro-3-methoxy-1,2,4-thiadiazole (**110**) is detailed by Marrano et al.³³⁹. At the recognition arm (C5) is attached one of the synthetically feasible linkers, with different length and flexibility, which in turn is attached to different positions of the benzoate anchor group. The amino acid residues of interest were located and the structures were energy minimised and docked simultaneously into the binding pocket using FlexX²⁶¹. The output database was visualised using MOE compute interactions option²⁶⁰, where the likelihood of the covalent S-S bond formation between Cys 431 thiol and the sulphur atom of the thiadiazole ring was assessed by measuring the distance between them.

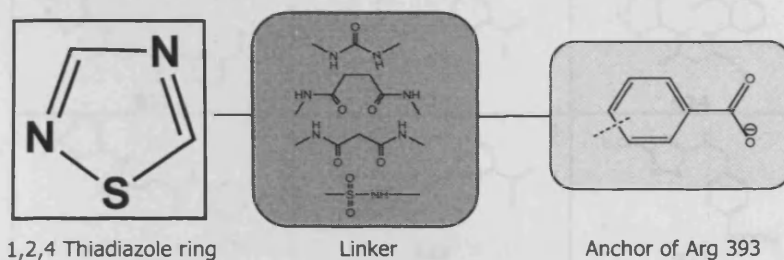


Figure 3.57 The general formula used to build the focused library of 1,2,4-thiadiazole derivatives

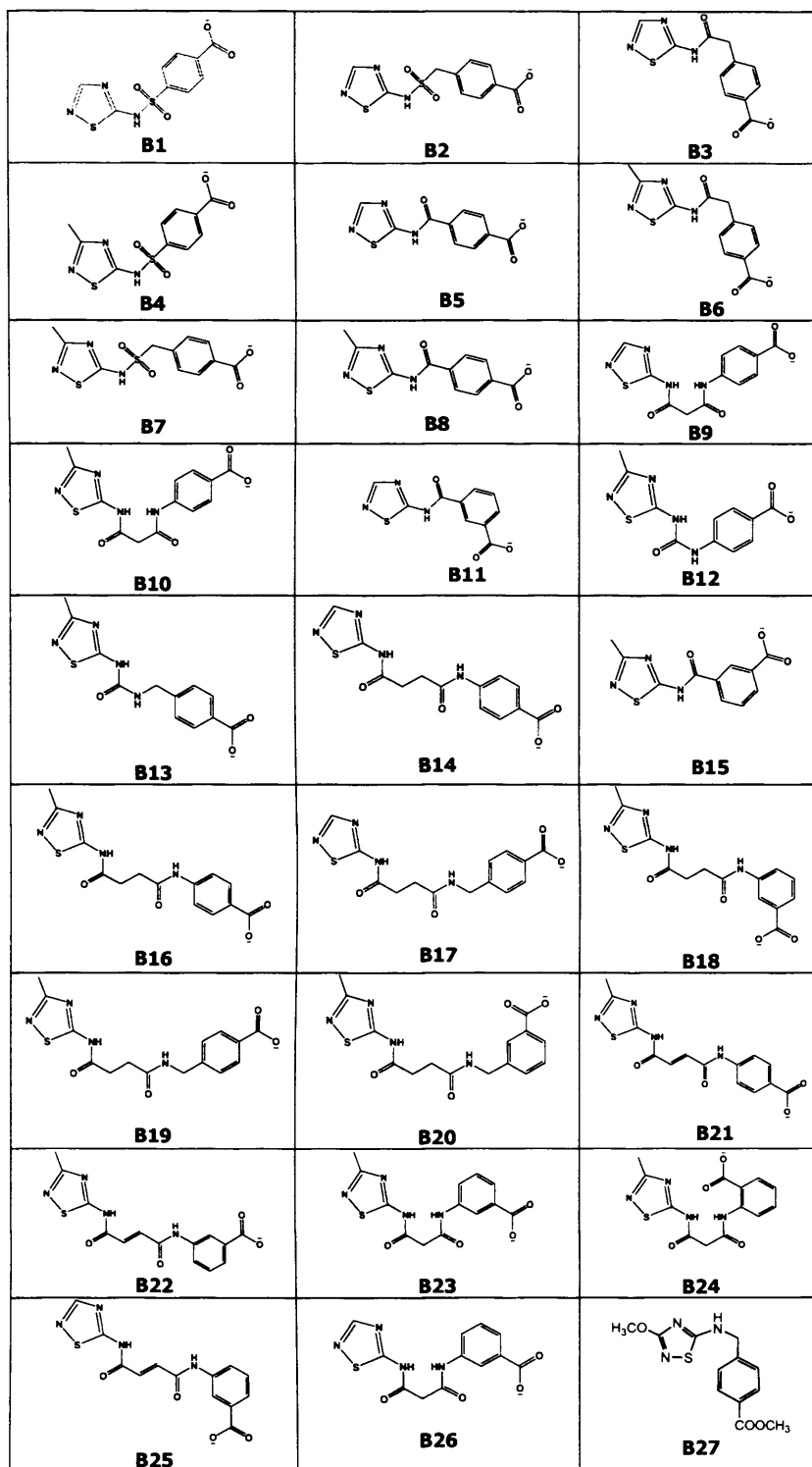
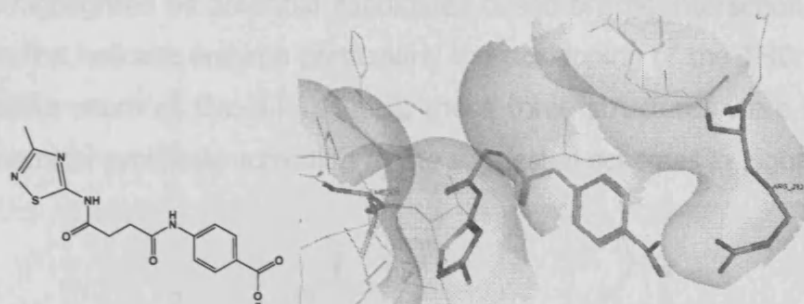
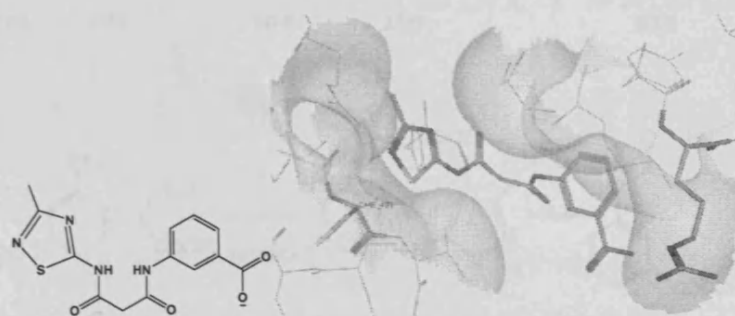


Figure 3.58 sample of the structural changes used in building the focused virtual library

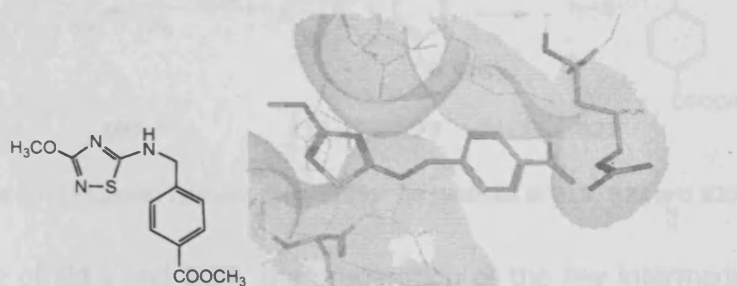
3.6.6 Selected compounds for chemical synthesis

**B16**

4-(4-(3-Methyl-1,2,4-thiazol-5-ylamino)-4-oxobutanamido)benzoate

**B23**

3-(3-(3-Methyl-1,2,4-thiazol-5-ylamino)-3-oxopropanamido)benzoate

**B27**

Methyl 4-((3-methoxy-1,2,4-thiazol-5-ylamino)methyl)benzoate

Figure 3.59 Promising candidates selected for subsequent synthesis step

From the docking simulation results compounds (**B16**, **B23** and **B27**) were highlighted as potential candidates based on the interaction pattern within the helicase enzyme particularly the positioning of the THD towards the sulfur atom of Cyc 431. Hence, these three structures were selected for chemical synthesis according to the suggested schemes in Figure 3.60.

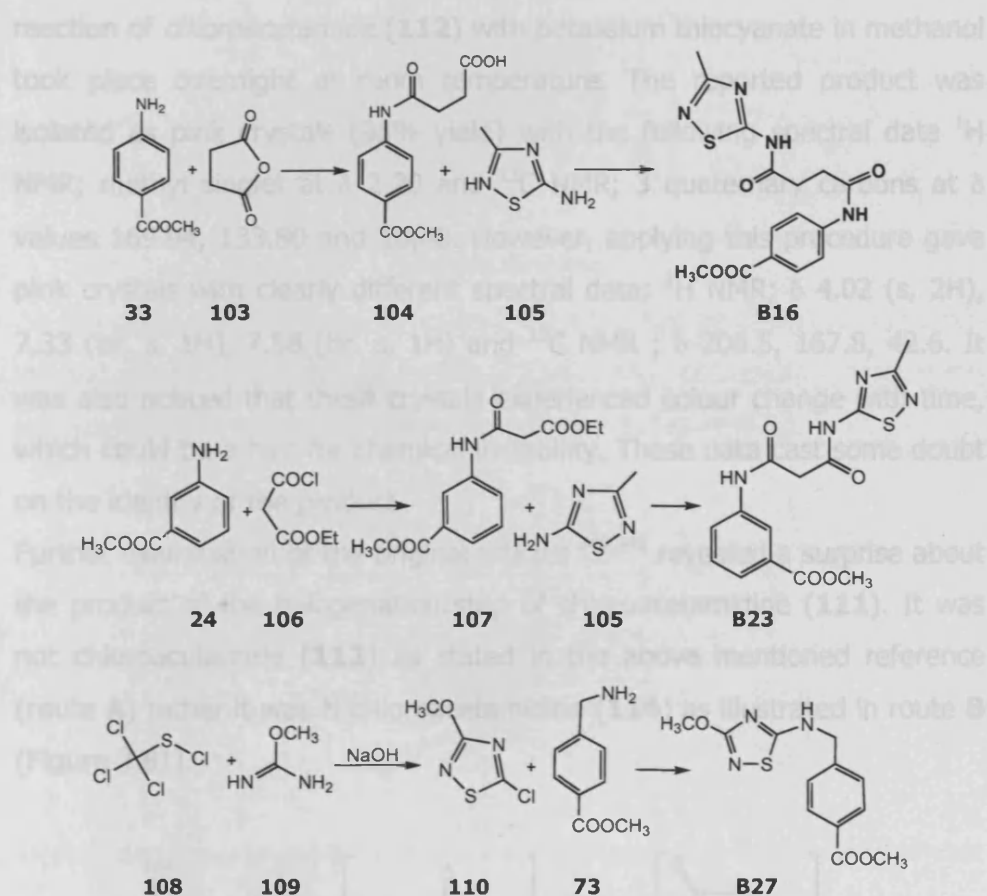


Figure 3.62 Proposed synthetic pathways for the synthesis of **B16**, **B23** and **B27**

In the case of **B16** and **B27**, the preparation of the key intermediate, 5-amino-3-methyl-1,2,4-thiadiazole (**101**), was devised according to a two step procedure reported by Marrano et al.³³⁹ (Figure 3.63). The first step is the halogenation of acetamidine hydrochloride (**111**) using sodium hypochlorite followed by extraction of "chloroacetamide".

In spite of the reported high yield for this reaction (82%), following the same procedure did not result in the expected product. There was no obvious reason for this, however chloroacetamide being commercially available, was purchased from Aldrich.

The second step was carried out according to the same reference, where a reaction of *chloroacetamide* (**112**) with potassium thiocyanate in methanol took place overnight at room temperature. The reported product was isolated as pink crystals (98% yield) with the following spectral data ^1H NMR; methyl singlet at δ 2.20 and ^{13}C NMR; 3 quaternary carbons at δ values 169.94, 133.80 and 18.48. However, applying this procedure gave pink crystals with clearly different spectral data; ^1H NMR; δ 4.02 (s, 2H), 7.33 (br. s, 1H), 7.58 (br. s, 1H) and ^{13}C NMR ; δ 206.5, 167.8, 42.6. It was also noticed that these crystals experienced colour change with time, which could be a hint for chemical instability. These data cast some doubt on the identity of the product.

Further examination of the original articles^{340,341} revealed a surprise about the product of the halogenation step of chloroacetamide (**111**). It was not chloroacetamide (**112**) as stated in the above mentioned reference (route **A**) rather it was N-chloroacetamide (**114**) as illustrated in route **B** (Figure 3.61).

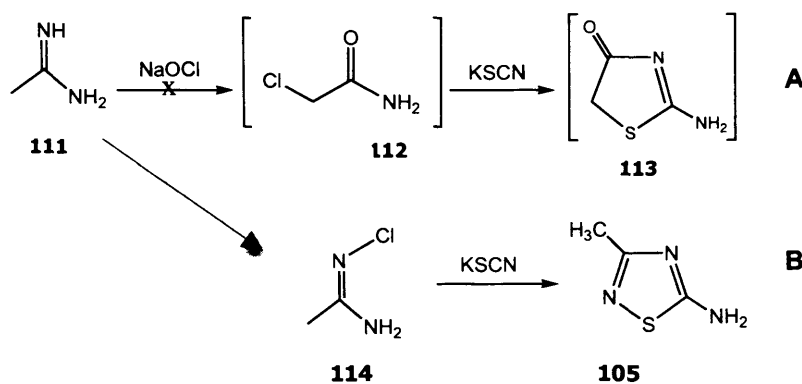


Figure 3.61 synthetic pathway of 105

The inconsistent spectral data of the product obtained from the reaction between chloracetamide (**112**) and potassium thiocyanate was attributed to the formation of pseudothiohydantoin (**113**) according to the mechanism depicted in Figure 3.62 and was proved by comparing the obtained NMR data with the reported ones for pseudothiohydantoin (**113**).

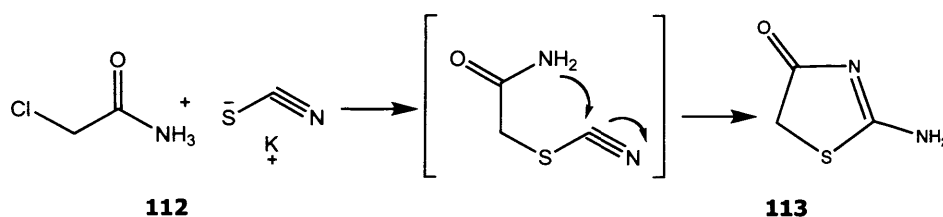


Figure 3.62 Proposed mechanism for pseudothiohydantoin (**113**) from chloracetamide (**112**) and potassium thiocyanate

On the other hand, 5-amino-3-methyl-1,2,4-thiadiazole (**105**) should have formed from N-chloroacetamidine (**114**) and potassium thiocyanate according to the suggested mechanism shown in Figure 3.63

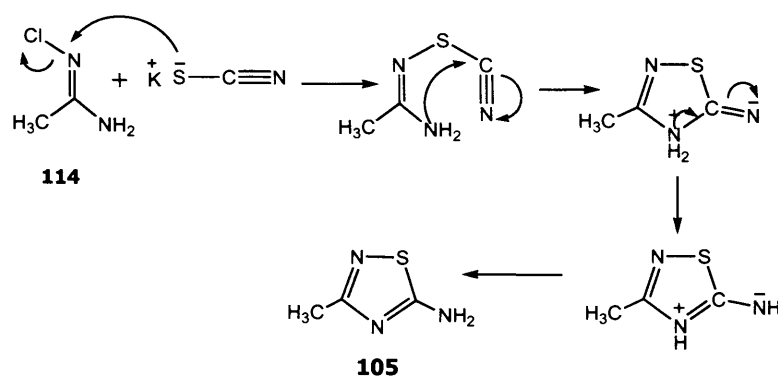


Figure 3.63 Proposed mechanism for the reaction of N-chloroacetamidine and potassium thiocyanate

Meanwhile, in the synthesis proposal of **B16**, the intermediate 4-(4-(methoxycarbonyl)phenylamino)-4-oxobutanoic acid (**104**) (Figure 3.60) was prepared following a reported procedure³⁴². The identity of **104** was assigned by the ^1H NMR spectrum; with a broad carboxylic acid singlet at δ 12.30, broad NH singlet at δ 10.27, two aromatic double doublets at δ 7.90 and 7.72, methyl singlet at δ 3.82, two methylene triplets at δ 2.61 and 2.54. ^{13}C NMR showed 3 carbonyl C=O signals at δ 174, 171 and 166, two

quaternary aromatic carbon at δ 143.2 and 123.0, two aromatic CH at δ 130.3 and 118, methyl carbon at δ 51.5, and two methylene carbons at 31.2 and 28.8.

Also, in the synthesis proposal of methyl 4-((3-methoxy-1,2,4-thiadiazol-5-ylamino)methyl)benzoate (**B27**), the reported synthetic procedure of the intermediate 5-chloro-3-methoxy-1,2,4-thiadiazole (**110**) by Marrano *et al* was adopted³³⁹. A mixture of methylisourea hydrogensulfate (**109**) and perchloromethyl mercaptan (**108**) was reacted with sodium hydroxide solution. The resulting product was obtained as yellow oil (42% yield) and its identity was confirmed by ^1H NMR spectral data; singlet methyl at δ 4.05 (reported 4.04) and ^{13}C NMR; 2 quaternary carbons at δ 173.64, 169.75 (reported 173.20 and 169.47) and methyl carbon at δ 56.46 (reported 56.84).

The product **110** was used directly in the second step to be coupled with methyl 4-(aminomethyl) benzoate (**73**) in the presence of triethylamine, this reaction lead to single product as shown as a clean spot in the TLC. However, the identity of the product was not clear since only the ester methyl group singlet was seen in the ^1H NMR spectra at δ 3.85, together with the methylene doublet at δ 4.32 (d, $J=6$ Hz, 2H, CH_2), the NH triplet at δ 6.66 ($J=6$) and two aromatic double doublets at 7.39 ($J=8.5$ Hz) and 7.92 ($J=8$ Hz).

3.6.7 Conclusion

Although the selected compounds of the virtual library are simple in structure and the synthetic pathways involved should have been uncomplicated using readily available starting materials, the lack of reliable synthetic pathway for the 1,2,4-thiadiazole ring presented a real challenge to prepare the compounds despite of using reported procedures.

Chapter 4

Experimental Chemistry

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 visualised by irradiation with UV light at 254 nm and 366 nm.

Column Chromatography

Glass columns were slurry packed in the appropriate eluent under gravity, with Woelm silica (32-63mm). Samples were applied as a concentrated solution in the same eluent. Fractions containing the product were identified by TLC and combined. The solvent removed in vacuo.

Flash Chromatography

Pre-filled ISOLUTE silica columns were washed with the appropriate solvent, and samples were applied as a concentrated solution in the same eluent. The samples were run on a FlashMaster Personal (K10607, Jones Chromatography Ltd.).

NMR Spectroscopy

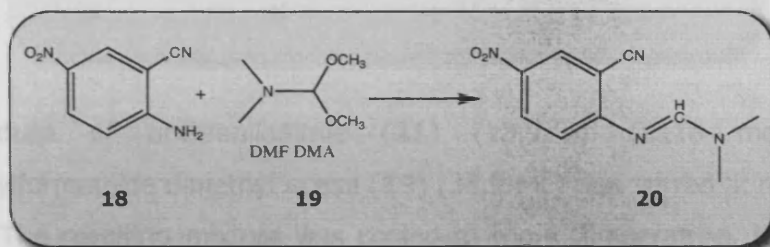
¹H, ¹³C, DEPT NMR spectra were recorded on a Bruker AVANCE 500 spectrometer (500 MHz and 125 MHz respectively) and auto calibrated to the deuterated solvent reference peak. Chemical shifts are given in δ relative to tetramethylsilane (TMS); the coupling constants (J) are given in Hertz (Hz). The spectra were recorded in CDCl₃ or DMSO at room temperature; TMS served as an internal standard ($\delta = 0$ ppm) for ¹H NMR and CDCl₃ was used as an internal standard ($\delta = 77.0$ ppm) for ¹³C NMR.

Chapter 4

Part 1

4.1 Quinazoline based structures

N'-(2-Cyano-4-nitrophenyl)-N,N-dimethylimido formamide



A mixture of 5-nitroanthranilonitrile (**18**) (1.92 g, 0.0118 mol) and dimethylformamide dimethyl acetal DMF DMA (**19**) (3.89 mL) was stirred at reflux temperature for 1.5 h. The resulting mixture was cooled to room temperature and refrigerated. The solid was filtered, washed with several portions of ether, and dried in vacuo (55 °C).

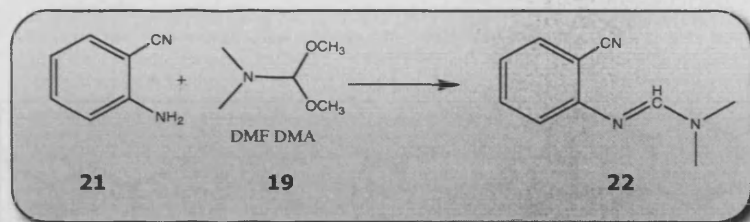
Yield; 2.38 g (98%)

mp; 154 °C (lit. mp 153-155 °C)²⁶⁷.

¹H NMR (DMSO-d₆) of **16**

δ 8.84 (d, *J* = 2.7 Hz, 1H), 8.27 (m, 2H), 7.38 (d, *J* = 9.3 Hz, 1H), 3.17 (s, 3H), 3.06 (s, 3H).

General *N'*-(2-Cyanophenyl)-*N,N*-dimethylimidoforamide



A mixture of anthranilonitrile (**21**) (13.9 g, 0.118 mol) and dimethylformamide dimethyl acetal (**19**) (38.9 mL) was stirred at reflux for 1.5 h. The resulting mixture was cooled to room temperature. Ice water was added and the mixture was extracted with EtOAc (3 x 300 mL). The organic layer was washed with water (3 x 300 mL), dried (Na_2SO_4) and evaporated under reduced pressure (55 °C). It was recrystallised from methanol.

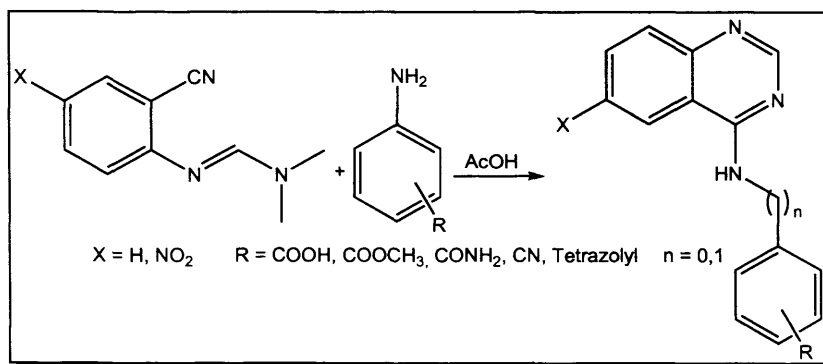
Yield; 18.4 g (90.5%).

mp; 45 °C (lit. mp 38.5 °C)²⁶⁸.

¹H NMR (CDCl_3 - d_1) of **22**

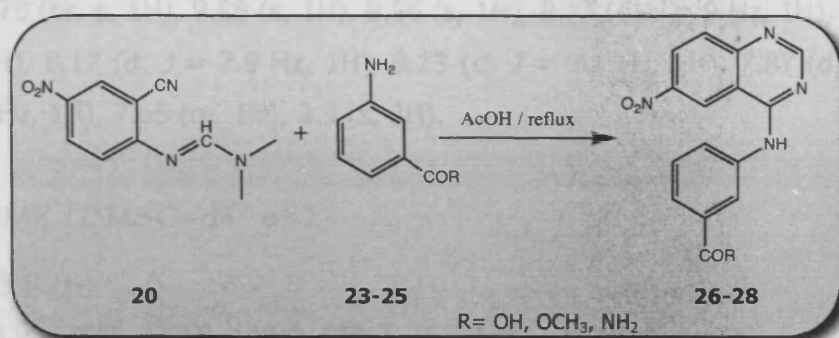
δ 6.69 (s, 1H), 6.62 (d, $J = 7.5$ Hz, 1H), 6.51 (m, 1H), 6.09 (m, 1H), 6.05 (d, $J = 8.15$ Hz, 1H), 2.19 (s, 3H), 2.17 (s, 3H).

General procedure



A mixture of N'-(2-cyanophenyl)-N,N-dimethylimidoforamide (**20**) or N'-(2-cyanophenyl)-N,N-dimethylimidoforamide (**22**) and the appropriate phenyl derivative [3-amino benzoic acid (**23**), methyl 3-aminobenzoate (**24**), 3-aminobenzamide (**25**), 4-amino benzoic acid (**32**), methyl 4-aminobenzoate (**33**) or 4-aminobenzamide (**34**), 4-aminomethylbenzoic acid (**41**), 3-(1H-tetrazol-5-yl)phenylamine (**47**), 4-(1H-tetrazol-5-yl)phenylamine (**49**), 4-aminobenzonitrile (**54**), anthranilic acid (**56**) or anthranilamide (**57**)] in HOAc was stirred at reflux for 1-3 h. The resulting solid was filtered, washed with ether then dried (50 °C) in vacuo.

3-(6-Nitroquinazolin-4-ylamino)benzoic acid derivatives



The general procedure was adopted using N'-(2-cyanophenyl)-N,N-dimethylimidoforamide (**20**) (2.28 g, 0.011 mol) and the appropriate benzoic acid derivative (**23-25**) (0.0121 mol each).

ID	R	M F	M w	Yield %	mp °C
26	OH	C ₁₅ H ₁₀ N ₄ O ₄	310	96.3	335
27	OCH ₃	C ₁₆ H ₁₂ N ₄ O ₄	324	87.5	176
28	NH ₂	C ₁₅ H ₁₁ N ₅ O ₃	309	84.7	258

¹H NMR (DMSO-d₆) of **26**

δ 10.56 (br. s, 1H), 9.7 (s, 1H), 8.74 (s, 1H), 8.56 (d, *J* = 8.95 Hz, 1H), 8.42 (s, 1H), 8.21 (d, *J* = 7.55 Hz, 1H), 7.94 (d, *J* = 9 Hz, 1H), 7.76 (d, *J* = 7.35 Hz, 1H), 7.55 (m, 1H).

¹³C NMR (DMSO-d₆) of **26**

δ 114.4, 131.5, 138.8, 144.6, 152.9, 167.2 (6 C, quaternary C).

δ 120.8, 123.3, 125.1, 126.7, 128.7, 129.5, 157.5 (7 C, aromatic CH).

¹H NMR (DMSO-d₆) of 27

δ 11.70 (br. s, 1H), 9.86 (s, 1H), 8.96 (s, 1H), 8.73 (d, *J* = 9 Hz, 1H), 8.39 (s, 1H), 8.17 (d, *J* = 7.9 Hz, 1H), 8.13 (d, *J* = 9.1 Hz, 1H), 7.87 (d, *J* = 7.65 Hz, 1H), 7.65 (m, 1H), 3.9 (s, 3H).

¹³C NMR (DMSO-d₆) of 27

δ 52.3 (CH₃).

δ 113.9, 130.2, 137.6, 145.4, 159.7, 165.8 (6 C, quaternary C).

δ 121.6, 124.3, 124.9, 126.4, 128.4, 128.6, 129.3, 154.9 (8 C, aromatic CH).

¹H NMR (DMSO-d₆) of 28

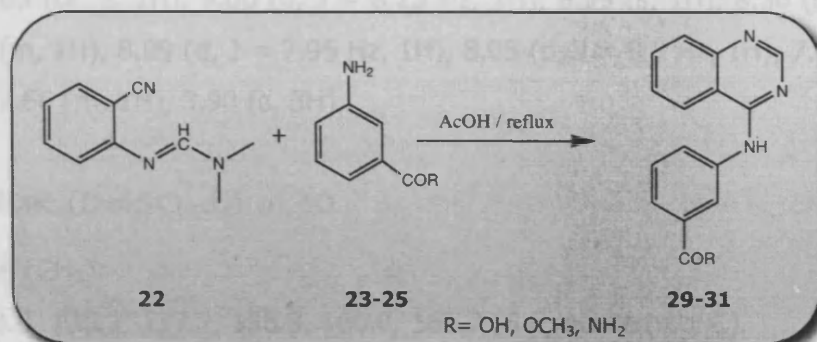
δ 9.73 (s, 1H), 8.74 (s, 1H), 8.57 (d, *J* = 8.95 Hz, 1H), 8.30 (s, 1H), 8.10 (d, *J* = 7.75 Hz, 1H), 8.02 (br. s, 1H), 7.95 (d, *J* = 9.15 Hz, 1H), 7.71 (d, *J* = 7.6 Hz, 1H), 7.51 (m, 1H), 7.40 (br. s, 1H), 4.8 (br. s, 1H).

¹³C NMR (DMSO-d₆) of 28

δ 114.4, 134.8, 138.6, 144.6, 153.1, 158.8, 167.6 (7 C, quaternary C).

δ 120.9, 122.3, 123.2, 125.5, 126.6, 128.4, 129.5, 157.6 (8 C, aromatic CH).

3-(Quinazolin-4-ylamino)benzoic acid derivatives



The general procedure was adopted using N'-(2-cyanophenyl)-N,N-dimethylidamidine (**22**) (1.9 g, 0.011 mol) and the appropriate benzoic acid derivative (**23-25**) (0.0121 mol each).

ID	R	M F	M w	Yield %	mp °C
29 ²⁶⁹	OH	C ₁₅ H ₁₁ N ₃ O ₂	265	97.8	264
30	OCH ₃	C ₁₆ H ₁₃ N ₃ O ₂	279	66.8	226
31	NH ₂	C ₁₅ H ₁₂ N ₄ O	264	92.9	236

¹H NMR (DMSO-d₆) of 29

δ 9.95 (br. s, 1H), 8.63 (s, 1H), 8.57 (d, J = 8.2 Hz, 1H), 8.46 (s, 1H), 8.19 (d, J = 7.8 Hz, 1H), 7.88 (m, 1H), 7.81 (d, J = 8.15 Hz, 1H), 7.72 (d, J = 7.35 Hz, 1H), 7.65 (m, 1H), 7.5 (m, 1H).

¹³C NMR (DMSO-d₆) of 29

δ 115.1, 131.1, 139.4, 149.5, 157.7, 167.3 (6 C, quaternary C).

δ 122.9, 123.3, 124.5, 126.5, 127.7, 128.7, 133.2, 154.3 (8 C, aromatic CH).

¹H NMR (DMSO-d₆) of 30

δ 12.05 (br. s, 1H), 9.06 (d, J = 8.25 Hz, 1H), 8.99 (s, 1H), 8.36 (s, 1H), 8.13 (m, 1H), 8.09 (d, J = 7.95 Hz, 1H), 8.05 (d, J = 8.3 Hz, 1H), 7.90 (m, 2H), 7.66 (m, 1H), 3.90 (s, 3H).

¹³C NMR (DMSO-d₆) of 30

δ 52.4 (CH₃).

δ 113.6, 130.1, 137.2, 138.8, 160.0, 165.7 (6 C, quaternary C).

δ 119.8, 125.1, 125.6, 127.1, 128.6, 129.2, 129.5, 136.3, 151.0 (9 C, aromatic CH).

¹H NMR (DMSO-d₆) of 31

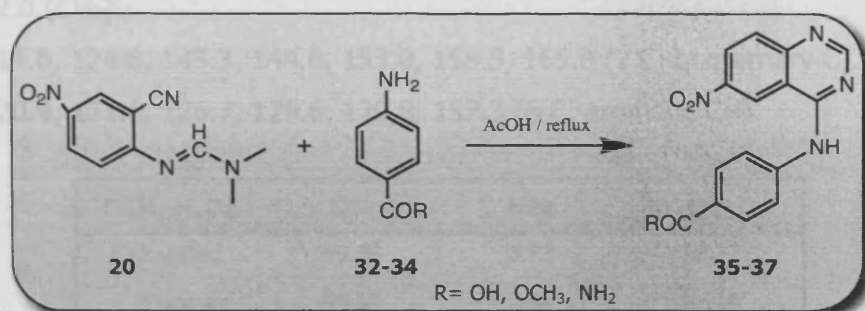
δ 10.00 (br. s, 1H), 8.61 (m, 2H), 8.32 (s, 1H), 8.10 (d, J = 7.9 Hz, 1H), 8.02 (br. s, 1H), 7.88 (m, 1H), 7.81 (d, J = 8.15 Hz, 1H), 7.66 (m, 2H), 7.48 (m, 1H), 7.40 (br. s, 1H).

¹³C NMR (DMSO-d₆) of 31

δ 115.1, 134.7, 139.3, 149.6, 157.8, 167.8 (6C, quaternary C).

δ 121.9, 122.5, 123.0, 125.2, 126.3, 127.8, 128.3, 133.1, 154.4 (9 C, aromatic CH).

4-(6-Nitroquinazolin-4-ylamino)benzoic acid derivatives



The general procedure was adopted using N'-(2-cyano-4-nitrophenyl)-N,N-dimethylimidamide (**20**) (2.28 g, 0.011 mol) and the appropriate benzoic acid derivative (**32-34**) (0.0121 mol each).

ID	R	M F	M W	Yield %	mp °C
35 ²⁷⁰	OH	C ₁₅ H ₁₀ N ₄ O ₄	310	91.4	350
36	OCH ₃	C ₁₆ H ₁₂ N ₄ O ₄	324	82	284
37 ²⁷¹	NH ₂	C ₁₅ H ₁₁ N ₅ O ₃	309	89.3	328

¹H NMR (DMSO-d₆) of **35**

δ 12.60 (br. s, 1H), 10.6 (br. s, 1H), 9.69 (s, 1H), 8.80 (s, 1H), 8.57 (d, J = 9.05 Hz, 1H), 8.06 (d, J = 7.95 Hz, 2H), 8.00 (d, J = 8.3 Hz, 2H), 7.96 (d, J = 9.1 Hz, 1H).

¹³C NMR (DMSO-d₆) of **35**

δ 114.5, 126.0, 142.7, 144.7, 153.0, 158.6, 166.9 (7 C, quaternary C).

δ 120.9, 121.6, 126.7, 129.6, 129.9, 157.4 (6 C, aromatic CH).

¹H NMR (DMSO-d₆) of **36**

δ 10.55 (br. s, 1H), 9.67 (s, 1H), 8.80 (s, 1H), 8.56 (d, J = 8.95 Hz, 1H), 8.08 (d, J = 8.3 Hz, 2H), 8.00 (d, J = 8.35 Hz, 2H), 7.95 (d, J = 9.05 Hz, 1H), 3.85 (s, 3H).

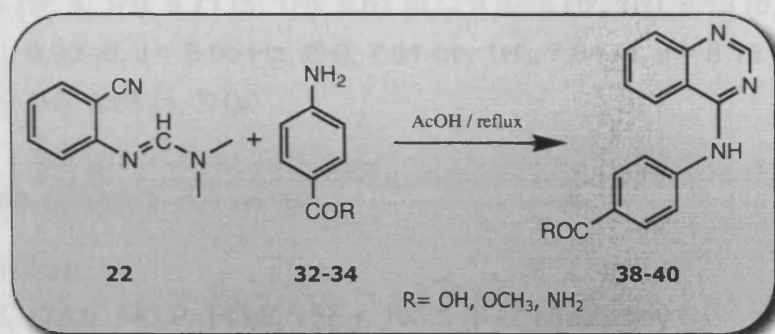
^{13}C NMR (DMSO- d_6) of 36 δ 52.0 (CH₃). δ 114.6, 124.6, 143.3, 144.6, 153.0, 158.5, 165.8 (7 C, quaternary C). δ 120.9, 121.6, 126.7, 129.6, 129.8, 157.3 (6 C, aromatic CH).

C₁₆H₁₂N₄O₄	C%	H%	N%
Calculated	59.26	3.73	17.27
Found	59.22	3.84	17.38

 ^1H NMR (DMSO- d_6) of 37 δ 10.05 (br. s, 1H), 9.67 (s, 1H), 8.77 (s, 1H), 8.56 (d, J = 9.1 Hz, 1H),
7.96 (m, 6H), 7.31 (s, 1H). ^{13}C NMR (DMSO- d_6) of 37 δ 114.5, 129.7, 141.2, 144.6, 153.0, 158.6, 167.4 (7 C, quaternary C). δ 120.8, 121.7, 126.7, 128.0, 129.5, 157.5 (6 C, aromatic CH).

C₁₅H₁₁N₅O₃	C%	H%	N%
Calculated	58.25	3.58	22.63
Found	58.25	3.42	22.64

4-(Quinazolin-4-ylamino)benzoic acid derivatives



The general procedure was adopted using *N'*-(2-cyanophenyl)-*N,N*-dimethylimidamide (**22**) (1.9 g, 0.011 mol) and the appropriate benzoic acid derivative (**32-34**) (0.0121 mol each).

ID	R	M F	M w	Yield %	mp °C
38 ²⁶⁹	OH	C ₁₅ H ₁₁ N ₃ O ₂	265	80.9	309
39	OCH ₃	C ₁₆ H ₁₃ N ₃ O ₂	279	74.3	234
40	NH ₂	C ₁₅ H ₁₂ N ₄ O	264	65.6	235

¹H NMR (DMSO-*d*₆) of **38**

δ 10.05 (br. s, 1H), 8.70 (s, 1H), 8.61 (d, *J* = 8.25 Hz, 1H), 8.09 (d, *J* = 8.05 Hz, 2H), 7.98 (d, *J* = 8.35 Hz, 2H), 7.90 (m, 1H), 7.84 (d, *J* = 8.15 Hz, 1H), 7.68 (m, 1H).

¹³C NMR (DMSO-*d*₆) of **38**

δ 115.3, 125.6, 143.4, 149.8, 157.5, 167.2 (6 C, quaternary C).

δ 120.9, 123.1, 126.5, 127.9, 129.9, 133.2, 154.2 (7 C, aromatic CH).

¹H NMR (DMSO-*d*₆) of 39

δ 10.05 (br. s, 1H), 8.71 (s, 1H), 8.61 (d, J = 8.25 Hz, 1H), 8.13 (d, J = 8.2 Hz, 2H), 8.00 (d, J = 8.55 Hz, 2H), 7.91 (m, 1H), 7.84 (d, J = 8.15 Hz, 1H), 7.69 (m, 1H), 3.85 (s, 3H).

¹³C NMR (DMSO-*d*₆) of 39

δ 51.9 (CH₃).

δ 115.3, 123.8, 143.9, 146.8, 157.4, 165.9 (6 C, quaternary C).

δ 120.9, 123.0, 126.6, 127.9, 129.9, 133.3, 154.1 (7 C, aromatic CH).

¹H NMR (DMSO-*d*₆) of 40

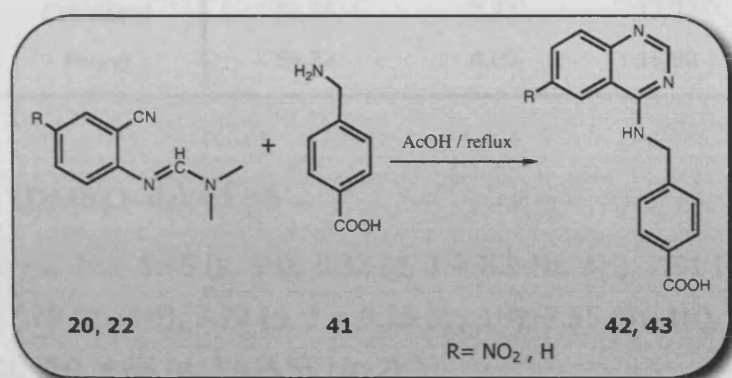
δ 9.95 (br. s, 1H), 8.68 (s, 1H), 8.60 (d, J = 8.25 Hz, 1H), 8.02 (d, J = 7.75 Hz, 2H), 7.91 (m, 4H), 7.83 (d, J = 8.2 Hz, 1H), 7.68 (m, 1H), 7.29 (s, 1H).

¹³C NMR (DMSO-*d*₆) of 40

δ 115.2, 128.9, 142.0, 149.7, 157.5, 167.5 (6 C, quaternary C).

δ 121.0, 123.0, 126.4, 127.9, 128.0, 133.2, 154.3 (7 C, aromatic CH).

4-((6-Nitroquinazolin-4-ylamino)methyl)benzoic acid and 4-((Quinazolin-4-ylamino)methyl)benzoic acid



The general procedure was adopted using *N'*-(2-cyano-4-nitrophenyl)-*N,N*-dimethylimidamide (**20**) or *N'*-(2-cyanophenyl)-*N,N*-dimethylimidamide (**22**) (0.011 mol) and 4-aminomethylbenzoic acid (**41**) (1.827g, 0.0121 mol).

ID	R	M F	M w	Yield %	mp °C
42	NO ₂	C ₁₆ H ₁₂ N ₄ O ₄	324	94.2	318
43 ²⁷²	H	C ₁₆ H ₁₃ N ₃ O ₂	279	99.2	320

¹H NMR (DMSO-*d*₆) of **42**

δ 9.65 (br. s, 1H, NH), 9.45 (s, 1H), 8.60 (s, 1H), 8.50 (d, *J* = 9.1 Hz, 1H), 7.91 (d, *J* = 7.5 Hz, 2H), 7.86 (d, *J* = 9.15 Hz, 1H), 7.49 (d, *J* = 7.75 Hz, 2H), 4.88 (d, *J* = 4.8 Hz, 2H).

¹³C NMR (DMSO-*d*₆) of **42**

δ 43.7 (CH₂)

δ 114.0, 130.2, 143.5, 144.1, 152.8, 160.3, 167.3 (7 C, quaternary C).

δ 120.6, 126.4, 127.3, 129.2, 129.4, 158.2 (6 C, aromatic CH).

C₁₆H₁₂N₄O₄	C%	H%	N%
Calculated	59.26	3.73	17.27
Found	59.22	4.00	16.90

¹H NMR (DMSO-d₆) of 43

δ 8.92 (br. s, 1H), 8.45 (s, 1H), 8.32 (d, J = 8.2 Hz, 1H), 7.91 (d, J = 7.3 Hz, 2H), 7.79 (m, 1H), 7.72 (d, J = 8.25 Hz, 1H), 7.55 (m, 1H), 7.47 (d, J = 7.95 Hz, 2H), 4.85 (d, J = 4.55 Hz, 2H).

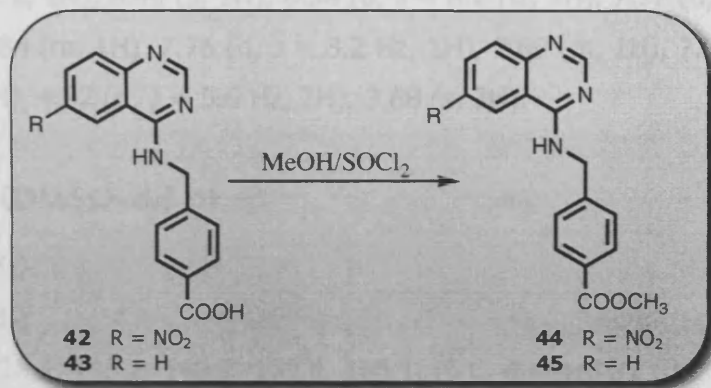
¹³C NMR (DMSO-d₆) of 43

δ 43.4 (CH₂)

δ 114.9, 129.3, 144.6, 149.1, 159.4, 167.2. (6 C, quaternary C).

δ 122.6, 125.8, 127.1, 127.5, 129.4, 132.7, 155.0 (7 C, aromatic CH).

Methyl 4-((quinazolin-4-ylamino)methyl)benzoate ester derivatives



A mixture of 4-((6-nitroquinazolin-4-ylamino)methyl)benzoic acid (**42**) or 4-((quinazolin-4-ylamino)methyl)benzoic acid (**43**) (0.0025 mol), thionyl chloride (4 mL) and dry methanol (70 mL) was stirred for 4 days. The reaction mixture was neutralised using saturated Na₂CO₃ solution, extracted with ethyl acetate (3 x 100 mL). The organic layer was washed with brine (2 x 50 mL), dried over anhydrous Na₂SO₄ and evaporated under reduced pressure.

ID	R	M w	M F	Yield %	mp °C
44	NO ₂	338	C ₁₇ H ₁₄ N ₄ O ₄	87.3	> 300 decmp.
45	H	293	C ₁₇ H ₁₅ N ₃ O ₂	81.4	155

¹H NMR (DMSO-d₆) of **44**

δ 9.62 (t, J = 5.65 Hz, 1H), 9.42 (d, J = 2.3 Hz, 1H), 8.59 (s, 1H), 8.51 (dd, J = 2.3, 9.2 Hz, 1H), 7.93 (d, J = 8.2 Hz, 2H), 7.87 (d, J = 9.15 Hz, 1H), 7.53 (d, J = 8.15 Hz, 2H), 4.89 (d, J = 5.7 Hz, 2H), 3.84 (s, 3H).

¹³C NMR (DMSO-d₆) of **44**

δ 43.7 (CH₂)

δ 52.0 (CH₃)

δ 114.0, 128.3, 144.1, 144.4, 152.8, 160.3, 171.8 (7 C, quaternary C).

δ 120.6, 126.4, 127.5, 129.2, 129.3, 158.2 (6 C, aromatic CH).

¹H NMR (DMSO-d₆) of 45

δ 8.97 (br. s, 1H), 8.49 (s, 1H), 8.36 (d, J = 8.2 Hz, 1H), 7.97 (d, J = 8.15 Hz 2H), 7.84 (m, 1H), 7.76 (d, J = 8.2 Hz, 1H), 7.60 (m, 1H), 7.54 (d, J = 8.05 Hz 2H), 4.92 (d, J = 5.6 Hz, 2H), 3.88 (s, 3H).

¹³C NMR (DMSO-d₆) of 45

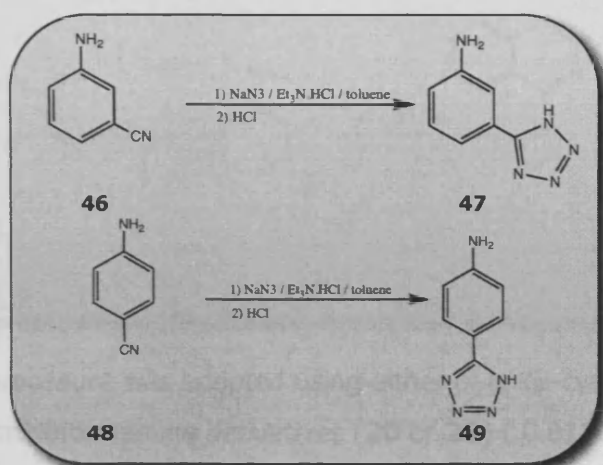
δ 43.3 (CH₂)

δ 52.0 (CH₃)

δ 114.9, 128.1, 145.2, 149.2, 159.4, 166.1. (6 C, quaternary C).

δ 122.6, 125.8, 127.2, 127.6, 129.2, 132.7, 155.0 (7 C, aromatic CH).

3-(1H-tetrazol-5-yl)benzenamine and 4-(1H-tetrazol-5-yl)benzenamine



The mixture of aminobenzonitrile (**46** or **48**) (1.18 g, 10 mmol), NaN₃ (0.85 g, 0.013 mol) and triethylamine hydrochloride salt (1.79 g, 0.013 mol) in toluene (100 mL) was heated to reflux for 28 h with stirring. After cooling, the product was extracted with water (3 x 50 mL). To the aqueous layer, 36 % HCl was added dropwise to salt out the produced tetrazole. After filtration, the solid was dried under reduced pressure²⁷³.

ID	Yield %	mp °C	ID	Yield %	m.p. °C
47	80.9	199	49	74.3	269

¹H NMR (DMSO-d₆) of **47**

δ 7.28 (m, 1H), 7.15 (d, J=7.55 Hz, 1H), 6.99 (m, 1H), 6.45 (m, 1H), 4.95 (br. s, 2H).

¹³C NMR (DMSO-d₆) of **47**

δ 133.1, 148.3 (2C, quaternaryC) 111.8, 112.7, 114.1, 128.5. (4 C, aromatic CH).

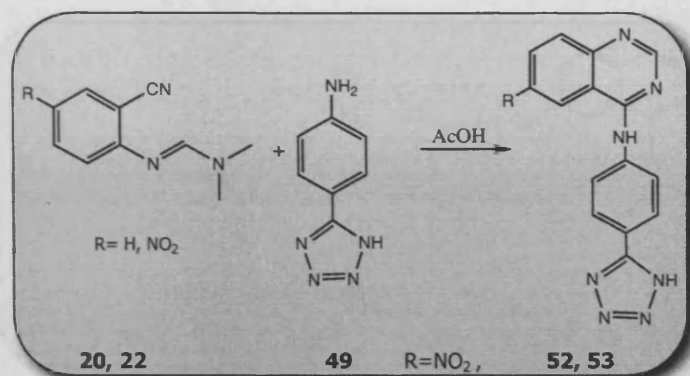
¹H NMR (DMSO-d₆) of **49**

δ 7.71 (d, J = 8.45 Hz, 2H), 6.69 (d, J = 8.55 Hz, 2H), 3.95 (br. s, 2H).

¹³C NMR (DMSO-d₆) of **49**

δ 110.7, 151.5, 155.3. (3C, quaternary C) 113.7, 128.2. (2C, aromatic CH).

4-Amino(4-(1H-tetrazol-5-yl)phenyl)quinazoline derivatives



The general procedure was adopted using *N'*-(2-cyanophenyl)-*N,N*-dimethylimidoformamide derivative (**20** or **22**) (0.011 mol) and 4-(1H-tetrazol-5-yl)phenylamine (**49**) (1.827g, 0.0121 mol).

ID	R	M w	M F	Yield %	mp °C
52	NO ₂	334	C ₁₅ H ₁₀ N ₈ O ₂	73.3	290
53	H	289	C ₁₅ H ₁₁ N ₇	59.9	279

¹H NMR (DMSO-*d*₆) of **52**

δ 12.15 (br. s, 1H, NH), 10.67 (br. s, 1H, NH), 9.73 (d, *J* = 2.25 Hz, 1H), 8.83 (s, 1H), 8.6 (dd, *J* = 2.4, 9.2 Hz, 1H), 8.15 (d, *J* = 8.7 Hz, 2H), 8.11 (d, *J* = 8.8 Hz, 2H), 7.98 (d, *J* = 9.15 Hz, 1H).

¹³C NMR (DMSO-*d*₆) of **52**

δ 121.1, 122.8, 126.9, 127.5, 129.3, 157.3 (6C, aromatic CH).

δ 115.2, 142.6, 144.7, 159.7, 167.8, 171.7 (6C, quaternary C).

¹H NMR (DMSO-*d*₆) of **53**

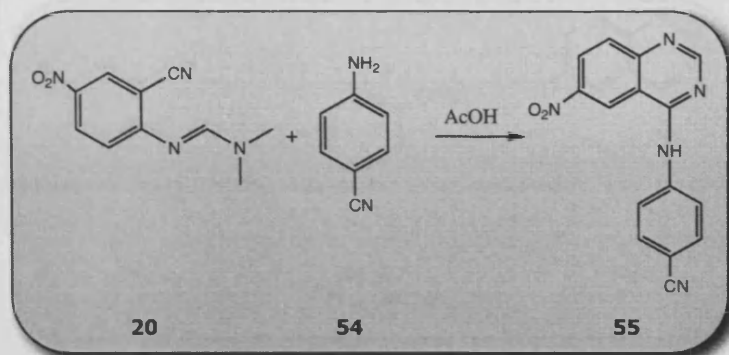
δ 10.3 (br. s, 1H, NH), 10.1 (br. s, 1H, NH), 9.07 (s, 1H), 8.7 (m, 2H), 8.39 (m, 2H), 8.11 (d, *J* = 8.3 Hz, 1H), 7.74 (m, 1H), 7.55 (m, 2H), 7.43 (m, 1H).

¹³C NMR (DMSO-*d*₆) of **53**

δ 122.1, 123.0, 126.5, 127.3, 127.8, 133.3, 154.2 (6 C, aromatic CH).

δ 115.2, 118.8, 142.0, 149.7, 157.5, 155.02 (6 C, quaternary C).

4-(6-Nitroquinazolin-4-ylamino)benzonitrile



The general procedure was adopted using N'-(2-cyano-4-nitrophenyl)-N,N-dimethylimidoforamide (**20**) (2.28 g, 0.011 mol) and 4-aminobenzonitrile (**54**) (1.427g, 0.0121 mol).

Yield; 1.96 g (64.3%)

mp; 275 °C. (C₁₅H₉N₅O₂).

¹H NMR (DMSO-d₆) of **55**

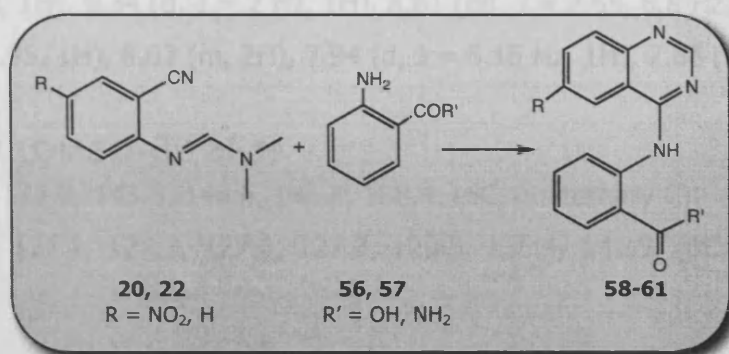
δ 10.4 (br. s, 1H), 9.65 (s, 1H), 8.80 (d, J = 1.05 Hz, 1H), 8.55 (m, 1H), 8.12 (d, J=7.65 Hz, 2H), 7.96 (d, J=9.15 Hz, 1H), 7.85 (d, J=7.6 Hz, 2H).

¹³C NMR (DMSO-d₆) of **55**

δ 114.7, 119.1, 143.4, 144.7, 153.0, 158.5, 172.2 (7 C, quaternary C).

δ 120.9, 122.2, 126.8, 129.5, 132.8, 157.1 (6 C, aromatic CH).

2-(Quinazolin-4-ylamino)benzoic acid derivatives



The general procedure was adopted using N'-(2-cyanophenyl)-N,N-dimethylimidamide derivatives (**20** or **22**) (0.011 mol) and the appropriate anthranilic acid derivative [anthranilic acid (**56**) or anthranilamide (**57**)] (0.0121 mol each).

ID	R	R'	M w	M F	Yield %	mp
58	NO ₂	OH	310	C ₁₅ H ₁₀ N ₄ O ₄	73.3	>360 decmp.
59	NO ₂	NH ₂	309	C ₁₅ H ₁₁ N ₅ O ₃	57.2	239
60 ²⁷⁴	H	OH	265	C ₁₅ H ₁₁ N ₃ O ₂	59.9	180
61 ²⁷⁵	H	NH ₂	264	C ₁₅ H ₁₂ N ₄ O	62.4	179

¹H NMR (DMSO-d₆) of **58**

δ 12.58 (br. s, 1H), 8.75 (s, 1H), 8.3 (m, 2H), 8.15 (d, J = 7.55 Hz, 1H), 8.06 (d, J = 9, 1H), 7.81 (m, 2H), 7.54 (m, 1H), 6.93 (d, J = 9.15 Hz, 1H).

¹³C NMR (DMSO-d₆) of **58**

δ 111.2, 120.8, 135.3, 147.5, 152.5, 154.6. (6 C, quaternary C).

δ 116.1, 125.7, 126.8, 127.1, 134.5. (5 C, aromatic CH).

¹H NMR (DMSO-*d*₆) of 59

δ 9.39 (s, 1H), 9.34 (d, J = 2 Hz, 1H), 8.61 (dd, J = 2.65, 8.8 Hz, 1H), 8.34 (d, J = 7.95, 1H), 8.02 (m, 2H), 7.94 (d, J = 8.15 Hz, 1H), 7.66 (m, 1H).

¹³C NMR (DMSO-*d*₆) of 59

δ 119.0, 122.2, 143.5, 146.6, 146.8, 158.4. (6C, quaternary C).

δ 120.7, 127.1, 127.3, 127.6, 127.8, 129.5, 136.4, 141.5. (8C, aromatic CH).

¹H NMR (DMSO-*d*₆) of 60

δ 12.45 (br. s, 1H), 9.31 (s, 1H), 8.74 (dd, J = 1.3, 8.05 Hz, 1H), 8.34 (dd, J = 1.35, 7.95 Hz, 1H), 8.00 (ddd, J = 1.55, 7.15, 8.5 Hz, 1H), 7.93 (ddd, J = 1.45, 7.15, 8.5 Hz, 1H), 7.95 (d, 7.95, 2H), 7.75 (ddd, J = 1.25, 6.95, 8.2 Hz, 1H), 7.63 (ddd, J = 1, 7.05, 8.05 Hz, 1H).

¹³C NMR (DMSO-*d*₆) of 60

δ 118.7, 121.2, 142.8, 144.4, 147.1, 158.7. (6C, quaternary C).

δ 125.4, 126.5, 126.9, 127.3, 127.7, 128.8, 133.8, 135.9, 138.1 (9C, aromatic CH).

¹H NMR (DMSO-*d*₆) of 61

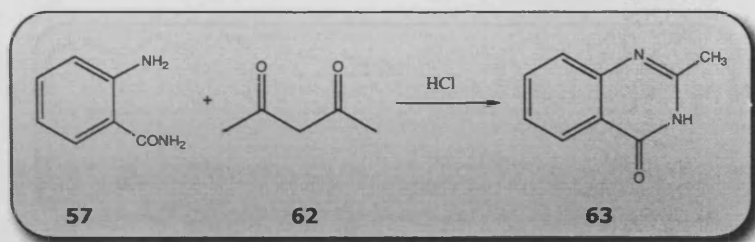
δ 9.28 (s, 1H), 8.71 (d, J = 7.95, 1H), 8.31 (d, J = 7.95 Hz, 1H), 7.97 (ddd, J = 1.3, 6.95, 8.25 Hz, 1H), 7.9 (ddd, J = 1.2, 7.05, 8.25 Hz, 1H), 7.83 (d, J = 8, 2H), 7.73 (m, 1H), 7.60 (m, 1H).

¹³C NMR (DMSO-*d*₆) of 61

δ 118.7, 121.2, 142.9, 144.4, 147.1, 158.7. (6 C, quaternary C).

δ 125.4, 126.4, 126.9, 127.3, 127.7, 128.9, 133.8, 135.9, 138.1. (9 C, aromatic CH).

2-Methylquinazolin-4(3H)-one



(1.36 g, 0.01 mol) of anthranilamide (**57**) was dissolved in (3 g, 0.03 mol) of acetylacetone (**62**), and one drop of concentrated hydrochloric acid was added on stirring. A white solid was developed after 1 h reflux ²⁷⁶. The reaction mixture was cooled and excess diketone was evaporated, the solid was washed with ether and recrystallised from ethanol.

Yield; 1.4 g (88.5%)

mp; 240°C. (lit 237-239) ²⁷⁶.

¹H NMR (CDCl₃-d₁) of **63**; (reported) ²⁷⁷

δ 11.5 (br. s, 1H), 8.29 (ddd, *J*=7.61, 1.47, 0.59 Hz, 1H), 7.77 (dt, *J*=8.49, 7.03, 1.47 Hz, 1H), 7.68 (ddd, *J*=8.79, 1.17, 0.59 Hz, 1H), 7.48 (dt, *J*=8.2, 7.03, 1.47 Hz, 1H), 2.59 (s, 3H).

¹H NMR (DMSO-d₆) of **63**

δ 12.20 (br. s, 1H), 8.07 (dd, *J* = 1.4, 7.95 Hz, 1H), 7.77 (ddd, *J* = 1.55, 6.85, 8.4 Hz, 1H), 7.57 (d, *J* = 8.05 Hz, 1H), 7.45 (ddd, *J* = 1, 7, 8 Hz, 1H), 2.36 (s, 3H).

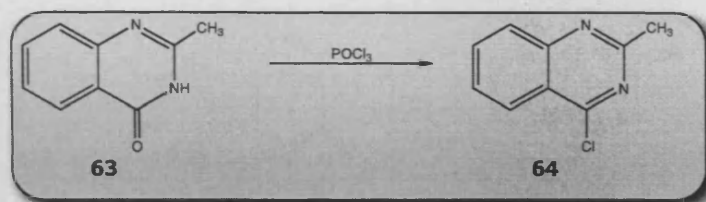
¹³C NMR (CDCl₃-d₁) of **63**

δ 22 (CH₃).

δ 164.2, 154.37, 149.8, 121.22 (4 C, quaternary C).

δ 134.5, 127.3, 126.5, 126.2 (4 C, aromatic CH).

Methyl 4-Chloro-2-methylquinazoline



A solution of 2-Methylquinazolin-4(3H)-one (**63**) (4.6 g, 0,0375 mol) and dimethylaminopyridine (DMAP) (0.182 g, 0,0561 mol) in dry toluene (100 mL) was refluxed for 5 min. Phosphorus oxychloride (2.5 mL) was added by syringe, and the resultant mixture was refluxed for 3 h (a clear orange colour developed after 1 h at reflux and turned deep red after 1.5 h). the solution was cooled to room temperature and filtered. The insoluble precipitate was washed with dry toluene (40 mL). the combined filtrate was rapidly washed sequentially with ice water (100 mL), ice-cooled 20% sodium hydroxide (2x100mL), ice water (100 mL), and saturated sodium chloride (100 mL). the organic layer was immediately washed with a HCl solution (1M, 100 mL) and water (100 mL) and dried over sodium sulfate. The solvent was removed in vacuo at 30 °C.

Yield; 2.3 g (35.1%)

mp; 85 (lit 85-86) °C²⁷⁸.

¹H NMR (CDCl₃-d₁) of **64**

δ 8.18 (d, J = 8.3 Hz, 1H), 7.95 (d, J = 9.5, 1H), 7.89 (m, 1H), 7.62 (m, 1H), 2.82 (s, 3H, CH₃).

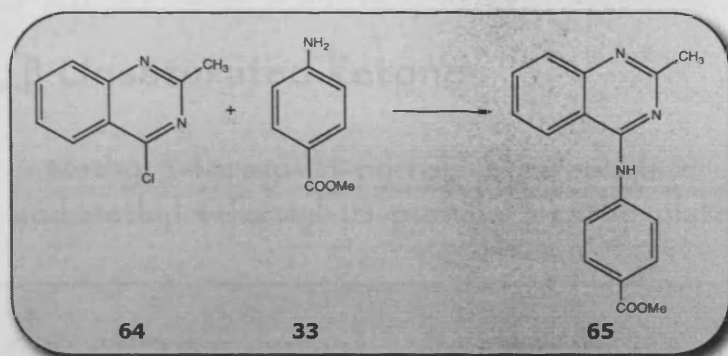
¹³C NMR (CDCl₃-d₁) of **64**

δ 26.1 (CH₃).

δ 163.6, 162.2, 151.5, 121.8 (4C, quaternary C).

δ 134.9, 128.1, 128, 125.7 (4C, aromatic CH).

Methyl 4-(2-methylquinazolin-4-ylamino)benzoate



A mixture of 4-chloro-2-methylquinazoline (**64**) (0.25 g, 0.0014 mol), methyl 4-aminobenzoate (**31**) (0.85 g, 0.0056 mol), triethylamine (0.99 g, 0.0098 mol) and isopropanol (50 mL) was refluxed for 12 days. After cooling to room temperature the mixture was extracted with ethyl acetate (3x50 mL), dried with anhydrous MgSO₄ and evaporated under vacuum. The residue was purified by column chromatography with ethyl acetate and hexane (70:30).

Yield; (0.2 g, 48.7%). mp ; 176 °C.

¹H NMR (DMSO-d₆) of **65**

δ 9.96 (br. s, 1H), 8.59 (d, J = 8.3 Hz, 1H), 8.21 (d, J = 8.65 Hz, 2H), 8.02 (d, J = 8.7 Hz, 2H), 7.88 (m, 1H), 7.78 (d, 1H), 7.63 (m, 1H), 3.89 (s, 3H), 2.63 (s, 3H).

¹³C NMR (DMSO-d₆) of **65**

δ 26.1 (CH₃)

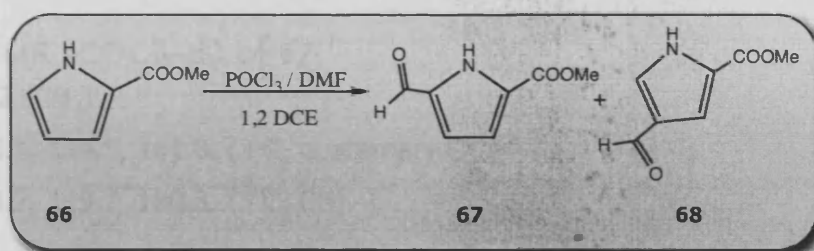
δ 51.8 (CH₃)

δ 113.3, 123.5, 144.2, 150.4, 157.3, 162.6, 165.9. (7 C, quaternary C).

δ 120.6, 122.9, 125.5, 127.3, 129.8, 133.1. (6 C, aromatic CH).

Chapter 4

Part 2

4.2 α , β Unsaturated ketonesMethyl 5-formyl-1H-pyrrole-2-carboxylate
and Methyl 4-formyl-1H-pyrrole-2-carboxylate

To RBF flushed with N₂ was added anhydrous DMF (7 mL, 0.09 mol) it was then cooled to 5-10 °C. To the cooled DMF was added POCl₃ (7.8 mL, 0.084 mol) dropwise over a couple of minutes. Dry 1,2-DCE (25 mL) was then added and the orange solution was cooled to 0-5°C during the addition of methyl pyrrole-2-carboxylate (**66**) (9.5 g, 0.076 mol) in dry 1,2-DCE (50 mL). The mixture was then heated to reflux for 15 min. the reaction was cooled to room temperature, treated with a mixture of ethyl acetate (60 mL) and water (75 mL), poured into saturated NaHCO₃ (350 mL), and separated. The aqueous layer was washed three times with ether, and the combined organic extracts were washed twice with aqueous saturated Na₂CO₃, dried over Na₂SO₄, and evaporated under vacuum. The resulting solid was chromatographed on a silica gel flash column, eluting with cyclohexane/ethyl acetate (70:30) to afford two compounds in a 2:1 ratio. The first to elute was Methyl 5-formyl-1H-pyrrole-2-carboxylate (**67**) (C₇H₇NO₃), which was recrystallised from cyclohexane. The second compound to elute was Methyl 4-formyl-1H-pyrrole-2-carboxylate (**68**) (C₇H₇NO₃), which was recrystallised from cyclohexane²⁸⁴.

Methyl 5-formyl-1H-pyrrole-2-carboxylate (67);

Yield; (7.6 g, 65.4%) mp; 93°C

¹H NMR (CDCl₃-d₁) of 67:

δ 9.92 (br. s, 1H), 9.70 (s, 1H), 6.97 (d, J = 4.2 Hz, 1H), 6.95 (d, J = 4.15 Hz, 1H), 3.94 (s, 3H).

¹³C NMR (CDCl₃-d₁) of 67:

δ 52.2 (CH₃)

δ 128.1, 134.5, 161.0. (3 C, quaternary C).

δ 115.7, 119.7, 180.3. (3 C, CH).

Methyl 4-formyl-1H-pyrrole-2-carboxylate (68);

Yield; (2.9 g, 25.1%) mp; 93 °C

¹H NMR (CDCl₃-d₁) of 68:

δ 9.81 (br. s, 1H), 9.78 (s, 1H), 7.51 (d, J = 1.25 Hz, 1H), 7.25 (d, J = 1.45 Hz, 1H), 3.83 (s, 3H).

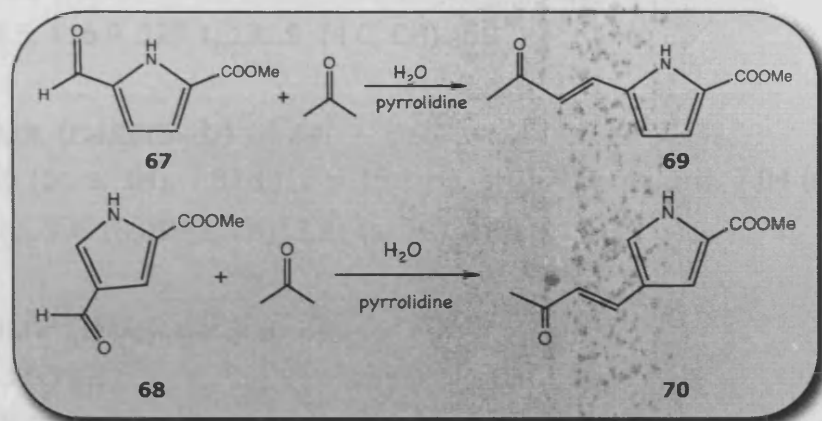
¹³C NMR (CDCl₃-d₁) of 68:

δ 52.0 (CH₃)

δ 124.8, 127.7, 161.2. (3 C, quaternary C).

δ 114.2, 128.5, 185.6. (3 C, CH).

(E)-Methyl 5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylate and
(E)-methyl 4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylate



To a stirring mixture of either of the methyl formyl-1H-pyrrole-2-carboxylate derivatives **67** or **68** (1.53 g, 0.01 mol), acetone (3.7 mL, 0.05 mol), water (50 mL) and amine (30 mol %) were added. The reaction was stirred overnight at room temperature and on completion as monitored by TLC was extracted with (3 x 50 mL) dichloromethane. The organic layer was washed with water (2 x 50 mL) and dried over anhydrous Na_2SO_4 and evaporated to obtain crude product. Column chromatography of the crude on silica gel using mixture of DCM/MeOH (99:1) as eluent gave pure products²⁸⁴. ($\text{C}_{10}\text{H}_{11}\text{NO}_3$) M.w. 193.

ID	Yield %	mp °C
69	43.9	155
70	59.3	168

^1H NMR ($\text{DMSO}-d_6$) of **69**:

δ 10.06 (br. s, 1H), 7.37 (d, $J = 16.3$ Hz, 1H), 6.85 (d, $J = 3.95$ Hz, 1H), 6.58 (d, $J = 16.25$ Hz, 1H), 6.51 (d, $J = 3.9$ Hz, 1H), 3.86 (s, 3H), 2.27 (s, 3H).

^{13}C NMR ($\text{DMSO}-d_6$) of **69**:

δ 27.5 (CH_3).

δ 52.1 (CH₃).

δ 125.7, 132.7, 161.6, 197.8. (4 C, quaternary C).

δ 114.5, 116.9, 125.1, 131.9. (4 C, CH).

¹H NMR (DMSO-d₆) of 69:

δ 9.56 (br. s, 1H), 7.37 (d, J = 16.1 Hz, 1H), 7.14 (m, 1H), 7.04 (s, 1H),

6.43 (d, J = 16.05 Hz, 1H), 3.81 (s, 3H), 2.25 (s, 3H).

¹³C NMR (DMSO-d₆) of 69:

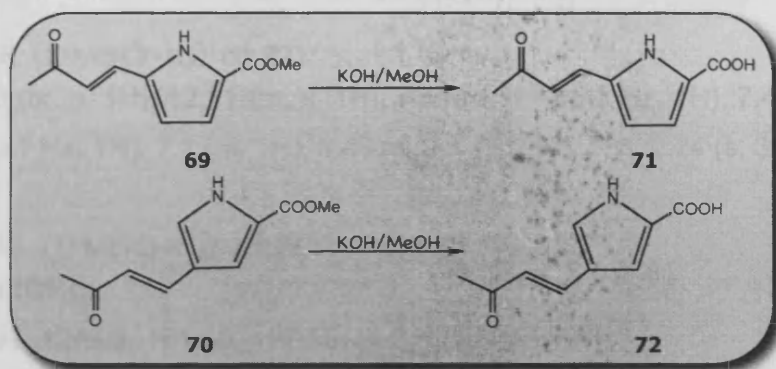
δ 27.3 (CH₃).

δ 52.8 (CH₃).

δ 122.2, 124.5, 161.2, 198.4. (4 C, quaternary C).

δ 113.0, 125.0, 136.5, 185.5. (4 C, CH).

(E)-5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylic acid and
(E)-methyl 4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxylic acid



A mixture of the corresponding Methyl pyrrole-2-carboxylate ester **69** or **70** (0.01 mol), KOH 1N solution (9 mL), MeOH (30 mL) and THF (30 mL) was refluxed 15 h. The resulting solution was cooled to room temperature, water was added and pH was adjusted to pH 2 using 1N HCl solution. This was extracted with ethyl acetate (3x 50mL). the organic layer was washed with brine solution and dried over dry Na₂SO₄ then evaporated under vacuum. The crude product was purified on silica gel using increasing proportions of MeOH in DCM. (C₉H₉NO₃) M.w. 179.

ID	Yield %	mp °C
71	68.4	210 decomp.
72	73.9	197 decomp.

¹H NMR (DMSO-d₆) of 71:

δ 12.62 (br. s, 1H), 12.21 (br. s, 1H), 7.43 (d, J = 16.35 Hz, 1H), 6.81 (d, J = 16.35 Hz, 1H), 6.78 (m, 1H), 6.69 (dd, J = 2.35, 3.75 Hz, 1H), 2.25 (s, 3H).

¹³C NMR (DMSO-d₆) of 71:

δ 27.3 (CH₃).

δ 126.6, 132.6, 161.5, 197.5. (4 C, quaternary C).

δ 113.7, 116.0, 124.9, 132.4. (4 C, CH).

$^1\text{H NMR}$ ($\text{DMSO}-d_6$) of **72**:

δ 12.73 (br. s, 1H), 12.11 (br. s, 1H), 7.51 (d, $J = 16.2$ Hz, 1H), 7.44 (dd, $J = 1.6, 3.1$ Hz, 1H), 7.1 (m, 1H), 6.49 (d, $J = 16.2$ Hz, 1H), 2.24 (s, 3H).

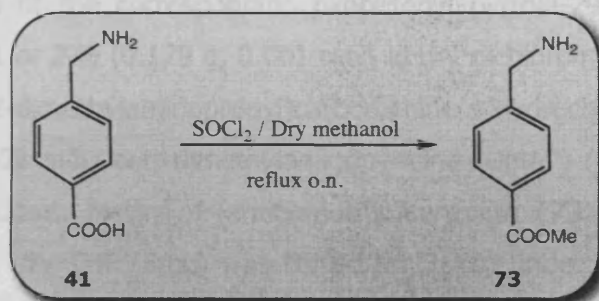
$^{13}\text{C NMR}$ ($\text{DMSO}-d_6$) of **72**:

δ 26.86 (CH_3).

δ 120.94, 124.96, 161.56, 197.55. (4 C, quaternary C).

δ 112.82, 123.80, 126.69, 137.65. (4 C, CH).

Methyl 4-(aminomethyl)benzoate



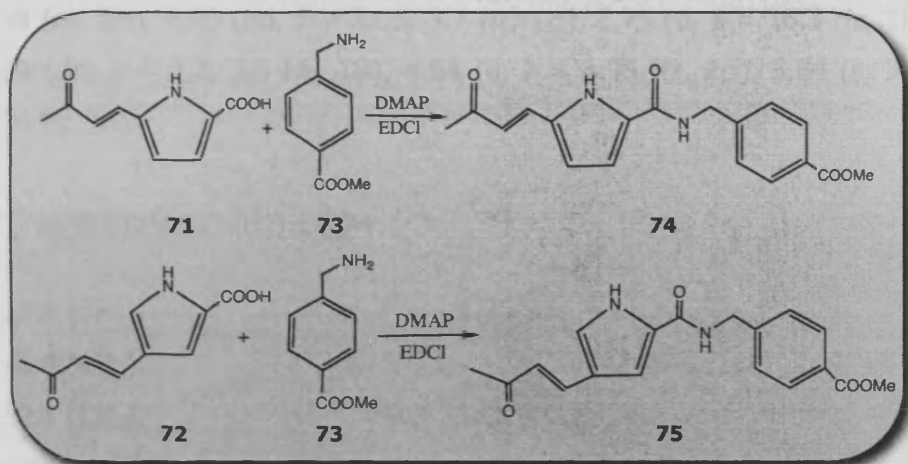
4-(Aminomethyl)benzoic acid (**41**) (5.0 g) was stirred with MeOH (100 mL) and SOCl_2 (5 mL) under reflux temperature overnight. Cooling to room temperature followed by evaporation gave (5.2 g, 96 %) of Methyl 4-(aminomethyl)benzoate hydrochloride (**73**).

mp 245 °C (lit 235-238 °C)²⁸⁵.

$^1\text{H NMR}$ ($\text{DMSO}-d_6$) of **61**:

δ 8.64 (br. s, 3H), 7.99 (d, $J = 8.35$ Hz, 2H), 7.66 (d, $J = 8.35$ Hz, 2H), 4.12 (s, 2H), 3.88 (s, 3H).

(E)-methyl 4-((5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoate and (E)-methyl 4-((4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoate



To a solution of the corresponding oxobutenyl pyrrole-2-carboxylic acid derivative (**71** or **72**) (0.179 g, 0.001 mol) in dry dichloromethane (5 mL), 1-ethyl 3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI) (0.383 g, 0.002 mol) and dimethylaminopyridine (DMAP) (0.244 g, 0.002 mol) were added. Methyl 4-(aminomethyl)benzoate (**73**)²⁸⁶ (0.165 g, 0.001 mol) in dry THF (5 mL) was stirred for 3 min under N₂ atmosphere then was added to the above mentioned DCM mixture. The resulting mixture was stirred at RT for 48 h under N₂. The resulting mixture was evaporated under reduced pressure. The crude product was chromatographed using ethyl acetate as eluent. (C₁₈H₁₈N₂O₄) M.w. 326.

ID	Yield %	mp °C
74	24.4	208
75	28.9	218

¹H NMR (DMSO-*d*₆) of 74:

δ 12.09 (br. s, 1H), 8.90 (t, J = 6.05 Hz, 1H), 7.94 (d, J = 8.3 Hz, 2H), 7.44 (m, 3H), 6.93 (dd, J = 2.05, 3.7 Hz, 1H), 6.75 (d, J = 16.3 Hz, 1H), 6.69 (dd, J = 2.2, 3.6 Hz, 1H), 4.54 (d, J = 5.95 Hz, 2H), 3.84 (s, 3H), 2.24 (s, 3H).

¹³C NMR (DMSO-*d*₆) of 74:

δ 27.2 (CH₃).

δ 41.8 (CH₂).

δ 52.1 (CH₃).

δ 128.1, 129.8, 131.1, 145.4, 160.1, 166.1, 197.5 (7 C, quaternary C).

δ 112.0, 113.8, 123.9, 127.4, 129.3, 132.8 (6 C, CH).

¹H NMR (DMSO-*d*₆) of 75:

δ 12.00 (br. s, 1H), 8.83 (t, J = 6 Hz, 1H), 7.98 (d, 8.3 Hz, 2H), 7.58 (d, J = 16.1 Hz, 1H), 7.49 (d, J = 8.35 Hz, 2H), 7.42 (dd, J = 1.6, 2.7 Hz, 1H), 7.22 (br. s, 1H), 6.37 (d, J = 16.1 Hz, 1H), 4.54 (d, J = 6.05 Hz, 2H), 3.89 (s, 3H), 2.29 (s, 3H).

¹³C NMR (DMSO-*d*₆) of 75:

δ 26.7 (CH₃).

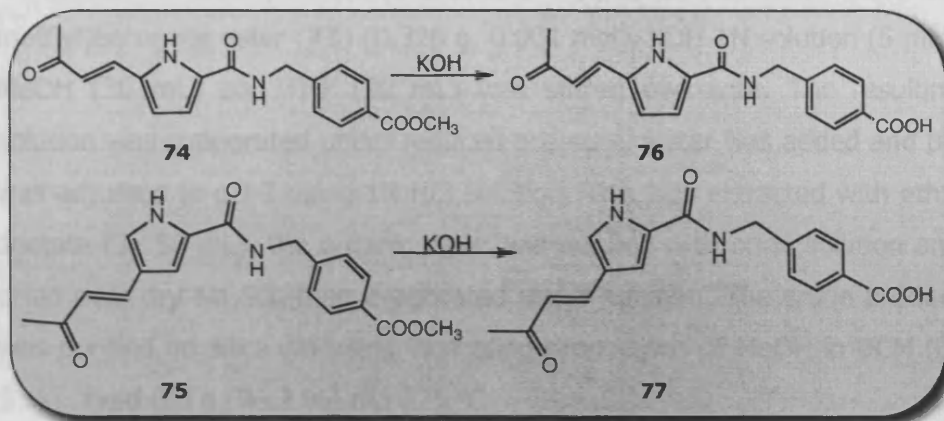
δ 41.8 (CH₂).

δ 52.0 (CH₃).

δ 120.5, 128.0, 128.1, 145.4, 160.3, 166.1, 197.4 (7 C, quaternary C).

δ 108.0, 123.2, 125.6, 127.3, 129.1, 138.2 (6 C, CH).

(E)-4-((5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoic acid and (E)-4-((4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoic acid



A mixture of methyl 4-((5-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl) benzoate ester (**74**) (0.326 g, 0.001 mol), KOH 1N solution (5 mL), MeOH (30 mL) and THF (30 mL) was stirred at 45 °C for 48 h. The resulting solution was evaporated under reduced pressure, water was added and pH was adjusted to pH 2 using 1N HCl solution. This was extracted with ethyl acetate (3 x 50 mL). the organic layer was washed with brine solution and dried over dry Na₂SO₄ then evaporated under vacuum. The crude product was purified on silica gel using increasing proportions of MeOH in DCM (0-3%). Yield 0.294 g (94.2%) mp 274 °C.

¹H NMR (DMSO-*d*₆) of **76**:

δ 12.75 (br. s, 1H), 12.08 (br. s, 1H), 8.87 (t, J = 6.55 Hz, 1H), 7.92 (d, J = 8.25 Hz, 2H), 7.44 (m, 3H), 6.93 (m, 1H), 6.75 (d, J = 16.3 Hz, 1H), 6.68 (m, 1H), 4.53 (d, J = 6 Hz, 2H), 2.24 (s, 3H).

¹³C NMR (DMSO-*d*₆) of **76**:

δ 27.2 (CH₃).

δ 41.8 (CH₂).

δ 129.3, 129.8, 131.1, 144.8, 160.1, 167.1, 197.4 (7 C, quaternary C).

δ 112.0, 113.7, 123.9, 127.1, 129.4, 132.8 (6 C, CH).

A mixture of methyl 4-((4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoate ester (**75**) (0.326 g, 0.001 mol), KOH 1N solution (5 mL), MeOH (30 mL) and THF (30 mL) was stirred overnight. The resulting solution was evaporated under reduced pressure, water was added and pH was adjusted to pH 2 using 1N HCl solution. This was extracted with ethyl acetate (3x 50 mL). the organic layer was washed with brine solution and dried over dry Na₂SO₄ then evaporated under vacuum. The crude product was purified on silica gel using increasing proportions of MeOH in DCM (0-3 %). Yield 0.3 g (96.2 %) mp 225 °C

¹H NMR (DMSO-d₆) of **77**:

δ 12.78 (br. s, 1H), 11.95 (br. s, 1H), 8.77 (t, J = 6.05 Hz, 1H), 7.91 (d, J = 8.25 Hz, 2H), 7.55 (d, J = 16.15 Hz, 1H), 7.43 (d, J = 8.3 Hz, 2H), 7.37 (dd, J = 1.4, 2.95 Hz, 1H), 7.18 (t, J = 1.85 Hz, 1H), 6.34 (d, J = 16.1 Hz, 1H), 4.52 (d, J = 6 Hz, 2H), 2.25 (s, 3H).

¹³C NMR (DMSO-d₆) of **77**:

δ 26.8 (CH₃).

δ 41.8 (CH₂).

δ 120.5, 128.1, 129.3, 144.9, 160.3, 167.2, 197.4 (7 C, quaternary C).

δ 108.0, 123.2, 125.6, 127.2, 129.3, 138.2 (6 C, CH).

Methyl 4-((4-iodobenzamido)methyl)benzoate of 80:

^1H NMR (DMSO-d_6) of 80:

δ 7.92 (d, $J = 8.35$ Hz, 2H), 7.71 (d, $J = 8.6$ Hz, 2H), 7.45 (d, $J = 8.5$ Hz, 2H), 7.31 (d, $J = 8.4$ Hz, 2H), 6.48 (br. s, 1H), 4.6 (d, $J = 5.9$ Hz, 2H), 3.84 (s, 3H).

^{13}C NMR (DMSO-d_6) of 80:

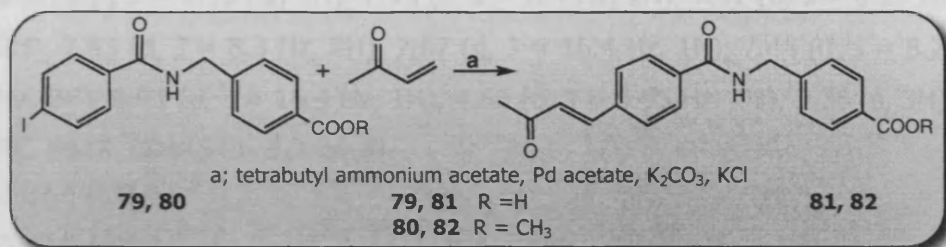
δ 43.8 (CH_2).

δ 52.2 (CH_2).

δ 98.7, 129.6, 133.5, 143.2, 166.7, 166.7. (6 C, quaternary C).

δ 127.6, 128.6, 130.1, 137.7, 137.9, 138.4. (6 C, aromatic CH).

(E)-4-((4-(3-Oxobut-1-enyl)benzamido) methyl)benzoic acid
and (E)-methyl 4-((4-(3-oxobut-1-enyl)benzamido)
methyl)benzoate



To a stirred solution of the appropriate iodo derivative [viz; 4-((4-iodobenzamido)methyl)benzoic acid (**79**) or methyl 4-((4-iodobenzamido) methyl) benzoate (**80**)] (0.003 mol) in DMF (5 mL) were added methyl vinyl ketone (0.63 g , 0.009 mol), tetrabutyl ammonium acetate (1.81 g, 0.006 mol), potassium carbonate (0.93 g, 0.0045 mol), potassium chloride (0.22 g, 0.003 mol) and palladium II acetate (0.02 g, 0.09 mmol). The mixture was stirred at 90 °C for 9 h. After the mixture was cooled, 2N HCl was slowly added and the reaction mixture was stirred at room temperature for 10 min. then, the mixture was diluted with ether and washed with water. The organic layer was dried over anhydrous MgSO₄ and concentrated in vacuo under reduced pressure. The residue was purified by column chromatography using (MeOH in EtOAc in increasing proportions up to 1%) as eluent in case of the free carboxylic acid derivative and (EtOAc:DCM 30:70) as eluent in case of the ester derivative.

ID	R	M w	M F	Yield %	mp
81	H	323	C ₁₉ H ₁₇ NO ₄	36.9	212
82	CH ₃	337	C ₂₀ H ₁₉ NO ₄	26.7	190

4-((4-(3-Oxobut-1-enyl) benzamido)methyl)benzoic acid (81):

^1H NMR ($\text{DMSO}-d_6$) of 81:

δ 9.19 (t, $J = 5.95$ Hz, 1H), 7.95 (d, $J = 8.4$ Hz, 2H), 7.91 (d, $J = 8.25$ Hz, 2H), 7.83 (d, $J = 8.3$ Hz, 2H), 7.67 (d, $J = 16.4$ Hz, 1H), 7.44 (d, $J = 8.25$ Hz, 2H), 6.91 (d, $J = 16.4$ Hz, 1H), 4.56 (d, $J = 5.95$ Hz, 2H), 2.36 (s, 3H).

^{13}C NMR ($\text{DMSO}-d_6$) of 81:

δ 27.4 (CH_3).

δ 42.5 (CH_2).

δ 128.5, 130.3, 135.2, 137.2, 166.5, 189.0, 198.3. (7 C, quaternary C).

δ 127.2, 127.8, 128.3, 128.7, 129.4, 141.9. (6 C, aromatic CH).

Methyl 4-((4-(3-oxobut-1-enyl)benzamido)methyl)benzoate (82):

^1H NMR ($\text{DMSO}-d_6$) of 82:

δ 9.21 (t, $J = 5.9$ Hz, 1H), 7.95 (m, 3H), 7.82 (d, $J = 8.35$ Hz, 2H), 7.66 (d, $J = 16.35$ Hz, 1H), 7.46 (d, $J = 8.3$ Hz, 1H), 6.90 (d, $J = 16.4$ Hz, 1H), 4.56 (d, $J = 5.95$ Hz, 2H), 3.84 (s, 3H), 2.36 (s, 3H).

^{13}C NMR ($\text{DMSO}-d_6$) of 82:

δ 27.4 (CH_3).

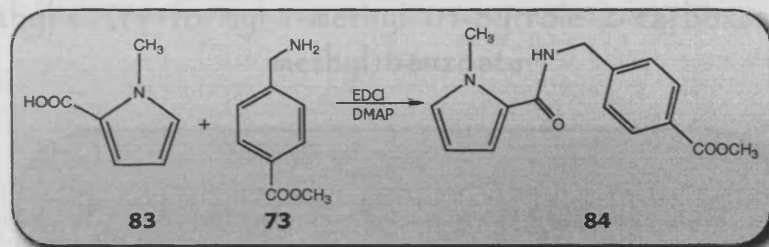
δ 42.5 (CH_2).

δ 52.0 (CH_3).

δ 128.4, 135.3, 137.2, 145.1, 165.7, 166.1, 198.2. (7 C, quaternary C).

δ 127.4, 127.8, 128.1, 128.3, 128.6, 129.1, 129.2, 129.7, 141.7, 141.9. (10 C, CH).

Methyl 4-((1-methyl-1H-pyrrole-2-carboxamido)methyl)benzoate



To a solution of 1-methyl-1H-pyrrole-2-carboxylic acid (**83**) (2.5 g, 0.02 mol) in dry dichloromethane (50 mL), 1-ethyl 3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDCI) (7.66 g, 0.04 mol) and dimethylaminopyridine (DMAP) (4.88 g, 0.04 mol) were added. Methyl 4-(aminomethyl)benzoate (**73**)²⁸⁶ (3.3 g, 0.02 mol) in dry THF (50 mL) was stirred for 3 min under N₂ atmosphere then was added to the above mentioned DCM mixture. The resulting mixture was stirred at RT for 48 h under N₂. The resulting mixture was evaporated under reduced pressure. The crude product was chromatographed using (Ethyl acetate/DCM) (50:50) as eluent. Yield (4.4 g) 81.6% (C₁₅H₁₆N₂O₃), mp 84 °C.

¹H NMR (CDCl₃-d₁) of **84**:

δ 8.03 (d, J = 8.25 Hz, 2H), 7.42 (d, J = 8.2 Hz, 2H), 6.77 (m, 1H), 6.59 (dd, J = 1.65, 3.95 Hz, 1H), 6.29 (br. s, 1H, NH), 6.12 (dd, J = 2.55, 3.85 Hz, 1H), 4.65 (d, J = 6.05 Hz, 2H, CH₂), 3.99 (s, 3H, CH₃), 3.93 (s, 3H, CH₃).

¹³C NMR (CDCl₃-d₁) of **84**:

δ 36.9 (CH₃).

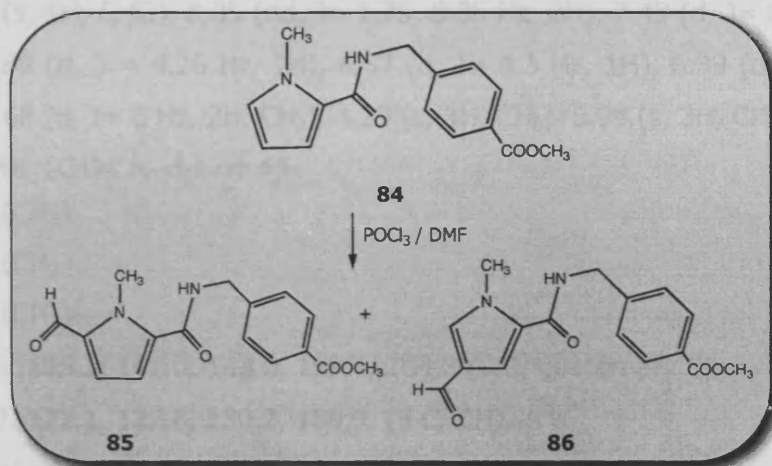
δ 41.8 (CH₂).

δ 52.2 (CH₃).

δ 129.2, 144.0, 161.0, 166.9 (4 C, quaternary C).

δ 107.3, 111.7, 127.4, 128.3, 130.0 (5 C, CH).

Methyl 4-((5-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)
methyl)benzoate &
Methyl 4-((4-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)
methyl)benzoate



To a 250 mL RBF was added anhydrous DMF (0.7 mL, 0.009 mol) it was then cooled to 5-10 °C. To the cooled DMF was added POCl₃ (0.78 mL, 0.0084 mol) dropwise over a couple of minutes. Dry 1,2-DCE (25 mL) was then added and the orange solution was cooled to 0-5 °C during the addition of Methyl 4-((1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate (**84**) (2.07 g, 0.0076 mol) in dry 1,2-DCE (25 mL). The mixture was then heated to reflux for 15 min. the reaction was cooled to room temperature, treated with a mixture of ethyl acetate (60 mL) and water (75 mL), poured into saturated NaHCO₃ (350 mL), and separated. The aqueous layer was washed three times with ether, and the combined organic extracts were washed twice with aqueous saturated Na₂CO₃, dried over anhydrous Na₂SO₄, and evaporated under vacuum. The resulting solid was chromatographed on a silica gel flash column using hexane/ethyl acetate (50:50) to elute two compounds in a 4:1 ratio and in the following order;

Methyl 4-((5-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate (**85**) ($C_{16}H_{16}N_2O_4$); Yield 1.6 g, 70.2%, mp 128 °C.

1H NMR ($CDCl_3-d_1$) **85**:

δ 9.72 (s, 1H, CHO), 8.05 (dd, $J = 1.75, 8.35$ Hz, 2H), 7.43 (d, $J = 8.45$ Hz, 2H), 6.89 (d, $J = 4.25$ Hz, 1H), 6.57 (d, $J = 4.3$ Hz, 1H), 6.39 (br. s, 1H, NH), 4.68 (d, $J = 6$ Hz, 2H, CH_2), 4.29 (s, 3H, CH_3), 3.94 (s, 3H, CH_3).

^{13}C NMR ($CDCl_3-d_1$) of **85**:

δ 34.6 (CH_3)

δ 43.3 (CH_2)

δ 52.2 (CH_3)

δ 123.6, 135.2, 141.0, 143.0, 159.7, 164.8 (6 C, quaternary C).

δ 111.3, 122.1, 127.6, 130.2, 180.9. (5 C, CH).

Methyl 4-((4-formyl-1-methyl-1*H*-pyrrole-2-carboxamido)methyl)benzoate; ($C_{16}H_{16}N_2O_4$) **82** Yield 0.41 g, 17.98 %, mp 126 °C.

1H NMR ($CDCl_3-d_1$) of **86**:

δ 9.76 (s, 1H, CHO), 8.05 (d, $J = 8.35$ Hz, 2H), 7.41 (d, $J = 8.45$ Hz, 2H), 7.38 (d, $J = 1.5$ Hz, 1H), 7.06 (d, $J = 1.7$ Hz, 1H), 6.41 (br. s, 1H, NH), 4.66 (d, $J = 6$ Hz, 2H, CH_2), 4.04 (s, 3H, CH_3), 3.94 (s, 3H, CH_3).

^{13}C NMR ($CDCl_3-d_1$) of **86**:

δ 37.7 (CH_3)

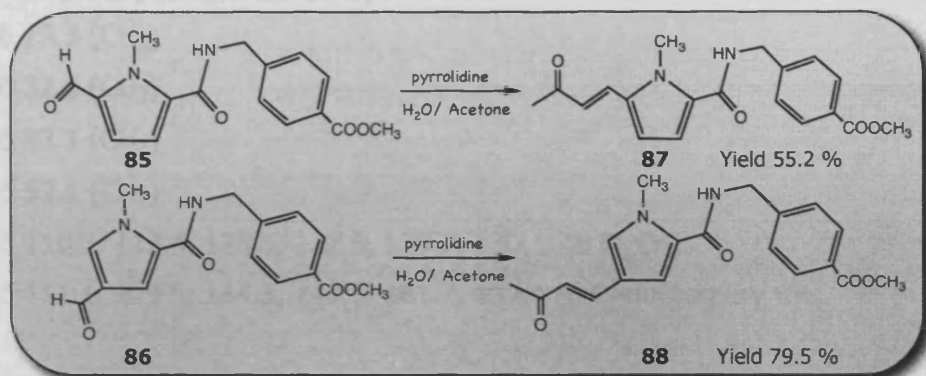
δ 43.1 (CH_2)

δ 52.1 (CH_3)

δ 116.4, 124.3, 148.6, 150.9, 157.4, 164.6 (6 C, quaternary C).

δ 110.6, 127.5, 130.1, 133.9, 184.8. (5 C, CH).

(*E*)-methyl 4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoate &
 (*E*)-methyl 4-((1-methyl-4-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoate



To a stirring mixture of either of the methyl formyl-1*H*-pyrrole-2-carboxamido) methyl) benzoate derivatives (**85** or **86**) (1.5 g, 0.005 mol), acetone (3.7 mL, 0.05 mol), water (50 mL) and pyrrolidine (30 mol %) were added. The reaction was stirred overnight at room temperature and on completion as monitored by TLC was extracted with (3x50 mL) dichloromethane. The organic layer was washed with water (2x 50 mL) and dried over anhydrous Na₂SO₄ and evaporated to obtain crude product. Column chromatography of the crude on silica gel using ethyl acetate as eluent gave pure products. (C₁₉H₂₀N₂O₄) M.w. 340.

ID	Yield %	mp °C
87	55.2	108
88	79.5	134

(*E*)-methyl 4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido)methyl)benzoate;

¹H NMR (CDCl₃-d₁) of **87**:

δ 8.03 (d, J = 8.25 Hz, 2H), 7.49 (d, J = 15.8 Hz, 1H), 7.40 (d, J = 8.2, 2H), 6.63 (m, 3H), 6.41 (br. s, 1H, NH), 4.65 (d, J = 6 Hz, 2H), 4.06 (s, 3H, CH₃), 3.93 (s, 3H, CH₃), 2.35 (s, 3H, CH₃).

¹³C NMR (CDCl₃-d₁) of **87**:

δ 29.3 (CH₃)

δ 32.6 (CH₃)

δ 43.1 (CH₂)

δ 52.1 (CH₃)

δ 110.3, 112.4, 125.6, 127.5, 129.7, 130.1. (6 C, CH).

δ 111.1, 129.5, 134.3, 143.5, 161.7, 197.4 (6 C, quaternary C).

(*E*)-methyl 4-((1-methyl-4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoate (**88**);

¹H NMR (DMSO-d₆) of **88**:

δ 8.04 (d, J = 8.3 Hz, 2H), 7.42 (d, J = 8.2, 2H), 7.39 (d, J = 16.1 Hz, 1H), 7.02 (d, J = 1.55, 1H), 6.83 (d, J = 1.7 Hz, 1H), 6.46 (br. t, J = 5.7, 1H, NH), 6.41 (d, J = 16.1, 1H), 4.65 (d, J = 6 Hz, 2H), 3.98 (s, 3H, CH₃), 3.94 (s, 3H, CH₃), 2.3 (s, 3H, CH₃).

¹³C NMR (DMSO-d₆) of **88**:

δ 27.2 (CH₃)

δ 37.2 (CH₃)

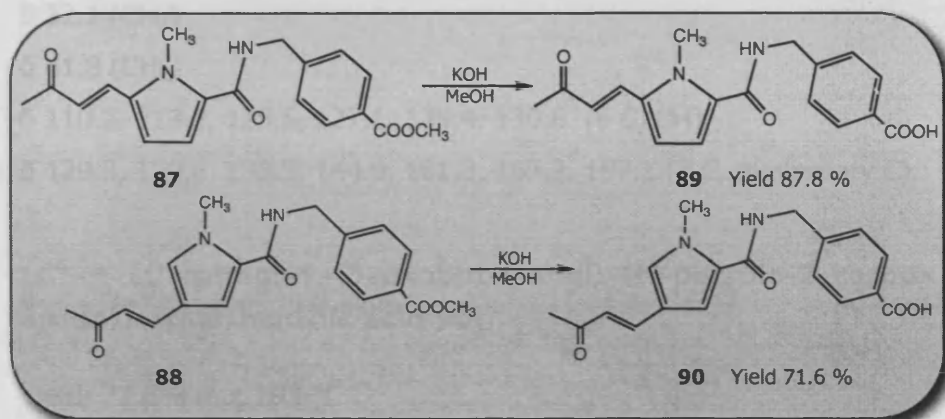
δ 43.0 (CH₂)

δ 52.1 (CH₃)

δ 109.7, 124.2, 127.5, 130.1, 130.3, 136.6. (6 C, CH).

δ 118.9, 129.4, 143.5, 161.2, 198.3 (5 C, quaternary C).

(*E*)-4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido) methyl)benzoic acid &
 (*E*)-4-((1-methyl-4-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido) methyl)benzoic acid



A mixture of either of the methyl (1-methyl-1*H*-pyrrole) benzoate derivatives (**87** or **88**) (0.34 g, 0.001 mol) KOH (0.056 g, 0.001 mol in 5 mL water), MeOH (30 mL) and THF (30 mL) was stirred at 45 °C for 48 h. The resulting solution was evaporated under reduced pressure, water was added and pH was adjusted to pH 2 using 1*N* HCl solution. This was extracted with ethyl acetate (3x 100 mL). the organic layer was washed with brine solution and dried over dry Na₂SO₄ then evaporated under vacuum. The crude product was purified on silica gel column chromatography using ethyl acetate as eluent.

(*E*)-4-((1-methyl-5-(3-oxobut-1-enyl)-1*H*-pyrrole-2-carboxamido) methyl)benzoic acid **89**

Yield; 87.8% (dark green crystals) mp 107 °C

¹H NMR (CDCl₃-d₁) of **89**:

δ 12.98 (br. s, 1H, COOH), 8.85 (br. t, J=6 Hz, NH, 1H), 7.91 (d, J=8.25 Hz, 2H), 7.56 (d, J=16 Hz, 1H), 6.92 (d, J=4.3, 1H), 6.82 (d, J=4.25, 1H),

6.63 (d, $J=15.95$, 1H), 4.49 (d, $J=6$ Hz, 2H, CH₂), 3.89 (s, 3H, CH₃), 2.43 (s, 3H, CH₃).

¹³C NMR (CDCl₃-d₁) of 89:

δ 27.2 (CH₃)

δ 32.1 (CH₃)

δ 41.8 (CH₂)

δ 110.2, 113.3, 125.5, 127.1, 129.4, 130.6. (6 C, CH).

δ 129.3, 129.6, 133.5, 144.9, 161.0, 167.2, 197.1 (7 C, quaternary C).

(*E*)-4-((1-methyl-4-(3-oxobut-1-enyl)-1H-pyrrole-2-carboxamido)methyl)benzoic acid 90;

Yield; 71.6% mp; 193 °C

¹H NMR (DMSO-d₆) of 90:

δ 12.8 (br. s, 1H, COOH), 8.8 (br. t, $J = 6$, 1H, NH), 7.9 (d, $J=8.2$ Hz, 2H), 7.5 (d, $J=16.2$ Hz, 1H), 7.4 (m, 3H), 7.2 (d, $J=1.6$ Hz, 1H), 6.3 (d, $J=16.1$, 1H), 4.5 (d, $J=6$, 2H, CH₂), 3.9 (s, 3H, CH₃), 2.3 (s, 3H, CH₃).

¹³C NMR (DMSO-d₆) of 90:

δ 26.8 (CH₃)

δ 36.6 (CH₃)

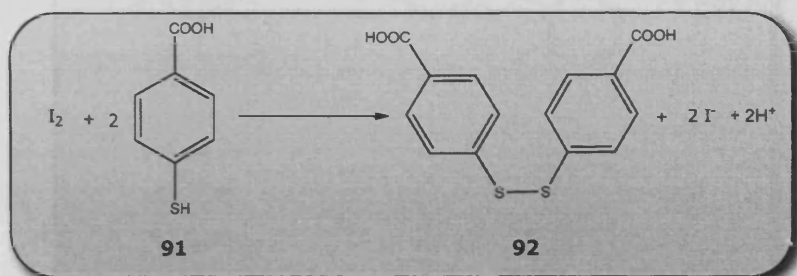
δ 41.7 (CH₂)

δ 110.5, 123.3, 127.1, 129.4, 131.0, 137.4. (6 C, CH).

δ 118.1, 129.5, 129.2, 145.0, 161.0, 167.3, 192.5 (7 C, quaternary C).

4.3 Thiol derivatives

4,4' Dithiobisbenzoic acid



4- Mercaptobenzoic acid (**91**) (2.5 g, 0.016 mol) and iodine (2.03 g, 0.008 mol) were dissolved in 30 ml ethanol. Triethylamine (3 ml, 0.022 mol) was added and the solution was stirred overnight. The cloudy solution was evaporated under vacuum then neutralised with 0.01 M HCl. the residual precipitate was filtered out, dried then washed with ether.

Yield; 2.2 g (89.7%). mp; >300 °C ²⁹⁶

¹H NMR (DMSO-d₆) of 92:

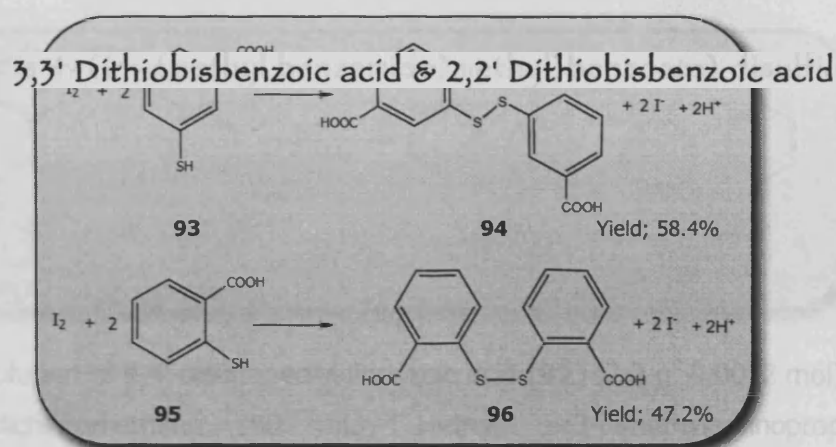
δ 13.04 (br. s, 2H, COOH), 7.94 (d, J=8.2 Hz, 4H), 7.65 (d, J=8.2 Hz, 4H).

¹³C NMR (DMSO-d₆) of 92:

δ 129.7, 130.3. (2 C, CH).

δ 126.1, 140.8, 166.6. (3 C, quaternary C).

3,3'-Dithiobisbenzoic acid & 2,2'-Dithiobisbenzoic acid



Same procedure was used to prepare the meta (**94**) and ortho (**96**) structural isomers of **92**

Yield; **94**; 58.4% Mp; 250°C²⁹⁷

¹H NMR (DMSO-*d*₆) of **94**:

δ 8.05 (dd (t), J=1.65, 1.65 Hz, 2H), 7.84 (ddd (dt), J=7.8, 1.2, 1.2 Hz, 2H), 7.75 (ddd, J=7.8, 1.98, 1 Hz, 2H), 7.52 (dd (dt), J=7.9, 1.3, 1.3 Hz, 2H).

¹³C NMR (DMSO-*d*₆) of **94**:

δ 127.5, 128.5, 129.8, 131.3. (4 C, CH).

δ 132.1, 136.3, 164.4. (3 C, quaternary C).

¹H NMR (DMSO-*d*₆) of **96**:

Yield; **96**; 47.2%. mp; 289°C

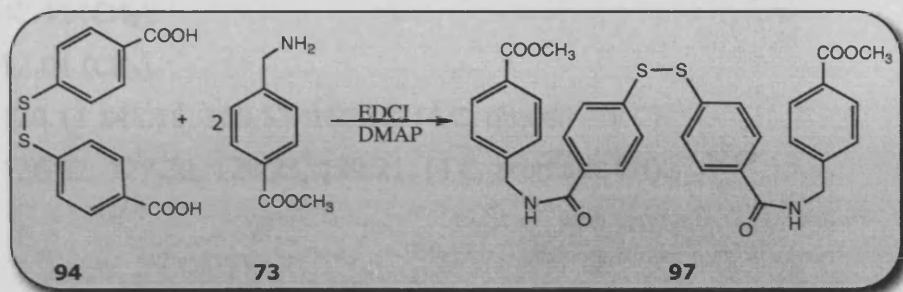
δ 8.05 (dd, J = 7.75, 1.4 Hz, 2H), 7.63 (dd, J = 8.15, 0.9 Hz, 2H), 7.56 (ddd, J = 7.25, 7.25, 1.45 Hz, 2H), 7.34 (ddd, J = 7.75, 1.05 Hz, 2H).

¹³C NMR (DMSO-*d*₆) of **92**:

δ 124.98, 125.98, 131.54, 133.22. (4 C, CH).

δ 128.09, 138.88, 167.58. (3 C, quaternary C).

Bis 4-((4-(methyl benzamido)methyl)benzoate) disulfide



To a solution of 4,4'-disulfanediyldibenzoic acid (**92**) (2.2 g, 0.0072 mol) in dry dichloromethane (50 mL), 1-ethyl 3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDCI) (2.76 g, 0.014 mol) and dimethylaminopyridine (DMAP) (1.76 g, 0.014 mol) were added. Methyl 4-(aminomethyl)benzoate (**73**) (2.4 g, 0.014 mol) in dry THF (50 mL) was stirred for 3 min under N₂ atmosphere then in dry dichloromethane (50 mL), 1-ethyl 3-(3-dimethylaminopropyl)carbodiimide was added to the above mentioned DCM mixture. This mixture was stirred at room temperature for 48 h under N₂. The resulting mixture was evaporated under reduced pressure. The crude solid was partitioned between water and ethylacetate. The combined organic layer was washed with brine, dried using anhydrous sodium sulfate, filtered, evaporated and chromatographed using ethyl acetate as eluent.

Yield; 3.96 g of white solid (91.5%).

mp; charring >300 °C

¹H NMR (DMSO-d₆) of 97:

δ 9.19 (t, J = 6 Hz, 2H, NH), 7.97 (d, J = 8.2 Hz, 4H), 7.96 (d, J = 8.5 Hz, 4H), 7.70 (d, J = 8.5 Hz, 4H), 7.49 (d, J = 8.2 Hz, 4H), 4.59 (d, J = 6 Hz, 4H, CH₂), 3.89 (s, 6H, COOCH₃).

^{13}C NMR (DMSO- d_6) of 97:

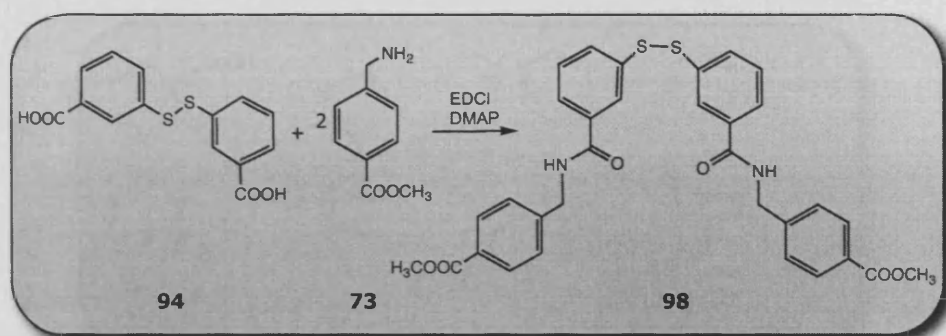
δ 42.42 (CH_2).

δ 52.01 (CH_3).

δ 128.11, 145.16, 165.53, 166.08. (4 C, quaternary C).

δ 126.42, 127.30, 128.33, 129.21. (4 C, aromatic CH).

Bis 3-((4-(methyl benzamido)methyl)benzoate) disulfide



Same procedure was used to prepare the meta (**98**) and ortho (**99**) isomers of **97**

(**98**) mp; 135 °C. yield; 41%

¹H NMR (DMSO-*d*₆) of **98**:

δ 9.21 (t, *J* = 5.9 Hz, 2H, NH), 8.08 (dd (t), *J* = 1.7, 1.7 Hz, 2H, C₂), 7.93 (d, *J* = 8.3 Hz, 4H, C_{2'}), 7.83 (ddd (dt), *J* = 7.9, 1.15, 1.15 Hz, 2H, C₆), 7.72 (ddd, *J* = 7.9, 1.95, 0.9 Hz, 2H, C₄), 7.52 (dd (t), *J* = 7.8, 7.8 Hz, 2H, C₅), 7.44 (d, *J* = 8.35 Hz, 4H, C_{3'}), 4.54 (d, *J* = 5.9 Hz, 4H, CH₂), 3.85 (s, 6H, 2COOCH₃).

¹³C NMR (DMSO-*d*₆) of **98**:

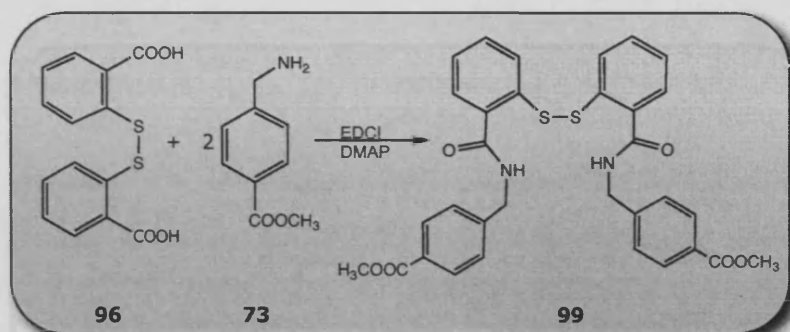
δ 42.50 (CH₂).

δ 52.01 (CH₃).

δ 126.21, 126.46, 127.34, 129.22, 129.64, 129.90. (6C, aromatic CH).

δ 128.14, 135.25, 136.17, 145.03, 165.33, 166.06. (6C, quaternary C).

Bis 2-((4-(methyl benzamido)methyl)benzoate) disulfide



(99) mp; 142 °C. yield; 23%

^1H NMR (DMSO- d_6):

δ 9.28 (t, $J = 5.9$ Hz, 2H, NH), 7.95 (d, $J = 8.5$ Hz, 4H), 7.74 (dd, $J = 7.5$, 1.5 Hz, 2H), 7.67 (dd, $J = 8.0$, 1.0 Hz, 2H), 7.53 (d, $J = 8.5$ Hz, 4H), 7.46 (m, 2H), 7.33 (ddd, $J = 7.5$, 7.5, 1 Hz, 2H), 4.58 (d, $J = 6.0$ Hz, 4H, CH_2), 3.85 (s, 6H, 2 COOCH_3).

^{13}C NMR (DMSO- d_6):

δ 42.44 (CH_2).

δ 52.03 (CH_3).

δ 125.82, 127.40, 128.00, 129.14, 131.26, 132.05. (6C, aromatic CH).

δ 128.22, 133.74, 136.86, 145.03, 166.07, 167.02. (6C, quaternary C).

^{13}C NMR (CDCl_3 - d_3) 100

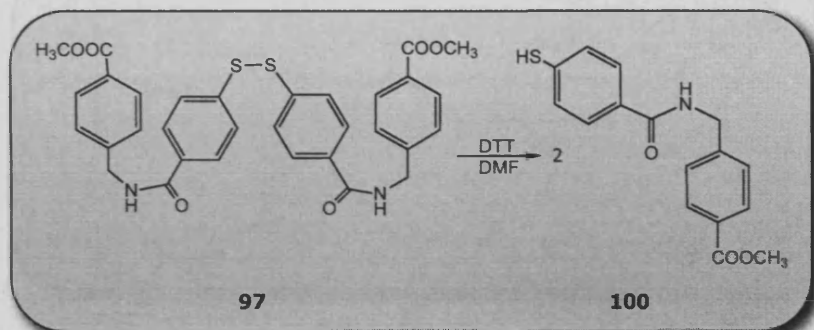
δ 43.50 (CH_2).

δ 52.13 (CH_3).

δ 127.58, 127.74, 129.43, 130.01 (6C, CH).

δ 128.98, 131.04, 136.58, 145.45, 166.77, 166.73 (6C, quaternary C).

Methyl 4-((4-mercaptobenzamido)methyl)benzoate



The dimer compound **97** (0.2 g, 0.00033 mol) was dissolved in 10 mL of DMF. Dithiothreitol (DTT) (62mg, 0.0004 mol) was added and the mixture was stirred at room temperature under N_2 atmosphere for 16 h. the solvent was removed under vacuum and the residue was purified by column chromatography ($CHCl_3/MeOH$ 99:1).

Yield; 0.14 g (70.0%).

1H NMR ($CDCl_3-d_1$) 100:

δ 7.91 (d, $J = 8.2$ Hz, 2H), 7.58 (d, $J = 8.3$ Hz, 2H), 7.29 (d, $J = 8.2$ Hz, 2H), 7.19 (d, $J = 8.3$ Hz, 2H), 6.57 (t, $J = 5.3$ Hz, 1H), 4.58 (d, $J = 5.3$ Hz, 2H, CH_2), 3.82 (s, 3H, OCH_3), 3.50 (s, 1H, SH).

^{13}C NMR ($CDCl_3-d_1$) 100:

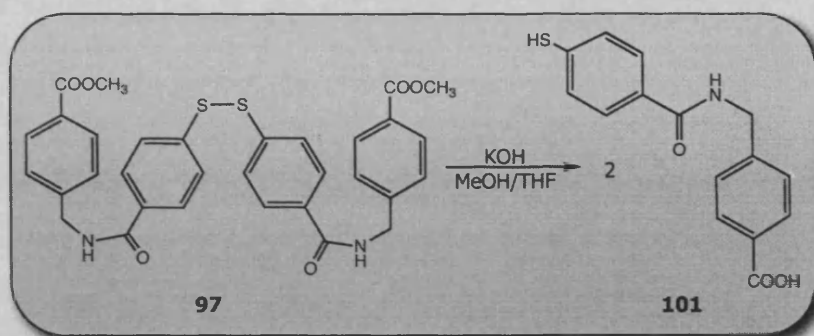
δ 43.60 (CH_2).

δ 52.13 (CH_3).

δ 127.58, 127.74, 129.42, 130.04. (4C, CH).

δ 128.58, 131.04, 136.58, 143.45, 166.77, 166.79. (6C, quaternary C).

4-((4-Mercaptobenzamido)methyl)benzoic acid



A mixture of Bis (methyl 4-((4-benzamido)methyl)benzoate) disulfide (**97**) (2 g, 0.0037 mol), KOH (0.41 g, 0.0074 mol) in 5 mL H₂O, MeOH (30 mL) and THF (30 mL) was stirred at 45 °C for 48 h. The resulting solution was evaporated under reduced pressure, water was added and pH was adjusted to pH 2 using 1N HCl solution. This was extracted with ethyl acetate (3x 50 mL). the organic layer was washed with brine solution and dried over anhydrous Na₂SO₄ then evaporated under vacuum. The crude product was purified by flash chromatography on silica gel using ethyl acetate. Yield; 1.7 g (92.1%) . mp ; 241 °C.

C ₁₅ H ₁₃ NO ₃ S	C%	H%	N%
Calculated	62.70	4.56	4.87
Found	63.07	4.66	4.70

¹H NMR (DMSO-d₆) of 101:

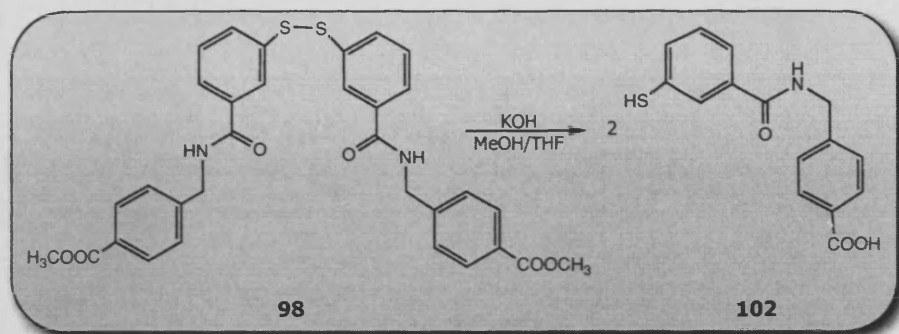
δ 12.91 (br. s, 1H, COOH), 9.08 (t, J = 5.7 Hz, 1H, NH), 7.96 (d, J = 8.3 Hz, 2H), 7.83 (d, J = 8.3 Hz, 2H), 7.47 (d, J = 8.3 Hz, 2H), 7.44 (d, J = 8.3 Hz, 2H), 5.79 (s, 1H, SH), 4.58 (d, J = 5.7 Hz, CH₂).

¹³C NMR (DMSO-d₆) of 101:

δ 42.44 (CH₂). δ 126.38, 127.13, 129.26, 129.35. (4C, CH).

δ 128.35, 133.18, 139.01, 144.63, 165.50, 167.13. (6C, quaternary C)

4-((3-Mercaptobenzamido)methyl)benzoic acid



Yield; (17.3 %) . mp ; 234 °C.

¹H NMR (DMSO-d₆) of 102:

δ 12.87 (br. s, 1H, COOH), 9.34 (t, J = 5.95 Hz, 1H, NH), 8.21 (m, 1H), 8.09 (ddd, J = 7.8 Hz, 1H), 7.92 (d, J = 8.3 Hz, 2H), 7.83 (m, 1H), 7.69 (dd, J = 7.7, 7.7 Hz, 1H), 7.44 (d, J = 8.4 Hz, 1H), 5.63 (s, 1H, SH), 4.57 (d, J = 5.9 Hz, CH₂).

¹³C NMR (DMSO-d₆) of 102:

δ 42.52 (CH₂).

δ 123.38, 127.18, 127.37, 129.21, 129.38, 130.09. (6 C, CH).

δ 134.75, 136.85, 137.14, 146.44, 166.21, 167.16. (6 C, quaternary C).

Chapter 5

Replicon & Enzyme assays: Overview and Findings

Chapter 5

HCV Replicon & Enzyme assays: Overview and Findings

5.1 Replicon

Replicon is a genetic element, which can be either DNA or RNA that can replicate under its own control in a cell. ³⁴³

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. ³⁴³

Initially HCV replicons were autonomously replicating genetically engineered HCV RNA 'mini-genomes' in which the region 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 the human hepatoma cell line Huh-7, followed by selection results in HCV cell clones. Recently, many different HCV replicons have been developed that allow screening of chemical compounds. ³⁴⁴ Replicon development had gone through the following stages;

5.2 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 efficiency. ^{345, 346} This low efficiency 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. ³⁴³

5.3 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 devised³⁴⁷⁻³⁵⁰. In these genetically modified HCV 'minigenomes' the region that encodes the structural proteins was replaced 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 used³⁴³ (Figure 5.1).

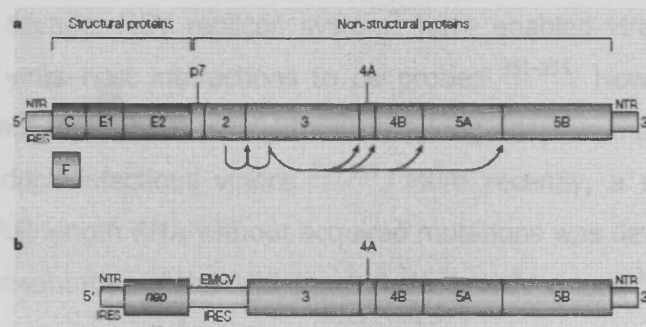


Figure 5.1 a) HCV genome encoding the core structural protein (C) and the envelope glycoproteins (E1 and E2), the non-structural (NS) proteins NS2 to NS5B. b) A subgenomic replicon by replacing up to the NS2-encoding region by the *neo* gene and the internal ribosome-entry site (IRES) of another virus (encephalomyocarditis virus; EMCV).

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)³⁴³. After transfection of the human hepatoma cell line Huh-7 with the subgenomic replicon RNA and subsequent selection with G418, only cells in which replicon was amplified to high levels expressed sufficient amounts of NPT, therefore, survive into a colony that can be isolated and expanded³⁴³ (Figure 5.2).

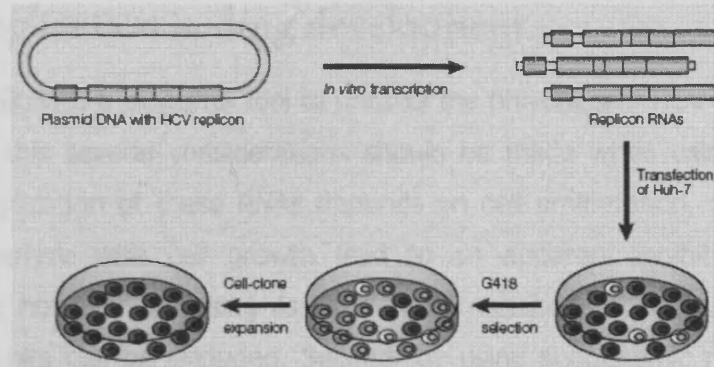


Figure 5.2 Establishment of cell clones. Cells that did not take up the RNA and cells in which the replicon RNA does not replicate will die because of the toxic effect of G418.

5.4 Breakthrough for HCV research

In the past decade, HCV replicon systems have enabled viral molecular biology and virus–host interactions to be probed³⁵¹⁻³⁵⁵. However, these systems cannot replicate *in vitro* without acquiring adaptive mutations, nor do they produce infectious virions³⁵⁶⁻³⁵⁸. 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 virions³⁵⁹⁻³⁶¹.

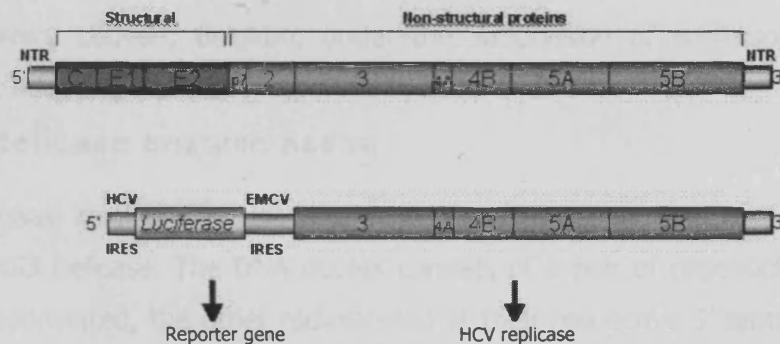


Figure 5.3 HCV subgenomic replicon encoding firefly luciferase instead of the structural proteins.

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 clones³⁶²⁻³⁶⁹ (Figures 5.3).

5.5 Application in drug development

HCV replicon is a powerful tool to unravel the principles of HCV replication. 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, a possible interference of the HCV structural proteins with the antiviral activity of a compound might be missed. With the availability of full-length genomes this possibility can be examined. Third, in bicistronic replicons, compounds that interfere with the internal IRES (from HCMV) also inhibit RNA replication. These compounds can now be excluded, because monocistronic replicons without an internal IRES are now available ^{370,371}.

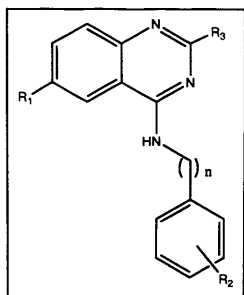
5.6 Subgenomic replicon assay

Both the subgenomic replicon and cytostatic assays for the prepared compounds was done thankfully in Rega Institute for Medical Research, KULeuven, Leuven, Belgium, under the supervision of professor Johan Neyts, adopting a protocol reported by Paeshuyse *et al* ³⁷².

5.7 Helicase enzyme assay

This assay involves the use of a DNA duplex substrate and recombinant HCV NS3 helicase. The DNA duplex consists of a pair of oligonucleotides, one biotinylated, the other radiolabeled at their respective 5' termini. The assay design is based on the fact that the release strand, a 20-mer oligonucleotide, is radiolabeled on its 5'-end, while the template strand, a 39-mer, is immobilized on a neutravidin-coated plate surface via its 5'-end biotin molecule. Helicase activity results in unwinding of the duplex DNA substrate and hence the release of the radiolabeled oligonucleotide, which translates in signal reduction with respect to control wells ³⁷³.

5.8 Anti-HCV assay of Quinazoline structures



EC₅₀ and CC₅₀ antiviral activity against HCV (genotype 1b)
SI selectivity index

Compound	R1	R2	R3	n	EC ₅₀ (μM)	CC ₅₀ (μM)	SI
60	H	2'-COOH	H	0	135.9	139.6	1.03
29	H	3'-COOH	H	0	> 188.7	> 188.7	-
38	H	4'-COOH	H	0	86.8	>188.7	2.17
30	H	3'-COOCH ₃	H	0	78.9	89.6	1.14
39	H	4'-COOCH ₃	H	0	32.3	>179.2	5.6
61	H	2'-CONH ₂	H	0	125	189.4	1.5
31	H	3'-CONH ₂	H	0	117.4	189.3	1.6
40	H	4'-CONH ₂	H	0	159.1	>189.3	1.2
58	NO ₂	2'-COOH	H	0	>161.3	>161.3	-
24	NO ₂	3'-COOH	H	0	64.5	>161.3	2.5
35	NO ₂	4'-COOH	H	0	45.2	>161.3	3.6
25	NO ₂	3'-COOCH ₃	H	0	98.8	>154.3	1.6
36	NO ₂	4'-COOCH ₃	H	0	9.2	154.3	16.8
59	NO ₂	2'-CONH ₂	H	0	19.4	35.6	1.8
26	NO ₂	3'-CONH ₂	H	0	16.2	106.8	6.6
37	NO ₂	4'-CONH ₂	H	0	19.4	>161.8	8.3
55	NO ₂	4'-CN	H	0	109.9	>171.8	1.6
41	H	4'-COOH	H	1	>179.2	>179.2	-
42	NO ₂	4'-COOH	H	1	9.3	148.2	15.9
45	H	4'-COOCH ₃	H	1	34.1	85.3	2.5
44	NO ₂	4'-COOCH ₃	H	1	124.3	147.9	1.2
51	H	3'-tetrazolyl	H	0	>173	>173	-
50	NO ₂	3'-tetrazolyl	H	0	>149.7	>149.7	-
54	H	4'-tetrazolyl	H	0	117.7	>173	1.5
52	NO ₂	4'-tetrazolyl	H	0	17.9	>149.7	8.3
65	H	4'-COOCH ₃	CH ₃	0	78.5	170.7	2.2

Table 5.1 Inhibition of HCV replication in the replicon system assay, showing 50% effective concentration (EC₅₀) and 50% cytostatic concentration (CC₅₀). EC₅₀: The effective concentration (μM) required to reduce luciferase signal by 50%. CC₅₀: The cytostatic concentration (μM) required to inhibit cell viability by 50%. SI, selectivity index= CC₅₀ / EC₅₀.

Among these 26 compounds, 6 showed EC₅₀ values less than 20 μ M, namely, **36, 59, 26, 37, 42 and 52** with relatively high selectivity index (16.8 in case of compound **36**). The general pattern of the activity of these compounds was in favour of the acidic group or its derivatives to be in position 4' of the phenyl ring like in the case of compounds **38, 39, 35, 36, 37, 42 and 52**, and the nitro group in position 6 of the quinazoline ring system rather than its absence as exhibited by compounds **36, 26, 37, 42 and 52**. Yet, some compounds, which lack the nitro group still can display moderate activity as in compounds **30, 39, 45 and 65**. Also it was noticed that the introduction of the methyl group in position 2 did not improve the activity (compound **39** vs. compound **65**).

It may also worth to note that in general the free carboxylic acid derivatives show less activity compared to the ester or amide counterparts probably due to the reduced cell permeability.

Interestingly, these compound were lacking any helicase inhibitory activity. This lack of conclusive prediction in the docking studies could be partly because of the high flexibility nature of NS3 helicase enzyme attributed to the relative free rotation of domain 2 in respect to domains 1 and 3. And partly to the approximation nature of molecular modelling techniques.

Instead these compounds might have some other viral or intracellular target to have antihepatitis C activity^{262,263}. Additional studies are to be undertaken to further elucidate the exact mechanism by which this series of quinazolines inhibited HCV replication in the replicon assay.

In summary, we designed, synthesized and evaluated 26 quinazoline derivatives as anti hepatitis C virus. The results showed that some of these compounds possessed significant anti-HCV activity. These novel inhibitors are amenable to more exploration and offer a reasonable starting point for further drug discovery efforts.

5.9 Anti-HCV assay of α , β Unsaturated ketones

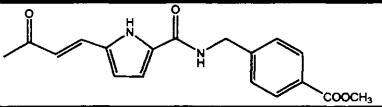
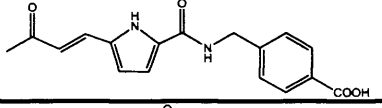
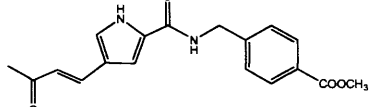
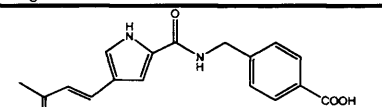
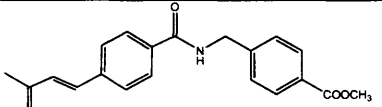
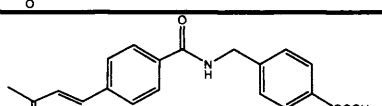
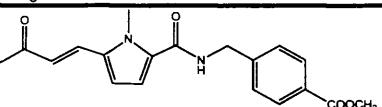
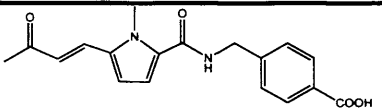
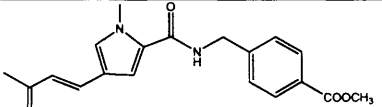
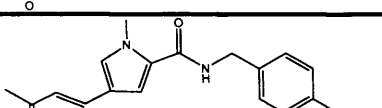
ID	Compound	EC ₅₀ (μ M)	CC ₅₀ (μ M)	IC ₅₀
74		9.2	30.7	0.26 μ M
76		93.6	>160.3	>300 μ M
75		76.7	>153.4	>300 μ M
77		>160.3	>160.3	>300 μ M
82		22.8	82.6	>300 μ M
81		>154.8	>154.8	>300 μ M
87		19.2	55.9	>300 μ M
89		>154.8	>154.8	>300 μ M
88		63.24	67.65	>300 μ M
90		1.23	>153.4	>300 μ M

Table 5.2 Inhibition of HCV replication in the replicon system and helicase unwinding assays, showing EC₅₀, CC₅₀ and IC₅₀

Michael acceptor based compounds were evaluated in a strand-displacement enzymatic assay based on the method of Hicham against purified recombinant HCV helicase.³⁷³ Compound **74** showed an IC₅₀ of 0.26 μM, while the other analogues did not show any significant activity at a concentration as high as 300 μM. For this, **74** was considered to be a potential hit and to further examine the importance of the cysteine residue to its binding to the helicase, the enzymatic assay was repeated in the presence of the thioactive agent; NEM (N-ethylmaleimide) (Figure 5.4). It was observed that the inhibitory activity of compound **74** was lost when added to the enzyme pre-incubated with NEM (case 4 Table 5.3). Not surprisingly, the enzyme preserved its helicase activity in these conditions, probably because NEM, as mercaptoethanol, is not big enough to impede the nucleic acid binding site.

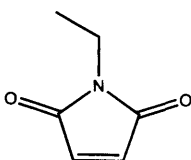


Figure 5.4 Chemical structure of the thioactive agent NEM (N-ethylmaleimide)

Exp.	Helicase	NEM	Compound 74	Radioactivity	Helicase activity
1	+	-	-	0.051+/-0.009	Active
2	+	+	-	0.053+/-0.017	Active
3	+	-	+	0.511+/-0.025	Inactive
4	+	+	+	0.062+/-0.031	Active
5	-	+	+	0.522+/-0.005	Inactive

Table 5.3 The effect of NEM on the anti helicase activity of **74**

It should also be noted that in the original NS3 helicase crystal structure used, two other cysteine residues beside Cys 431 appeared to have reacted with mercaptoethanol (Cys 279 and Cys 499), but these are placed on the enzyme surface well away from the nucleic acid binding site (>20 Å), thus, making their involvement in the activity of these compounds less probable. The active molecule was also evaluated for a potential inhibitory

effect on HCV subgenomic replicon replication as reported earlier ²⁷². Compound **74** proved however rather cytostatic to the hepatoma cells (EC_{50} 9.2 μ M; CC_{50} 30.7 μ M, which was expected, given the fact that the vinyl ketone group is a known toxicophore. It is worth noting that the ester analogues were prepared and evaluated for biological activity. It is possible to speculate from the model that the free acid analogue **76** might have a better interaction with Arg 393. However, it failed to inhibit NS3 helicase in this particular assay conditions.

5.10 Anti-HCV assay Thiol derivatives

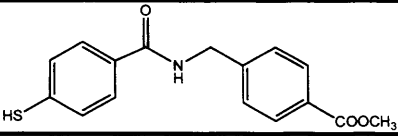
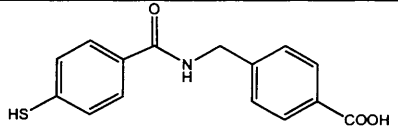
ID	Compound	EC_{50} (μ M)	CC_{50} (μ M)	IC_{50}
100		71.4	>166	>300 μ M
101		68.3	>174	0.37 μ M

Table 5.4 Inhibition of HCV replication in the replicon system assay, showing EC_{50} , CC_{50} and IC_{50}

In the case of the two thiol compounds (**100** and **101**) examined for antihelicase activity, **101** exhibited submicromolar activity (0.37 μ M) in the helicase enzyme assay. As it could be anticipated that the free acid analogue **101** might have a better interaction with Arg 393 than the ester **100**.

Further studies are underway aimed for obtaining a co-crystallised ligand/protein complex, which might give better insight about the proposed mechanism of action for compounds **74** and **101**.

Chapter 6
Bibliography

Chapter 6

Bibliography

1. Feinstone S., Kapikian A., Purcell R., Alter H., Holland P. Transfusion associated hepatitis not due to viral hepatitis type A or B. *New Engl. J. Med.* **292**, 767–770 (1975).
 2. Choo Q. *et al.* isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359-362 (1989).
 3. World Health organisation, Hepatitis C- Global surveillance update *weekly epidemiological record* **75**; 17-28 (2000).
 4. Anonymous, hepatitis C – global prevalence (update), *World Health Org. Weekly Epidemiol. Rec.* **75**, 18-19 (2000).
 5. Antonelli A., Ferri C., Ferrari S., Colaci M., Sansonno D., Fallahi P. Endocrine manifestations of hepatitis C virus infection, nature clinical practice, *Nature Clinical practice: Endocrinology & Metabolism*, **5**, 26-34 (2009).
 6. Craxi A., Laffib G., Zignego A. Hepatitis C virus (HCV) infection: A systemic disease. *Molecular Aspects of Medicine*, **29**, 85–95 (2008).
 7. Hoofnagle J. Course and outcome of hepatitis C. *Hepatology* **36**, S21–S29 (2002).
 8. Bigger C., Brasky K., Lanford R. DNA micro array analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J. Virol.* **75**, 7059–7066 (2001).
 9. Bigger C. *et al.* Intrahepatic gene expression during chronic hepatitis C virus infection in chimpanzees. *J. Virol.* **78**, 13779–13792 (2004).
 10. Su A. *et al.* Genomic analysis of the host response to hepatitis C virus infection. *Proc. Natl Acad. Sci. USA* **99**, 15669–15674 (2002).
 11. Thimme R. *et al.* Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proc. Natl Acad. Sci. USA* **99**, 15661–15668 (2002).
 12. Lunel F., Musset L. Hepatitis C virus infection and cryoglobulinemia. *Viral Hepatitis Rev.* **2**:111-124 (1996).
 13. Agnello V., Chung R., Kaplan L. A role of hepatitis C virus infection in type II cryoglobulinemia. *New Engl. J. Med.*; **327**, 1490-1495 (1992).
-

14. Dammacco F. et al. The lymphoid system in hepatitis C virus infection: autoimmunity, mixed cryoglobulinemia, and overt B-cell malignancy. *Semin. Liv. Dis.* **20**, 143–157 (2000).
 15. Deforges S. et al. Expression of hepatitis C virus proteins in epithelial intestinal cells in vivo. *J. Gen. Virol.* **85**, 2515–2523 (2004).
 16. Forton D., Karayiannis P., Mahmud N., Taylor-Robinson S., Thomas H. C. Identification of unique hepatitis C virus quasispecies in the central nervous system and comparative analysis of internal translational efficiency of brain, liver, and serum variants. *J. Virol.* **78**, 5170–5183 (2004).
 17. Rosenberg S. Recent advances in the molecular biology of hepatitis C virus. *J. Mol. Biol.* **313**, 451–464 (2001).
 18. Friebe P., Lohmann V., Krieger N., Bartenschlager R. Sequences in the 5' non-translated region of hepatitis C virus required for RNA replication. *J. Virol.* **75**, 12047–12057 (2001).
 19. Buckwold V., Beer B., Donis R. Bovine viral diarrhea virus as a surrogate model of hepatitis C virus for the evaluation of antiviral agents. *Antiviral Res.* **60**, 1–15 (2003).
 20. Lohmann V. et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*. **285**, 110–113 (1999).
 21. Blight K., Kolykhalov A., Rice C. Efficient initiation of HCV RNA replication in cell culture. *Science* **290**, 1972–1974 (2000).
 22. Bukh J., Miller R., Purcell R. Genetic heterogeneity of hepatitis C virus — quasispecies and genotypes. *Semin. Liv. Dis.* **15**, 41–63 (1995).
 23. Neumann A. et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon- α therapy. *Science* **282**, 103–107 (1998).
 24. Lindenbach B., Rice C. Unravelling hepatitis C virus replication from genome to function. *NATURE*, **436 (18)**, 933–938 (2005).
 25. Simmonds P. Genetic diversity and evolution of hepatitis C virus—15 years on. *J. Gen. Virol.* **85**, 3173–3188 (2004).
 26. Farci P., Purcell R. Clinical significance of hepatitis C virus genotypes and quasispecies. *Semin. Liv. Dis.* **20**, 103–126 (2000).
 27. Reherrmann B., Nascimbeni M. Immunology of hepatitis B virus and hepatitis C virus infection. *Nature Rev. Immunol.* **5**, 215–229 (2005).
-

-
28. Shoukry N., Cawthon A., Walker C. Cell-mediated immunity and the outcome of hepatitis C virus infection. *Annu. Rev. Microbiol.* **58**, 391–424 (2004).
 29. Moradpour D., Penin F., Rice C., Replication of hepatitis C virus, *Nature reviews microbiology*, **5**, 453-463 (2007).
 30. Feld J., Hoofnagle J. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *NATURE*, **436 (18)**, 967-972 (2005).
 31. Moussali J., Opolon P., Poynard T. Management of hepatitis C. *J. Hepatitis* **5**, 73-82 (1998).
 32. Wang C., Pflugheber J., Sumpter R., Sodora D., Hui D., Sen G., Gale M. Alpha interferon induces distinct translational control programs to suppress hepatitis C virus RNA replication. *J. Virol.*, **77**, 3898-3912 (2003).
 33. Sharara A., Perkins D., Misukonis M., Chan S., Dominitz J., Weinberg J. Interferon (IFN)- α activation of human blood mononuclear cells in vitro and in vivo for nitric oxide synthase (NOS) type 2 mRNA and protein expression: possible relationship of induced NOS2 to the anti-hepatitis C effects of IFN-R in vivo. *J. Exp. Med.*, **186**, 1495-1502 (1997).
 34. Di Bisceglie A., Hoofnagle J. Optimal therapy of hepatitis C. *Hepatology*, **36**, S121–S127 (2002).
 35. Zein N. Clinical significance of hepatitis C virus genotypes. *Clin. Microbiol. Rev.* **13**, 223-235 (2000).
 36. McHutchison J., Poynard T. Combination therapy with interferon plus ribavirin for the initial treatment of chronic hepatitis C. *Semin. Liver Dis.* **19** (Suppl. 1), 57–65 (1999).
 37. Strader D. Understudied populations with hepatitis C. *Hepatology* **36**, S226–S236 (2002).
 38. Scott L., Perry C. Interferon-R-2b plus ribavirin: A review of its use in the management of chronic hepatitis C. *Drugs* **62**, 507-556 (2002).
 39. Huang Z., Murray M., Secrist J., recent development of therapeutics for chronic HCV infection, *Antiviral research*, **71**, 351-362 (2006).
 40. Crotty S., Cameron C., Andino R. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl Acad. Sci. USA* **98**, 6895–6900 (2001).
-

-
41. Crotty S. *et al.* The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nature Med.* **6**, 1375–1379 (2000).
 42. Glue P. *et al.* Pegylated interferon-alpha2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. Hepatitis C Intervention Therapy Group. *Clin. Pharmacol. Ther.* **68**, 556–567 (2000).
 43. Zeuzem S. *et al.* PEG interferon α -2a in patients with chronic hepatitis C. *New Engl. J. Med.* **343**, 1666–1672 (2000).
 44. Heathcote E. *et al.* PEG interferon α -2a in patients with chronic hepatitis C and cirrhosis. *New Engl. J. Med.* **343**, 1673–1680 (2000).
 45. Lindsay K. *et al.* A randomized, double blind trial comparing PEGylated interferon alfa-2b to interferon α -2b as initial treatment for chronic hepatitis C. *Hepatology* **34**, 395–403 (2001).
 46. Manns M. *et al.* PEG interferon α -2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* **358**, 958–965 (2001).
 47. Fried M. *et al.* PEG interferon α -2a plus ribavirin for chronic hepatitis C virus infection. *New Engl. J. Med.* **347**, 975–982 (2002).
 48. Hadziyannis S. *et al.* PEG interferon α -2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann. Intern. Med.* **140**, 346–355 (2004).
 49. Davis G., Wong J., McHutchison J., Manns M., Harvey J., Albrecht J. Early virologic response to treatment with peg interferon α -2b plus ribavirin in patients with chronic hepatitis C. *Hepatology* **38**, 645–652 (2003).
 50. Medina J., Garcia-Buey L., Moreno-Monteagudo J., TraperoMarugan M., Moreno-Otero R. Therapeutic advantages of pegylation of interferon alpha in chronic hepatitis C. *Rev. Esp. Enferm. Apar. Dig.* **95**, 561–574 (2003).
 51. Perez-Olmeda M., Nunez M., Romero M., Gonzalez J., Castro A., Arribas J., Pedreira J., Barreiro P., Garcia-Samaniego J., Martin-Carbonero L., Jimenez-Nacher I., Soriano V. Pegylated IFN-R2b plus ribavirin as therapy for chronic hepatitis C in HIV-infected patients. *AIDS* **17**, 1023–1028, (2003).
 52. Matsuura Y. Expression and characterization of hepatitis C virus (HCV) proteins. *Virusu* **45**, 105–115 (1995).
 53. <http://www.expertreviews.org/> expert reviews in molecular medicine, Vol. **5**; 19 November 2003 Cambridge University Press.
-

-
54. Westaway E. Flavivirus replication strategy. *Adv Virus Res.*; **33**: 45–90 (1987).
 55. Forns X. and Bukh J. The Molecular Biology of Hepatitis C Virus: Genotypes and Quasispecies. *Clinics in Liver Disease*, **3**, 693-716 (1999).
 56. Tellinghuisen T., Evans M., Hahn T., You S., Rice C, Studying Hepatitis C Virus: Making the Best of a Bad Virus, *Journal of Virology*, **81** (17), 8853–8867 (2007).
 57. Tan S., Pause A., Shi Y. and Sonenberg N. Hepatitis C Therapeutics: current status and emerging strategies. *Nature Reviews Drug Discovery* **1**; 867-881 (2002).
 58. Bartenschlager R. *et al.* Novel insights into hepatitis C virus replication and persistence. *Adv. Virus Res.* **63**, 71–180 (2004).
 59. Griffin S. *et al.* The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.* **535**, 34–38 (2003).
 60. Pavlovic D. *et al.* The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain imino sugar derivatives. *Proc. Natl Acad. Sci. USA* **100**, 6104–6108 (2003).
 61. Bartenschlager R., Lohmann V. Replication of hepatitis C virus. *J. Gen. Virol.* **81**, 1631–1648 (2000).
 62. Ogata N., Alter H., Miller R., Purcell R. Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc. Nat. Acad. Sci. U.S.A.* **88**, 3392-3396 (1991).
 63. Egger D. *et al.* Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J. Virol.* **76**, 5974–5984 (2002).
 64. Kim D., Gwack Y., Han J., Choe J. C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem. Biophys. Res. Commun.* **215**, 160–166 (1995).
 65. Suzich J. *et al.* Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with the related pestivirus and flavivirus enzymes. *J. Virol.* **67**, 6152–6158 (1993).
-

-
66. Gwack Y., Kim D., Han J., Choe J. Characterization of RNA binding activity and RNA helicase activity of the hepatitis C virus NS3 protein. *Biochem. Biophys. Res. Commun.* **225**, 654–659 (1996).
 67. Tai C., Chi W., Chen D., Hwang L. The helicase activity associated with hepatitis C virus non structural protein 3 (NS3). *J. Virol.* **70**, 8477–8484 (1996).
 68. Behrens S., Tomei L., De F. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15**, 12–22 (1996).
 69. Lohmann V., Korner F., Herian U., Bartenschlager R. Biochemical properties of hepatitis C virus NS5B RNA dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J. Virol.* **71**, 8416–8428 (1997).
 70. Qureshi S., Hepatitis C Virus-Biology, Host Evasion Strategies, and Promising New Therapies on the Horizon, *Medicinal Research Reviews*, **27**, 353-373 (2006).
 71. Houghton M., Abrignanis S. Prospects for a vaccine against the hepatitis C virus *Nature* **436**, 961-966 (2005).
 72. Folgori A. *et al.* A T-cell HCV vaccine eliciting effective immunity against heterologous virus challenge in chimpanzees. *Nature Med.* **12**,190–197 (2006).
 73. Nevens F. *et al.* A pilot study of therapeutic vaccination with envelope protein E1 in 35 patients with chronic hepatitis C. *Hepatology* **38**, 1289–1296 (2003).
 74. Leroux-Roels, G. *et al.* Immunogenicity and tolerability of intradermal administration of an HCV E1-based vaccine candidate in healthy volunteers and patients with resolved or ongoing chronic HCV infection. *Hum. Vaccin.* **1**, 61–65 (2005).
 75. Manns M. *et al.* Immunization with the therapeutic hepatitis C virus peptide vaccine IC41 in 66 chronic hepatitis C non-responder patients. *Hepatology* **40** (Suppl. 1), 251A (2004).
 76. Klade C. *et al.* Therapeutic peptide vaccination against chronic hepatitis C virus infection. *J. Hepatol.* **46** (Suppl. 1), S229 (2007).
 77. Polakos N., *et al.* Characterization of hepatitis C virus core-specific immune responses primed in rhesus macaques by a non classical ISCOM vaccine. *J. Immunol.* **166**, 3589–3598 (2001).
-

-
78. Li Y., Kang H., Babiuk L., Liu Q., Elicitation of strong immune responses by a DNA vaccine expressing a secreted form of hepatitis C virus envelope protein E2 in murine and porcine animal models. *World J. Gastroenterol.* **12**, 7126–7135 (2006).
79. Chapel C. *et al.* Reduction of the infectivity of hepatitis C virus pseudoparticles by incorporation of misfolded glycoproteins induced by glucosidase inhibitors. *J. Gen Virol.* **88**, 1133–1143 (2007).
80. Reed K., Rice C. Expression and characterization of the HCV NS2 protease. *Method Mol. Med.*, **19**, 331–342 (1998).
81. Pallaoro M., Lahm A., Biasiol G., Brunetti, M., Nardella C., Orsatti L., Bonelli F., Orru S., Narjes F., Steinkuhler C. Characterization of the hepatitis C virus NS2/3 processing reaction by using a purified precursor protein. *J. Virol.*, **75**, 9939–9946 (2001).
82. Lamarre D. *et al.* An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature* **426**, 186–189 (2003).
83. Lawitz E. *et al.* 28 Days of the hepatitis C protease inhibitor VX-950, in combination with PEG-Interferon- α -2a and ribavirin, is well-tolerated and demonstrates robust antiviral effects. *Gastroenterology* **131**, 950–951 (2006).
84. Hinrichsen H. *et al.* Short-term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients. *Gastroenterology* **127**, 1347–1355 (2004).
85. Reiser M. *et al.* Antiviral efficacy of NS3-serine protease inhibitor BILN-2061 in patients with chronic genotype 2 and 3 hepatitis C. *Hepatology* **41**, 832–835 (2005).
86. Sarrazin C. *et al.* SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon α -2b for genotype 1 non responders. *Gastroenterology* **132**, 1270–1278 (2007).
87. Pawlotsky J., McHutchison J. Hepatitis C. Development of new drugs and clinical trials: promises and pitfalls. Summary of an AASLD hepatitis single topic conference, Chicago, IL, February 27-March 1, 2003. *Hepatology* **39**, 554–567 (2004).
88. Wilkinson T. Hepatitis C virus: Prospects for future therapies. *Curr. Opin. Invest. Drugs*, **2**, 1516–1522 (2001).
-

-
89. Olsen D. *et al.* HCV antiviral activity and resistance analysis in chronically infected chimpanzees treated with NS3/4A protease and NS5B polymerase inhibitors. *J. Hepatol.* **46** (Suppl. 1), S298 (2007).
90. Roberts S. *et al.* Interim results of a multiple ascending dose study of R1626, a novel nucleoside analog targeting HCV polymerase in chronic HCV patients. *J. Hepatol.* **44** (Suppl. 2), S269 (2006).
91. Afdhal N. *et al.* Valopicitabine (NM283), alone or with peg-interferon, compared to peg-interferon/ribavirin (pegIFN/RBV) retreatment in hepatitis C patients with prior non-response to peg IFN/RBV: week 24 results. *J. Hepatol.* **44** (Suppl. 2), S19 (2006).
92. Dieterich D. *et al.* Early clearance of HCV RNA with valopicitabine (NM283) plus peg-interferon in treatment-naive patients with HCV-1 infection: first results from a Phase IIb trial. *J. Hepatol.* **44** (Suppl.2), S271 (2006).
93. McHutchison J., Bartenschlager R., Patel K. Pawlotsky J. The face of future hepatitis C antiviral drug development: recent biological and virologic advances and their translation to drug development and clinical practice. *J. Hepatol.* **44**, 411–421 (2006).
94. McHutchison J. *et al.* Phase 1B, randomized, double-blind, dose-escalation trial of CPG 10101 in patients with chronic hepatitis C virus. *Hepatology* **46**, 1341–1349 (2007).
95. Goto K. *et al.* Evaluation of the anti-hepatitis C virus effects of cyclophilin inhibitors, cyclosporin A, and NIM811. *Biochem. Biophys. Res. Commun.* **343**, 879–884 (2006).
96. Ma S. *et al.* NIM811, a cyclophilin inhibitor, exhibits potent *in vitro* activity against hepatitis C virus alone or in combination with α interferon. *Antimicrob. Agents Chemother.* **50**, 2976–2982 (2006).
97. Flisiak R. *et al.* The cyclophilin inhibitor DEBIO-025 has a potent dual anti-HIV and anti-HCV activity in treatment-naive HIV/HCV co-infected subjects. *National AIDS Treatment Advocacy Project web site* [online], <http://www.natap.org/2006/AASLD/_30.htm> (2006).
98. Pottage J. *et al.* Short-term antiviral activity and safety of ACH-806 (GS-9132), an NS4A antagonist, in HCV genotype 1 infected individuals. *J. Hepatol.* **46** (Suppl. 1), S294 (2007).
99. Melcher S., Wilson T., Lilley D. The dynamic nature of the four-way junction of the hepatitis C virus IRES. *RNA*, **9**, 809-820 (2003).
-

-
100. Buratti E., Gerotto M., Pontisso P., Alberti A., Tisminetzky S., Baralle F. In vivo translational efficiency of different hepatitis C virus 5'-UTRs. *FEBS Lett.*, **411**, 275-280 (1997).
 101. McHutchison J. *et al.* A phase I trial of an antisense inhibitor of hepatitis C virus (ISIS 14803), administered to chronic hepatitis C patients. *J. Hepatol.* **44**, 88-96 (2006).
 102. Jarczak D., Korf M., Beger C., Manns M., Kruger M. Hairpin ribozymes in combination with siRNAs against highly conserved hepatitis C virus sequence inhibit RNA replication and protein translation from hepatitis C virus subgenomic replicons. *FEBS J.* **272**, 5910-5922 (2005).
 103. Kanda T., Steele R., Ray R., Ray R. Small interfering RNA targeted to hepatitis C virus 5' nontranslated region exerts potent antiviral effect. *J. Virol.*, **81**, 669-676 (2007).
 104. Guang X., Helicases as Antiviral and Anticancer Drug Targets, *Current Medicinal Chemistry*, **14**, 883-915 (2007).
 105. Lohman T., Bjornson K. Mechanisms of helicase catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169-214 (1996).
 106. Soutanas P., Wigley D. Unwinding the 'Gordian knot' of helicase action. *Trends Biochem. Sci.* **26**, 47-54 (2001).
 107. Singleton M., Wigley D. Modularity and specialization in superfamily 1 and 2 helicases. *J. Bacteriol.* **184**, 1819-1826 (2002).
 108. Levin M., Patel S. in *Molecular Motors* (ed. Schliwa, M.) 179-198 (Wiley-VCH Weinheim (2002).
 109. Gorbalenya A., Koonin E. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429 (1993).
 110. Kwong A., Govinda B., Jeang K. Viral and Cellular RNA helicases as antiviral targets, *Nature Reviews Drug Discovery*, **4** (10), 845-53 (2005).
 111. Caruthers J., McKay D. Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123-13 (2002).
 112. Subramanya H., Bird L., Brannigan J., Wigley D. Crystal structure of a DExx box DNA helicase. *Nature*, **384**, 379-383 (1996).
-

113. Walker J., M., S., Runswick M., Gay N. Distantly related sequences in the α - and β -subunits of ATP synthetase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951 (1982).
 114. Daune M., Duffin W. J. *Molecular physics: Structures in motion*. Oxford University Press: New York, (1999).
 115. Levin M., Gurjar M., Patel S. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase *Nat. Struct. Mol. Biol.*, **12**, 429 (2005).
 116. Kim J., Morgenstern K., Griffith J., Dwyer M., Thomson J., Murcko M., Lin C., Caron P. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure*, **6**, 89-100 (1998).
 117. Bianco P., Brewer L., Corzett M., Balhorn R., Yeh Y., Kowalczykowski S., Baskin R. Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature*, **409**, 374-378 (2001).
 118. Breyer W., Matthews B. A structural basis for processivity. *Protein Sci.*, **10**, 1699 (2001).
 119. Kadare G., Haenni A. Virus-encoded RNA helicases. *J. Virol.*, **71**, 2583-2590 (1997).
 120. Kleymann G. *et al.* New helicase-primase inhibitors as drug candidates for the treatment of herpes simplex disease. *Nature Med.* **8**, 392–398 (2002).
 121. Crumpacker C., Schaffer P. New anti-HSV therapeutics target the helicase-primase complex. *Nature Med.* **8**, 327–328 (2002).
 122. Jones P. Strategies for antiviral drug discovery. *Antivir. Chem. Chemother.* **9**, 283–302 (1998).
 123. Frick D. Helicases as antiviral drug targets. *Drug News Perspect.* **16**, 355–362 (2003).
 124. Gu B., Liu C., Lin-Goerke J., Maley D., Gutshall L., Fletenberger C., Del Vecchio A. The RNA helicase and nucleotide triphosphatase activities of the bovine viral diarrhea virus NS3 protein are essential for viral replication. *J. Virol.*, **74**, 1794-1800 (2000).
 125. Matusan A., Pryor M., Davidson A., Wright P. Mutagenesis of the Dengue Virus Type 2 NS3 Protein within and outside Helicase Motifs: Effects on Enzyme Activity and Virus Replication. *J. Virol.*, **75**, 9633-9643 (2001).
-

-
126. Borowski P., Schalinski S., Schmitz H. Nucleotide triphosphatase/ helicase of hepatitis C virus as a target for antiviral therapy. *Antiviral Res.*, **55**, 397-412 (2002).
 127. Samuel C. Antiviral actions of interferons. *Clin. Microbiol. Rev.* **14**, 778–809 (2001).
 128. Yoneyama M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunol.* **5**, 730–737 (2004).
 129. Spector F., Liang L., Giordano H., Sivaraja M., Peerson M. T157602, a 2-amino-thiazole inhibits HSV replication by interacting with the UL5 component of the UL5/8/52 helicase primase complex. *Antiviral Res.* **37**, A43 (1998).
 130. Spector F., Liang L., Giordano H., Sivaraja M., Peterson M. Inhibition of herpes simplex virus replication by a 2-amino thiazole via interactions with the helicase component of the UL5–UL8–UL52 complex. *J. Virol.* **72**, 6979–6987 (1998).
 131. Crute J. *et al.* Inhibition of herpes simplex virus type 1 helicase–primase by (dichloroanilino) purines and pyrimidines. *J. Med. Chem.* **38**, 1820–1825 (1995).
 132. Betz U., Fischer R., Kleymann G., Hendrix M., Rubsamen-Waigmann H. Potent in vivo antiviral activity of the herpes simplex virus primase–helicase inhibitor BAY 57–1293. *Antimicrob. Agents Chemother.* **46**, 1766–1772 (2002).
 133. Faucher A., White P., Brochu C., Grand-Maitre C., Rancourt J., Fazal G. Discovery of Small-Molecule Inhibitors of the ATPase Activity of Human Papillomavirus E1 Helicase *J. Med. Chem.*, **47**, 18-21 (2004).
 134. Tanner J., Zheng B., Zhou J., Watt R., Jiang J., Wong K., Lin Y., Lu L., He M., Kung H., Kesel A., Huang J. The Adamantane-Derived Bananins Are Potent Inhibitors of the Helicase Activities and Replication of SARS Coronavirus *Chem. Biol.*, **12**, 303-311 (2005).
 135. Borowski P., Deinert J., Schalinski S., Bretner M., Ginalski K., Kulikowski T., Shugar D. Halogenated benzimidazoles and benzotriazoles as inhibitors of the NTPase/helicase activities of hepatitis C and related viruses. *Eur. J. Biochem.*, **270**, 1645-1653 (2003).
 136. Bretner M., Baier A., Kopanska K., Najda A., Schoof A., Reinholz M., Lipniacki A., Piasek A., Kulikowski T., Borowski P., Synthesis and biological
-

- activity of 1 H-benzotriazole and 1 H-benzimidazole analogues. *Antivir. Chem. Chemother.*, **16**, 315 (2005).
137. Maga G., Gemma S., Fattorusso C., Locatelli G., Butini S., Persico M., Kukreja G., Romano M., Chiasserini L., Savini L., Novellino E., Nacci, Spdari S., Campiani G. Specific targeting of hepatitis C virus NS3 RNA helicase. Discovery of the potent and selective competitive nucleotide-mimicking inhibitor QU633. *Biochemistry*, **44**, 9637-9644 (2005).
138. Frick D. The Hepatitis C Virus NS3 Protein: A Model RNA Helicase and potential Drug Target. *Curr. Issues Mol. Biol.* **9**, 1-20 (2007).
139. Li H., Clum S., You S., Ebner K., 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.* **73**, 3108-3116 (1999).
140. Suzich J., Tamura J., Palmer H., Warrener P., Grakoui A., Rice C., Feinstone S., Collett M. Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with the related pestivirus and flavivirus enzymes. *J. Virol.* **67**, 6152-6158 (1993).
141. Tamura J., Warrener P., Collett M. RNA-stimulated NTPase activity associated with the p80 protein of the pestivirus bovine viral diarrhea virus. *Virology*, **193**, 1-10 (1993).
142. Warrener P., Tamura J., Collett M. RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. *J. Virol.* **67**, 989-996 (1993).
143. Yao N., Hesson T., Cable M., Hong Z., Kwong A., Le H., Weber P. Structure of the hepatitis C virus RNA helicase domain. *Nat. Struct. Biol.*, **4**, 463-467 (1997).
144. Cho H., Ha N., Kang L., Chung K., Bac S., Jang S., Oh B. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. *J. Biol. Chem.*, **273**, 15045-15052 (1998).
145. Yao N., Reichert P., Taremi S., Prosis W., Weber P. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure Fold. Des.* **7**, 1353-1363 (1999).
146. Mackintosh S., Lu J., Jordan J., Harrison M., Sikora B., Sharma S., Cameron C., Raney K., Sakon J. Structural and biological identification of residues on the surface of NS3 helicase required for optimal replication of the hepatitis C virus. *J. Biol. Chem.* **281**, 3528-3535 (2006).
-

-
147. Soultanas P., Dillingham M., Velankar S., Wigley D. DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J. Mol. Biol.* **290**, 137–148 (1999).
 148. Velankar S., Soultanas P., Dillingham M., Subramanya H., Wigley D. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84 (1999).
 149. Bernstein D., Zittel M., Keck J. High resolution structure of the E.coli RecQ helicase catalytic core. *EMBO J.* **22**, 4910–4921 (2003).
 150. Bird L., Subramanya H., Wigley D. Helicases: a unifying structural theme? *Curr. Opin. Struct. Biol.*, **8**, 1418 (1998).
 151. Theis K., Chen P., Skorvaga M., Van Houten B., Kisker C. Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *EMBO J.*, **18**, 6899–6907 (1999).
 152. Gorbalenya A., Kunin E. Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res.*, **17**, 8413–8440 (1989).
 153. Walker J., Saraste M., Runswick M., Gay N., Distantly related sequences in the α - and β -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.*, **1**, 945–951 (1982).
 154. Tai C., Pan W., Liaw S., Yang U., Hwang L., Chen D. Structure-based mutational analysis of the hepatitis C virus NS3 helicase. *J. Virol.*, **75**, 8289–8297 (2001).
 155. Du M., Johnson R., Sun K., Staschke C., Wang Q. Comparative characterization of two DEAD-box RNA helicases in superfamily II: human translation-initiation factor 4A and hepatitis C virus non-structural protein 3 (NS3) helicase. *Biochem. J.*, **363**, 147–155 (2002).
 156. Borowski P., Mueller O., Niebuhr A., Kalitzky M., Hwang L., Schmitz H., Siwecka M., Kulikowski T. ATP-binding domain of NTPase/helicase as a target for hepatitis C antiviral therapy. *Acta Biochim. Pol.*, **47**, 173–180 (2000).
 157. Zhang N., Chen H., Koch V., Schmitz H., Liao C., Bretner M., Bhadti V., Fattom A., Naso R., Hosmane R., Borowski P. Ring-expanded (“fat”) nucleoside and nucleotide analogues exhibit potent in vitro activity against flaviviridae NTPases/helicases, including those of the West Nile virus, hepatitis C virus, and Japanese encephalitis virus. *J. Med. Chem.*, **46**, 4149–4164 (2003).
-

-
158. Kuang W., Lin Y., Jean F., Huang Y., Tai C., Chen D., Chen P., Hwang L. Hepatitis C virus NS3 RNA helicase activity is modulated by the two domains of NS3 and NS4A. *Biochem. Biophys. Res. Commun.*, **317**, 211-217 (2004).
159. Lun L., Sun P., Trubey C., Bachur N. Antihelicase action of CI-958, a new drug for prostate cancer. *Cancer Chemother. Pharmacol.*, **42**, 447-453 (1998).
160. Zhu K., Henning D., Iwakuma T., Valdez B., Busch H. Adriamycin inhibits human RH II/Gu RNA helicase activity by binding to its substrate. *Biochem. Biophys. Res. Commun.*, **266**, 361-365 (1999).
161. Bachur N., Yu F., Johnson R., Hickey R., Wu Y., Malkas L. Helicase inhibition by anthracycline anticancer agents. *Mol. Pharmacol.*, **41**, 993-998 (1992).
162. a) Diana G., Bailey T., US5633388, 1997. b) Diana G., Bailey T., Nitz T. WO9736554, 1997.
163. Phoon C., Ng P., Ting A., Yeo S., Sim M. Biological evaluation of hepatitis C virus helicase inhibitors. *Bioorg. Med. Chem. Lett.*, **11**, 1647-1650 (2001).
164. Boguszewska A., Krawczyk M., Najda Z., Kopanska K., Stankiewicz A., Ostoja W., Bretner M. Searching for a new anti-HCV therapy: Synthesis and properties of tropolone derivatives, *Biochemical and Biophysical Research Communications* **341**, 641-647 (2006).
165. Stankiewicz A., Palchykovska L., Kostina V., Alexeeva I., Shved A., Boguszewska A. New acridone-4-carboxylic acid derivatives as potential inhibitors of *Hepatitis C virus* infection, *Bioorganic and Medicinal chemistry*, **16** (19), 8846-8852 (2008).
166. McHutchison J., Fried M. Current therapy for hepatitis C: pegylated interferon and ribavirin. *Clin. Liver Dis.* **7**, 149-161 (2003).
167. Pawlotsky J. Mechanisms of antiviral treatment efficacy and failure in chronic hepatitis C. *Antiviral Res.* **59**, 1-11 (2003).
168. Strader D., Wright T., Thomas D., Seeff L. Diagnosis, management, and treatment of hepatitis C. *Hepatology* **39**, 1147-1171 (2004).
169. Chisari F., Unscrambling hepatitis C virus-host interactions. *Nature*, **436** (18), 930-932 (2005).
170. Borden E., Sen G., Uze G., Silverman R., Ransohoff R., Foster G., Stark G. Interferons at age 50: past, current and future impact on biomedicine. *Nature reviews drug discovery*, **6**, 975 - 990 (2007).
-

-
171. Boyd D. drug design. In *Encyclopedia of Medicinal Chemistry*. Schleyer P., Allinger N., Clark T., Gastiger J., Kollman P., Schaefer H. and Schreiner P. Eds.; John Wiley and sons Inc.: Chichester, UK, 1998; pp 795-804.
172. Richon A., Current Status and Future direction of the molecular modeling industry, *Drug Discovery Today*, **13**, 15/16, 665-669 (2008).
173. Leach A., *Molecular Modelling: Principles and Applications*; Prentice Hall: Harlow, 2001.
174. Kitchen D., Decornez H., Furr J., Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature* **3**, 935-949 (2004).
175. Goodman J., *Chemical applications of molecular modelling*, the royal Society of Chemistry: Cambridge, 1998.
176. Allinger N. Force fields: A brief introduction. In *Encyclopedia of medicinal chemistry*. Schleyer P. Eds. : John Wiley and Sons inc.: Chichester, UK, 1998 (1013-1015).
177. Grant G. and Richards W. *Computational Chemistry*; oxford university Press: oxford, 1995.
178. 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. Am. Chem. Soc.*, **117**, 5179-5197 (1995).
179. Halgren T. Merck molecular force field. 1. Force, form, scope, parameterisation and performance of MMFF94. *J. Comp. Chem.*, **17**, 490-519 (1996).
180. Kuntz I., Blaney J., Oatley S., Langridge R., Ferrin T. A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* **161**, 269–288 (1982).
181. Alexandre M., Bonvin J. Flexible protein-protein docking. *Current Opinion in Structural Biology*, **16** (2), 194-200 (2006).
182. Venhorst J. *et al.* Homology modeling of rat and human cytochrome P450 2D (CYP2D) isoforms and computational rationalization of experimental ligand-binding specificities. *J. Med. Chem.* **46**, 74–86 (2003).
-

-
183. Williams P. *et al.* Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **424**, 464–468 (2003).
 184. Krüger 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; pp 443-476.
 185. Leach A., Shoichet B., Peishoff C. Docking and Scoring Perspective; Prediction of Protein-Ligand Interactions. Docking and Scoring: Successes and Gaps, **49** (20), 251-255 (2006).
 186. Taylor R., Jewsbury P., Essex J. A review of protein-small molecule docking methods. *J Comput Aided Mol Des*, **16** (3), 151-166 (2002).
 187. DesJarlais R. Docking flexible ligands to macromolecular receptors by shape. *J. Med Chem.* **29**, 2149–2153 (1986).
 188. Klebe G. & Rarey M. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489 (1996).
 189. Kuntz I. & Leach A. Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comput. Chem.* **13**, 730–748 (1992).
 190. Rarey M., Kramer B., Lengauer T., Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*, **261**(3), 470-489 (1996).
 191. Olson A. & Goodsell D. Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins* **17**, 1–10 (1993).
 192. Verdonk M., Cole J., Hartshorn M., Murray C., Taylor R. Improved protein-ligand docking using GOLD. *Proteins*, **52**(4), 609-23 (2003).
 193. Dixon J. & Oshiro C. Flexible ligand docking using a genetic algorithm. *J. Comput. Aided Mol. Des.* **9**, 113–130 (1995).
 194. Morris G., Goodsell D., Halliday R., Huey R., Hart W. Automated docking using a Lamarckian genetic algorithm and an empirical free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
 195. Jones G., Willet P., Glen R., Leach A. & Taylor R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748 (1997).
-

-
196. Westhead D., Clark D., Murray C. A comparison of heuristic search algorithms for molecular docking. *J. Comput. Aided Mol. Des.* **11**, 209–228 (1997).
197. Baxter C., Murray C., Clark D., Westhead D., Eldridge M. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins* **33**, 367–382 (1998).
198. Baxter C., Murray C., Clark D., Westhead R., Eldridge M. Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins*, **33** (3), 367-382 (1998).
199. Brooijmans N., Kuntz I. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 335–373 (2003).
200. Di Nola A., Berendsen H., Roccatano D. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins* **19**, 174–182 (1994).
201. Trosset J., Scheraga H. Reaching the global minimum in docking simulations: a Monte Carlo energy minimization approach using Bezier Splines. *Proc. Natl. Acad. Sci. USA* **95**, 8011–8015 (1995).
202. Rester U., Dock around the Clock, current status of small molecule docking and scoring. *QSAR & Combinatorial Science*, **25**, 605-615 (2006).
203. Meek P., Liu Z., Tian L., Wang C., Welsh W., Zauhar R., Shape Signatures: speeding up computer aided drug discovery, *Drug discovery today*, **11**, 895-904 (2006).
204. Carlson H., McGammon J. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* **57**, 213–218 (2000).
205. Read R., Hart T. A multiple-start Monte Carlo docking method. *Proteins* **13**, 206–222 (1992).
206. Leach A. Ligand docking to proteins with discrete sidechain flexibility. *J. Mol. Biol.* **235**, 245–356 (1994).
207. Desmet J., Maeyer M., Hazes B., Lasters I. The dead end elimination theorem and its use in protein side-chain positioning. *Nature* **356**, 539–542 (1992).
208. Knegt R., Kuntz I., Oshiro C. Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **266**, 242–440 (1997).
-

-
209. Goodford P. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **28**, 849–857 (1985).
210. Mohan V., Gibbs A., Cummings M., Jaeger E., DesJarlais R. Docking: Successes and Challenges, *Current Pharmaceutical Design.* **11**, 323–333 (2005).
211. 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.* **9**, 13–32 (1995).
212. Murray C. *et al.* PRO_SELECT: combining structure based drug design and combinatorial chemistry for rapid lead discovery. 1. Technology. *J. Comp. Aided Mol. Des.* **11**, 193–207 (1997).
213. Bohacek R., McMartin C. Multiple highly diverse structures complementary to enzyme binding sites: results of extensive application of a *de novo* design method incorporating combinatorial growth. *J. Am. Chem. Soc.* **116**, 5560–5571 (1994).
214. Wang R., Gao Y., Lai L. LigBuilder: a multi-purpose program for structure-based drug design. *J. Mol. Model.* **6**, 498–516 (2000).
215. Pearlman D., Murcko M. CONCEPTS: new dynamic algorithm for *de novo* design suggestion. *J. Comput. Chem.* **14**, 1184–1193 (1993).
216. 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.* **11**, 425–445 (1997).
217. Charifson P., Corkery J., Murcko M., Walters W. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **42**, 5100–5109 (1999).
218. Irwin J., Scoichet B., ZINC- A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.*, **45**, 177–182 (2005).
219. Masciocchi J, Frau G., Fanton M., Sturlese M., Ioris M., Pireddu L., Palla P., Cedrati F., Rodriguez P., Moro S. MMsINC: a large-scale chemoinformatics database, *Nucleic Acids Research*, 1–7 (2008).
220. Lipiniski C., Lombardo F., Dominy B., Feeny P., experimental and computational approaches to estimate solubility and permeability in drug
-

- discovery and development settings. *Advanced drug Delivery Reviews*, **23**, 3-25 (1997).
221. Wermuth C., Gannelin C., Lindberg P., Mitscher L. Glossary of terms used in medicinal chemistry. *Pure Appl. Chem.* **70**, 1129–1143 (1998).
222. Schneider G., Fechner U. Computer-Based De Novo Design of Drug-Like molecules, *Nature Reviews Drug Discovery*, **4**, 649-663 (2005).
223. Dobson C. Chemical space and biology. *Nature* **432**, 824–828 (2004).
224. Lipinski C., Hopkins A. Navigating chemical space for biology and medicine. *Nature* **432**, 855–861 (2004).
225. Schneider G. Trends in virtual combinatorial library design. *Curr. Med. Chem.* **9**, 2095–2101 (2002).
226. Lewis R. *et al.* Automated site-directed drug design using molecular lattices. *J. Mol. Graphics* **10**, 66–78 (1992).
227. Roe D., Kuntz I. BUILDER v.2: improving the chemistry of a *de novo* design strategy. *J. Comput. Aided Mol. Des.* **9**, 269–282 (1995).
228. Tschinke V., Cohen N. The NEWLEAD program: a new method for the design of candidate structures from pharmacophoric hypothesis. *J. Med. Chem.* **36**, 3863–3870 (1993).
229. Teague S. *et al.* The design of lead like combinatorial libraries. *Angew. Chem. Int. Ed. Engl.* **38**, 3743–3747 (1999).
230. Lewis R. Automated site-directed drug design: approaches to the formation of 3D molecular graphs. *J. Comput. Aided Mol. Des.* **4**, 205–210 (1990).
231. Pearlman D., Murcko M. CONCERTS: dynamic connection of fragments as an approach to *de novo* ligand design. *J. Med. Chem.* **39**, 1651–1663 (1996).
232. Liu H., Duan Z., Luo Q., Shi Y. Structure-based ligand design by dynamically assembling molecular building blocks at binding site. *Proteins* **36**, 462–470 (1999).
233. Zhu J., Yu H., Fan H., Liu H., Shi Y. Design of selective inhibitors of cyclooxygenase-2 dynamic assembly of molecular building blocks. *J. Comput. Aided Mol. Des.* **15**, 447–463 (2001).
-

-
234. Zhu J., Fan H., Liu H., Shi Y. Structure-based ligand design for flexible proteins: application of new F-DycoBlock. *J. Comput. Aided Mol. Des.* **15**, 979–996 (2001).
235. Zonta N., Grimstead J., Avis N., Brancale A., Accessible haptic technology for drug design applications. *J. Mol. Model.* **15** (2), 193–6 (2009).
236. Moraitakis G., Purkiss A., Goodfellow J. Simulated dynamics and biological macromolecules. *Rep. Prog. Phys.* **66** 383–406 (2003).
237. Adcock S., McCammon J. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins, *Chem. Rev.*, **106**, 1589–1615 (2006).
238. Waszkowycz B. *et al.* PRO LIGAND: an approach to *de novo* molecular design. 2. design of novel molecules from molecular field analysis (MFA) models and pharmacophores. *J. Med. Chem.* **37**, 3994–4002 (1994).
239. Nachbar R. Molecular evolution: automated manipulation of hierarchical chemical topology and its application to average molecular structures. *Genet. Programming Evolvable Machines* **1**, 57–94 (2000).
240. Pellegrini E., Field M. J. Development and testing of a *de novo* drug-design algorithm. *J. Comp. Aided Mol. Des.* **17**, 621–641 (2003).
241. Douguet D., Thoreau E., Grassy G. A genetic algorithm for the automated generation of small organic molecules: drug design using an evolutionary algorithm. *J. Comput. Aided Mol. Des.* **14**, 449–466 (2000).
242. Schneider G., Lee M., Stahl M., Schneider P. *De novo* design of molecular architectures by evolutionary assembly of drug-derived building blocks. *J. Comput. Aided Mol. Des.* **14**, 487–494 (2000).
243. Globus A., Lawton J., Wipke W. Automatic Molecular design using evolutionary algorithms. *Nanotechnology* **10**, 290–299 (1999).
244. Brown N., McKay B., Gilardoni F., Gasteiger J. A graph-based genetic algorithm and its application to the multiobjective evolution of median molecules. *J. Chem. Inf. Comput. Sci.* **44**, 1079–1087 (2004).
245. Liu D., Wang Y., Gesell J., Wyss D. Solution structure and backbone dynamics of an engineered arginine-rich subdomain 2 of the hepatitis C virus NS3 RNA helicase. *J. Mol. Biol.*, **314**, 543–561 (2001).
246. Liu D., Windsor W., Wyss D. Double stranded DNA-induced localized unfolding of HCV NS3 helicase subdomain 2. *Protein Sci.* **12**, 2757–2767 (2003).
-

-
247. Lam A., Keeney D., Frick D. Two novel conserved motifs in the hepatitis C virus NS3 protein critical for helicase action. *J. Biol. Chem.* **278**, 44514–44524 (2003).
248. Jin L., Peterson D. Expression, isolation, and characterization of the hepatitis C virus ATPase/ RNA helicase. *Arch. Biochem. Biophys.* **323**, 47–53 (1995).
249. Kim D., Gwack Y., Han J., Choe J. Towards defining a minimal functional domain for NTPase and RNA helicase activities of the hepatitis C virus NS3 protein. *Virus Res.* **49**, 17–25 (1997).
250. Lin C., Kim J. Structure-based mutagenesis study of hepatitis C virus NS3 helicase. *J. Virol.* **73**, 8798–8807 (1999).
251. Preugschat F., Danger D., Carter L., 3rd, Davis R., Porter D. Kinetic analysis of the effects of mutagenesis of W501 and V432 of the hepatitis C virus NS3 helicase domain on ATPase and strand-separating activity. *Biochemistry*, **39**, 5174–5183 (2000).
252. Kim J., Seo M., Shelat A., Kim C., Kwon T., Lu H., Moustakas D., Sun J., Han J. Structurally conserved amino acid W501 is required for RNA helicase activity but is not essential for DNA helicase activity of hepatitis C virus NS3 protein. *J. Virol.* **77**, 571–582 (2003).
253. Frick D., Rypma R., Lam A., Frenz C. Electrostatic analysis of the hepatitis C virus NS3 helicase reveals both active and allosteric site locations. *Nucleic Acids Res.* **32**, 5519–5528 (2004).
254. Lam A., Rypma R., Frick D. Enhanced nucleic acid binding to ATP-bound hepatitis C virus NS3 helicase at low pH activates RNA unwinding. *Nucl. Acids Res.* **32**, 4060–4070 (2004).
255. Maria Chiara Barbera. Internal Report 2005. Welsh School of Pharmacy, Cardiff University.
256. a) Anna-Claire Cummins, Internal report, (2003) Welsh School of Pharmacy b) Dimitrios Vlachakis. Internal Report (2005). Welsh School of Pharmacy, Cardiff University.
257. Patrick G., *An introduction to medicinal chemistry*, third edition Oxford University press (2005).
258. Tripos SYBYL 7.1; Tripos Inc., 1699 South Hanley Rd, St. Louis, Missouri 63144, USA. <http://www.tripos.com>.
-

-
259. Molecular Operating Environment (MOE). Chemical Computing Group, Inc. Montreal, Quebec, Canada. <http://www.chemcomp.com>.
260. Raboisson P, Lenz O, Lin T *et al*, Evaluation of the antihepatitis C virus effect of novel potent, selective, and orally bioavailable JNK and VEGFR kinase inhibitors, *Bio. Med. Chem. Lett.*, **17**,1843-1849 (2007).
261. Donghi M., Ferrara M., Koch U., Narjes F., Ontoria O., Summa V., Quinazoline derivatives as antiviral agents, WO/2007/028789.
262. Herr J., 5-Substituted-1H-tetrazoles as carboxylic acid isosteres: medicinal chemistry and synthetic methods, *Bioorganic & Medicinal Chemistry* 10, 11, 3379-3393 (2002).
263. Tominey A., Docherty P., Rosair G., Quenardelle R., Kraft A., Unusually Weak Binding Interactions in Tetrazole-Amidine Complexes, *Org. Lett.*, **8**, 7, 1279-1282 (2006).
264. Rewcastle G., Palmer B., Bridges A., Showalter H., Sun L., Nelson J., McMichael A., Kraker A., Fry D., Denny W. Tyrosine Kinase Inhibitors. 9. Synthesis and Evaluation of Fused Tricyclic Quinazoline Analogues as ATP Site Inhibitors of the Tyrosine Kinase Activity of the Epidermal Growth Factor Receptor. *J. Med. Chem.*, **39**, 918-928 (1996).
265. Rewcastle G., Palmer B., Thompson A., Bridges A. Cody D., Zhou H., Fry D., McMichael A., Denny W. Tyrosine Kinase Inhibitors. 10. Isomeric 4-[(3-Bromophenyl)amino]pyrido[d]pyrimidines Are Potent ATP Binding Site Inhibitors of the Tyrosine Kinase Function of the Epidermal Growth Factor Receptor. *J. Med. Chem.*, **39**, 1823-1835 (1996).
266. Gibson K., Grundy W., Godfrey A., Woodburn J., Ashton S., Curry B. J., Scarlett L., Barker A., Brown D. Epidermal Growth Factor Receptor Tyrosine Kinase: Structure-Activity Relationships and Antitumour Activity of Novel Quinazolines. *Bioorg. Med. Chem. Lett.*, **7(21)**, 2723-2728 (1997).
267. Tsou H., Mamuya N., Johnson B., Reich M., Gruber B., Fei Ye, Nilakantan R., Ru Shen, Discafani C., DeBlanc R., Davis R., Koehn F., Greenberger L., Wang Y., Wissner A. 6-Substituted-4-(3-bromophenyl amino)quinazolines as Putative Irreversible Inhibitors of the Epidermal Growth Factor Receptor (EGFR) and Human Epidermal Growth Factor Receptor (HER-2) Tyrosine Kinases with Enhanced Antitumor Activity *J. Med. Chem.*, **44**, 2719-2734 (2001).
-

-
268. Gazit A., Levitzki A. Preparation of 4-anilino substituted quinazolines as inhibitors of epidermal growth factor receptor kinases. PCT Int. Appl., 85pp. WO 2004013091 (2004).
269. Solonskaya N., Bliznyukov V., Spectrophotometric examination of N-quinazolyl derivatives of aminobenzoic acid, *Farmatsevtichnii Zhurnal* (Kiev), **26 (3)**, 21-5 (1971).
270. Botros S., Ghoneim K., Khalifa M., Synthesis of certain nitroquinazoline derivatives structurally related to some chemotherapeutic agents. *Egyptian Journal of pharmaceutical sciences*, **13(1)**, 11-21 (1972).
271. Adrian A., v-Triazolo[4,5-d]pyrimidines (8-azapurines). X. New routes to v-triazolo[4,5-d]pyrimidines via 4-[[[(dimethylamino)methylene] amino]-1,2,3-triazole-5-carbonitriles, *Journal of the Chemical Society*, Perkin Transactions 1: Organic and Bio-Organic Chemistry (1972-1999), **4**; 461-7, (1972).
272. Baker S., Firestine S., Smithrud D., Salinas F., Benkovic S., Synthesis of two bi-functional ligands for the Quest three- hybrid system; *Tetrahedron letters*, **41(36)** 7009-7012 (2000).
273. Koguro K., Oga T., Mitsui S., Orita R., Novel synthesis of 5-substituted tetrazoles from nitriles, *synthesis*, 910-914 (1998).
274. Stephen T., Stephen H., Syntheses in the quinazolone series. II. Synthesis of quino- and quinazoquinazolones, *J. Chem. Soc.*, 4173-7 (1956).
275. Gineinah M., El-sherbeny M., Nasr M., Maarouf A., Synthesis and anti-inflammatory screening of some quinazoline and quinazolyl-4-oxoquinazoline derivatives. *Archiv der Pharmazie*, **335** (11-12), 556-562 (2003).
276. Maloshitskaya O., Sinkkonen J., Aleksyev V., Zelenin K., Pihlaja K. A comparison of ring-chain tautomerism in heterocycles derived from 2-aminobenzenesulfonamide and anthranilamide, *Tetrahedron*, **61**, 7294-7303 (2005).
277. Silvana R., Jose E., Roberto R., Syntheses of Some 4-Anilinoquinazoline Derivatives, *Syntheses*, **3**, 0429-0435 (2004).
278. Connolly D., Lacey P., McCarthy M., Saunders C., Carroll A., Goddard R., Guiry P. Preparation and Resolution of a Modular Class of Axially Chiral Quinazoline-Containing Ligands and Their Application in Asymmetric Rhodium-Catalyzed Olefin Hydroboration, *J. Org. Chem*, **69**, 6572-6589, (2004).
279. Hardy J., Wells J., Searching for new allosteric sites in enzymes, *Cur. Op. Str. Bio.*, **14**, 1-10 (2004).
-

280. Sonia Biondaro and Andrea Brancale "Computer-aided design and synthesis of novel hepatitis C helicase inhibitors" -Degree of Pharmacy (2006). Molecular dynamics simulations were performed by Dr. Andrea Brancale using GROMACS 3.1 (www.gromacs.org), using a triclinic water box and NPT conditions.
281. Kandil S., Biondaro S., Vlachakisa D., Cummins A., Coluccia A., Berry C., Pieter P., Neyts J., Brancale A., Discovery of a novel HCV helicase inhibitor by a de novo drug design approach, *Bioorganic & Medicinal Chemistry Letters*, **19**, 11, (2009).
282. <http://www.enamine.net/index.php?option=content&task=view&id=8>
283. Deng X., Mani N., A facile, environmentally benign sulfonamide synthesis in water, *Green Chem.*, **8**, 835-838 (2006).
284. A) Khan M., Morgan K., Morrey D. Carbonyl derivatives of heterocyclic compounds-III, *Tetrahedron*, **22**, 2095-2105, (1966). B) David M., Sam H. *et al.* Rational tetraarylporphirin syntheses; tetraarylporphirins from the Mac Donald Route, *J. Org. Chem.*, **53**, 7245-7257, (1993).
285. Chimni S., Mahajan D. Electron deficiency of aldehydes controls the pyrrolidine catalyzed direct cross-aldol reaction of aromatic/heterocyclic aldehydes and ketones in water, *Tetrahedron*, **61**, 5019-5025 (2005).
286. Goodyer C., Chinje E., Jaffar M., Stratford I. and Threadgill M., Synthesis of *N*-benzyl- and *N*-phenyl-2-amino-4, 5-dihydrothiazoles and thioureas and evaluation as modulators of the isoforms of nitric oxide synthase, *Bioorganic & medicinal chemistry*, **19**, 4189-4206 (2003).
287. Boger D., Fink E., Hedrick P. Total Synthesis of Distamycin A and 2640 Analogues: A Solution-Phase Combinatorial Approach to the discovery of New, Bioactive DNA Binding Agents and Development of a Rapid, High-Throughput Screen for Determining Relative DNA Binding Sequence Selectivity. *M. J. Am. Chem. Soc.*, **122**, 6382 (2000).
288. Battistuzzi G., Cacchi S., Fabrizi G., An Efficient Palladium-Catalyzed Synthesis of Cinnamaldehydes from Acrolein Diethyl Acetal and Aryl Iodides and Bromides. *Organic Letters*, **5**, 5, 777-780 (2003).
289. Mizoroki T., Mori K., Ozaki A., Arylation of Olefin with Aryl Iodide Catalyzed by Palladium. *Bull. Chem. Soc. Jpn.*, **44**, 581 (1971).
-

-
290. Heck R., Nolley J. Palladium-catalyzed vinylic hydrogen substitution reactions with aryl, benzyl, and styryl halides. *J. Org. Chem.*, **37**(14), 2320-2322 (1972).
291. Beletskaya I., Cheprakov A. The Heck Reaction as a Sharpening Stone of Palladium Catalysis, *Chem. Rev.*, **100**, 3009-3066 (2000).
292. Heck R. Palladium-Catalyzed Vinylation of Organic Halides, *Org. React.*, **27**, 345-390 (1982).
293. Wolfgang K., Prockl K. Heck Reactions of Aryl Chlorides Catalyzed by Ligand Free Palladium Salts, *Catal. Lett.* **125**,197–200 (2008).
294. Svennebring A., Sjöberg P., Larhed M., Nilsson P. A mechanistic study on modern palladium catalyst precursors as new gateways to Pd(0) in cationic Heck reactions, *Tetrahedron*, **64**, 1808-1812 (2008).
295. Tanja Miletic and Andrea Brancale "Rational design and synthesis of HCV helicase inhibitors" -Degree of pharmacy (2008).
296. Alagic A., Koprianiuk A., Kluge R. Hemoglobin-Superoxide Dismutases-Chemical Linkages That Create a Dual-Function Protein, *J. Am. Chem. Soc.*, **127**, 8036-8043 (2005).
297. Herbert M., Konnerth U., Graw S., Echter T. Untersuchungen über das Benzothiiren, *Chemische Berichte*, **117** (1), 107-126 (1984).
298. Hu Q., Tay L., Noestheden M., Pezacki J. Mammalian Cell Surface Imaging with Nitrile-Functionalized Nanoprobes: Biophysical Characterization of Aggregation and Polarization Anisotropy in SERS Imaging. *J. Am. Chem. Soc.*, **129** (1), 14-15 (2007).
299. Cleland W., Dithiothreitol, a New Protective Reagent for SH Groups, *Biochemistry*, **3** (4), 480-482 (1964).
300. Moyer M., Feldman P., Rapoport H., Intramolecular N-H, O-H and S-H insertion reactions. Synthesis of heterocycles from α -Diazo β -Keto Esters, *J. org. Chem.*, **50**, 5223-5230 (1985).
301. Oae, S. Ed.; *Organic Sulfur Chemistry: Structure and Mechanism*, CRC Press: Boca Raton, (1992).
302. Leung-Toung R., Li W., Tam T., Karimian K. Thiol-Dependent Enzymes and Their Inhibitors: A Review. *Current Medicinal Chemistry*, **9**, 979-1002 (2002).
-

-
303. Powers J., Asgian J., Ekici Ö., James K. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases, *Chem. Rev.*, **102**, 4639-4750 (2002).
304. Otto H-H., Schirmeister T. Cysteine Proteases and Their Inhibitors, *Chem. Rev.*, **97** (1), 133-171 (1997).
305. Leung D., Abbenante G., Fairlie D., Protease Inhibitors: Current Status and Future Prospects, *J. Med. Chem.*, **43** (3), 305-341 (2000).
306. Talanian R., Brady K., Cryns V., Caspases as Targets for Anti-Inflammatory and Anti-Apoptotic Drug Discovery *J. Med. Chem.* **43** (18), 3351-3371 (2000).
307. Scozzafava A., Mastrolorenzo A., Supuran C., Agents that target cysteine residues of biomolecules and their therapeutic potential. *Expert Opin. Ther. Pat.*, **11**(5), 765-787 (2001).
308. Grzonka Z., Jankowska E., Kasprzykowski F., Kasprzykowska R., Lankiewicz L., Wiczak W., Wieczerek E., Ciarkowski J., Drabik P., Janowski R., Kozak M., Jaskolski M., Grubb A., Structural studies of cysteine proteases and their inhibitors. Review, *Acta Biochim. Pol.*, **48**(1), 1-20 (2001).
309. Turk B., Turk D., Turk V., Lysosomal cysteine proteases: more than scavengers, *Biochim. Biophys. Acta*, **1477**(1-2), 98-111 (2000).
310. Chen J., Mehta K., Tissue transglutaminase: an enzyme with a split personality *Int. J. Biochem. Cell Biol.*, **31**(8), 817-836 (1999).
311. Cooper A., Jeitner T., Blass J. The role of transglutaminases in neurodegenerative diseases: overview, *Neurochem. Int.*, **40** (1), 1-5 (2002).
312. Sárkány Z., Szeltner Z., Polgár L. Thiolate-Imidazolium Ion Pair Is Not an Obligatory Catalytic Entity of Cysteine Peptidases: The Active Site of Picornain 3C, *Biochemistry*, **40** (35), 10601-10606 (2001).
313. Apotex Inc., US Patent 6,162,791.
314. Donkor I. A survey of Calpain Inhibitors, *Curr. Med. Chem.*, **7**, 1171-1188 (2000).
315. Tam T., Leung-Toung R., Li W., Spino M., Karimian K. Medicinal Chemistry and Properties of 1,2,4-Thiadiazoles, *Mini-Reviews in Medicinal Chemistry*, **5**, 367-379 (2005).
316. Iizawa Y., Okonogi K, Hayashi R., Iwahi T., Yamazaki T., Imada A., Therapeutic effect of cefozopran (SCE-2787), a new parenteral cephalosporin,
-

- against experimental infections in mice. *Antimicrob. Agents Chemother.*, **37**, 100 (1993).
317. Fawzi A., Macdonald D., Benbow L., Smith-Torhan A., Zhang H., Weig B., Ho G., Tulshian D., Linder M., Graziano M. Sch-202676: An allosteric Modulator of both Agonist and Antagonist Binding to G Protein-Coupled Receptors. *Mol. Pharmacol.*, **59**, 30 (2001).
318. Castro A., Castan T., Encinas A., Porcal W., Gil C. Advances in the synthesis and recent therapeutic applications of 1,2,4-thiadiazole heterocycles, *Bioorganic & Medicinal Chemistry*, **14**, 1644–1652 (2006).
319. Martí'nez A., Alonso M., Castro A., Pe'rez C., Moreno F., First Non-ATP Competitive Glycogen Synthase Kinase 3 β (GSK-3 β) inhibitors: Thiadiazolidinones (TDZD) as potential Drugs for the treatment of Alzheimer's Disease. *J. Med. Chem.*, **45**, 1292–1299 (2002).
320. Collins I., Moyes C., Davey W., Rowley M., Bromidge F., Quirk K., Atack J., McKernan R., Thompson S., Wafford K., Dawson G., Pike A., Sohal B., Tsou N., Ball R., Castro J. 3-Heteroaryl-2-pyridones: Benzodiazepine Site Ligands with Functional Selectivity for $\alpha 2/\alpha 3$ -Subtypes of Human GABAA Receptor-Ion Channels *J. Med. Chem.* **45**, 1887-1900 (2002).
321. Unangst P., Shrum G., Connor D., Dyer R., Schrier D., Novel 1,2,4-oxadiazoles and 1,2,4-thiadiazoles as dual 5-lipoxygenase and cyclooxygenase inhibitors, *J. Med. Chem.*, **35**, 3691-3698 (1992).
322. Song Y., Connor D., Sercel A., Sorenson R., Doubleday R., Unangst P., Roth B., Beylin V., Gilbertsen R., Chan K., Schrier D., Guglietta A., Bornemeier D., Dyer R. Synthesis, Structure–Activity Relationships, and in Vivo Evaluations of Substituted Di-*tert*-butylphenols as a Novel Class of Potent, Selective, and Orally Active Cyclooxygenase-2 Inhibitors. 2. 1,3,4- and 1,2,4-Thiadiazole Series. *J. Med. Chem.*, **42**, 1161-1169 (1999).
323. van Muijlwijk-Koezen J., Timmerman H., Vollinga R., Frijtag von Drabbe Kunzel J., de Groote M., Visser S., Ijzerman A. Thiazole and Thiadiazole Analogues as a Novel Class of Adenosine Receptor Antagonists *J. Med. Chem.*, **44**, 749-762 (2001).
324. Antonio Y., Camargo C., Galeazzi E., Iriarte J., Guzman M., Muchowski J., Gerrity K., Liu F., Miller L., Strosberg A. Synthesis of heteroaromatic potential beta-adrenergic antagonists by the glycidol route *J. Med. Chem.*, **21**, 123-126 (1978).
-

-
325. MacLeod A., Baker R., Freedman S., Patel S., Merchant K., Ros M., Saunders J. Synthesis and muscarinic activities of 1,2,4-thiadiazoles. *J. Med. Chem.*, **33**, 2052-2059 (1990).
326. Karimian K., Tam T., Leung-Toung R., Li W., Bryson S., Wodzinska J., Thiadiazole compounds useful as inhibitors of cysteine activity dependent enzymes, *US6, 468,977*, (2002).
327. Karimian K., Tam T., Desilets D., Lee S., Cappelletto T., Li W., Process for scavenging thiols, *US6, 114,537*, (2000).
328. Karimian K., Tam T., Desilets D., Lee S., Cappelletto T., Li W., Proton pump inhibitors, *US6, 093,738*, (2000).
329. Karimian K., Tam T., Leung-Toung R., Li W., Thiadiazole compounds useful as inhibitors of H⁺ /K⁺ Atpase, *US6, 060,472*, (2000).
330. Franz J., Dhingra O. *Comp. Heterocycl. Chem. I*, **6**, 463-511 (1984).
331. Wilkins D., Bradley P., *Comp. Heterocycl. Chem. II*, **4**, 307-354 (1996).
332. Kurzer F. 1,2,4-Thiadiazoles. *Adv. Heterocycl. Chem.*, **32**, 285-398 (1982).
333. Crook S., Sykes P. Reactions of nucleophiles with some *N*-methyl-1,2,4-thiadiazolium salts *J. Chem. Soc. Perkin I*, 1791-1796 (1977).
334. Scozzafava A., Casini A., Supuran C. Targeting Cysteine Residues of Biomolecules: New Approaches for the Design of Antiviral and Anticancer Drugs. *Curr. Med. Chem.*, **9**, 1167-1185 (2002).
335. Goerdeler J., Über 1.2.4-Thiodiazole, I. Mitteil. Darstellung und Eigenschaften der 5-Amino-1.2.4-thiodiazole (Mitbearbeitet von Kurt Wember und Gerhard Worsch) *J. Chem. Ber.*, **87**, 57-67 (1954).
336. Hennrich G., Sonnenchein H., Resch-Genger U., Redox Switchable Fluorescent Probe Selective for Either Hg(II) or Cd(II) and Zn(II). *J. Am. Chem. Soc.* **121**, 5073-5074 (1999).
337. Bryson S., Pai E., Desilets D., Li W., Tam T., Karimian K. Evidence for the formation of a disulfide bond between cysteine proteases and 1,2,4-thiadiazoles; PENCE AGM **2000** Meeting, Montebello, Quebec, Canada, June 24-26; PENCE Inc. National Business Centre, 750 Heritage Medical Research Centre, Edmonton, AB T6G 2S2, Canada.
338. Tam T., Li W., Leung-Toung R., Karimian K. *IDrugs* , **3** (9), 1064-1074 (2000).
-

-
339. Marrano C., de Macedo P., Gagnon P., Lapierre D., Gravel C., Keillor J., Synthesis and Evaluation of Novel Dipeptide-Bound 1,2,4-Thiadiazoles as Irreversible Inhibitors of Guinea Pig Liver Transglutaminase. *Bioorganic & Medicinal Chemistry*, **9**, 3231–3241 (2001).
340. Goerdeler J. *et al*, N-halogenated amines. I. Preparation and properties of N-chloroform and acetamidine. *Chemische Berichte*, **86**, 400-403 (1953).
341. Goerdeler J. *et al*, 1,2,4-Thiadiazoles. I. Preparation and properties of 5-amino-1,2,4-thiadiazoles. *Chemische Berichte*, **87**, 57-67 (1954).
342. Correa-Basurto J., Vázquez Alcántara I., Espinoza-Fonseca L., Trujillo-Ferrara J. *p*-Aminobenzoic acid derivatives as acetylcholinesterase inhibitors, *European Journal of Medicinal Chemistry*, **40** (7), 732-735 (2005).
343. Bartenschlager R., Hepatitis C virus replicons: potential role for drug development, *Nature reviews drug discovery*, **1**, 911-916 (2002).
344. Mc Hutchison 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*, **44**, 411- 421 (2006).
345. Bartenschlager R., Lohmann V. Novel cell culture systems for the hepatitis C virus. *Antiviral Res.* **52**, 1–17 (2001).
346. Kato N., Shimotohno K. Systems to culture hepatitis C virus. *Curr. Top. Microbiol. Immunol.* **242**, 261–278 (2000).
347. Blight K., Kolykhalov A., Rice C. Efficient initiation of HCV RNA replication in cell culture. *Science* **290**, 1972-1974 (2000).
348. Krieger N., Lohmann V., Bartenschlager R. Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J. Virol.* **75**, 4614–4624 (2001).
349. Lohmann V., Körner F., Dobierzewska A., Bartenschlager R. Mutations in hepatitis C virus RNAs conferring cell culture adaptation. *J. Virol.* **75**, 1437–1449 (2001).
350. Guo J., Bichko V., Seeger C. Effect of α -interferon on the hepatitis C virus replicon. *J. Virol.* **75**, 8516–8523 (2001).
351. Blight K., Kolykhalov A., Rice C. Efficient initiation of HCV RNA replication in cell culture. *Science* **290**, 1972-1974 (2000).
-

-
352. Krieger N., Lohmann V., Bartenschlager R. Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J. Virol.* **75**, 4614–4624 (2001).
353. Lohmann V., Körner F., Dobierzewska A., Bartenschlager R. Mutations in hepatitis C virus RNAs conferring cell culture adaptation. *J. Virol.* **75**, 1437–1449 (2001).
354. Guo J., Bichko V., Seeger C. Effect of α -interferon on the hepatitis C virus replicon. *J. Virol.* **75**, 8516–8523 (2001).
355. Ikeda M., Yi M., Li K., Lemon S. Selectable subgenomic and genome length dicistronic RNAs derived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently in cultured Huh7 cells. *J. Virol.* **76**, 2997–3006 (2002).
356. 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.* **82**, 723–733 (2001).
357. Pietschmann T. *et al.* Persistent and transient replication of full-length hepatitis C virus genomes in cell culture. *J. Virol.* **76**, 4008–4021 (2002).
358. Jones S., Antiviral drugs: Breakthrough for HCV research, *Nature Reviews Drug Discovery*, **4**, 629 (2005).
359. Wakita W. *et al.* Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature Med.* **11**, 791–796 (2005).
360. Lindenbach B. *et al.* Complete replication of hepatitis C virus in cell culture. *Science*, **309**, 623–626 (2005).
361. Zhong J. *et al.* Robust hepatitis C virus infection in vitro. *Proc. Natl Acad. Sci.*, **102**, 9294–9299 (2005).
362. Friebe P., Lohmann V., Krieger N., Bartenschlager R. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J. Virol.* **75**, 12047–12057 (2001).
363. 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.* **76**, 5326–5338 (2002).
-

-
364. Cheney I. *et al.* Mutations in NS5B polymerase of hepatitis C virus: impacts on in vitro enzymatic activity and viral RNA replication in the subgenomic replicon cell culture. *Virology* **297**, 298–306 (2002).
365. Frese M. *et al.* Interferon- γ inhibits replication of subgenomic and genomic hepatitis C virus RNAs. *Hepatology* **35**, 694–703 (2002).
366. Tardif K., Mori K., Siddiqui A. Hepatitis C virus subgenomic replicons induce endoplasmic reticulum stress activating an intracellular signaling pathway. *J. Virol.* **76**, 7453–7459 (2002).
367. Pflugheber J. *et al.* Regulation of PKR and IRF-1 during hepatitis C virus RNA replication. *Proc. Natl Acad. Sci.* **99**, 4650–4655 (2002).
368. Bartenschlager R., Pietschmann T., Efficient hepatitis C virus cell culture system: what a difference the host cell makes. *Proc. Natl. Acad. Sci. USA* **102**, 9739–9740 (2005).
369. Yi M., Bodola F., Lemon S. Subgenomic hepatitis C virus (HCV) replicons inducing the expression of a secreted enzymatic reporter protein, *Virology*, **304**, 197–210 (2002).
370. Pietschmann T., Lohmann V., Rutter G., Kurpanek K., Bartenschlager R. Characterization of cell lines carrying self-replicating hepatitis C virus RNAs. *J. Virol.* **75**, 1252–1264 (2001).
371. Young S. Inhibition of HIV-1 integrase by small molecules: the potential for a new class of AIDS chemotherapeutics. *Curr. Opin. Drug Discov. Dev.* **4**, 402–410 (2001).
372. Paeshuysse J., Vliegen I., Coelmont L., Leyssen P., Tabarrini O., Herdewijn P., Mittendorfer H., Easmon J., Cecchetti V., Bartenschlager R., Puerstinger G., Neyts J. Comparative In Vitro Anti-Hepatitis C Virus Activities of a Selected Series of Polymerase, Protease, and Helicase Inhibitors, Antimicrobial Agents and Chemotherapy, **52**, 9, 3433–3437 (2008).
373. Hicham Alaoui-Ismaili M., Gervais C., Brunette S., Gouin G., Hamel M., Rando R., Jean Bedard J., A novel high throughput screening assay for HCV NS3 helicase activity, *Antiviral Research* **46**, 181–193 (2000).
-