Identification and characterisation of small molecule inhibitors targeted to the hepatitis C virus NS2 autoprotease.

Joseph Charles Shaw (BSc (Hons))

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds

Astbury Centre for Structural Molecular Biology

September 2014

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2014 The University of Leeds Joseph Charles Shaw

The right of Joseph Charles Shaw to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgments

The completion of this thesis would not have been possible without a number of people.

Firstly, I would like to thank my supervisors Prof. Mark Harris and Prof. Colin Fishwick for their continued advice, support and enthusiasm with this project.

I would also like to thank past and present members of Garstang 8.61. Thanks in particular to Doug for helpful discussion and to Barney and Bruno for the same as well as occasionally taking my mind off science. Special thanks to Hazel for advice and for her relentless optimism and ability to cheer me up during the more frustrating times.

I am also grateful to Katie, Martin, Sarah and Ricky for providing a great deal of help to a biologist in a chemistry lab.

Thank you to my family for their support throughout the highs and lows of the last four years.

Finally I would like to acknowledge the Wellcome Trust for the funding which made this project possible.

Abstract

Hepatitis C virus (HCV) is a positive-strand RNA virus present in 2-3% of the global population and commonly establishing a chronic infection, leading to long term diseases such as liver cirrhosis and hepatocellular carcinoma. Recent advances have led to the development of a range of direct-acting anti-viral drugs (DAAs), some of which are already improving outcomes in the clinic. It is clear however, that effective therapy for the treatment of HCV will most likely require a combination of DAAs to overcome the rapid onset of viral resistance. In this regard additional inhibitors of the virus lifecycle, which act through a novel molecular target, are required.

The autoprotease activity encoded within the C-terminus of the non-structural 2 (NS2) protein is essential for processing of a precursor to the mature viral proteins, and as a consequence is also required for the onset of viral genome replication and the establishment of HCV infection. Despite representing an attractive target for anti-virals, no inhibitors of the NS2 autoprotease have been reported.

In order to identify small molecule inhibitors of the NS2 autoprotease, two independent assays were optimised as a measure of NS2-mediated proteolysis. These assays were employed to demonstrate that inhibitors of the NS2 autoprotease were able to block HCV genome replication. The assays were subsequently used to identify a lead-like small molecule inhibitor by screening an *in silico* enriched library. This compound was further characterised in the context of NS2 activity *in vitro* and cell culture models of the virus lifecycle. The resultant series represent the first documented inhibitors capable of exerting an anti-viral effect by targeting the NS2 autoprotease.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

µM	micromolar
A	Ala, alanine
Å	Angstrom
ADME	Absorption, distribution, metabolism and elimination
ALT	alanine aminotransferase
ApoB	Apolipoprotein B
ApoE	Apolipoprotein E
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BSA	Bovine serum albumin
BSL3	Biosafety level 3
BVDV	Bovine viral diarrhoea virus
C	Cys, Cysteine
CADD	Computer aided drug design
CAM	Culture adapted mutation
CC_{50}	50% cytotoxic concentration
cDNA	complementary deoxyribonucleic acid
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate
CLDN	Claudin
CO_2	Carbon dioxide
CsA	Cyclosporine A
Cyp	Cyclophilin
CypA	Cyclophilin A
D	Asp, aspartic acid
DAA	Direct acting anti-viral
DENV	Dengue virus
DEPC	Diethylpyrocarbonate
dH_2O	de-ionised water
DMEM	Dulbecco's modified Eagle's medium
DMF	Dimethylformamide
DMSO	Dimethyl sulphoxide
DMV	double membrane vesicle

DNA	deoxyribonucleic acid
dNTP	deoxy-nucleotide triphosphate
dsRNA	double stranded ribonucleic acid
DTT	Dithiothreitol
E	Glu, glutamic acid
E. Coli	*Eschericia Coli*
E1	Envelope protein 1
E2	Envelope protein 2
EC_{50}	50% effective concentration
EDTA	ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EGTA	ethylene glycol tetraacetic acid
EMCV	Encephalomyocarditis virus
ER	Endoplasmic reticulum
ERK	Extracellular signal-regulated kinase
EtOAc	Ethyl acetate
F	Frameshift protein
FBS	Fetal bovine serum
ffLuc	firefly luciferase
fM	femtomolar
G	Gly, glycine
GAG	glycosaminoglycan
GBV-B	George Barker virus – B
GdnHCl	Guanidine hydrochloride
GFP	Green fluorescent protein
GST	Glutathione S-transferase
gt	genotype
H	His, histidine
h.p.e.	hours post electroporation
HAV	Hepatitis A virus
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HIV	Human Immunodeficiency virus

HPLC	High pressure liquid chromatography
Hsp90	Heat shock protein 90
HTLV	Human T-cell leukemia virus
HTS	High-throughput screening
HVR	Hyper-variable region
IAA	Iodoacetic acid
IAM	Iodoacteamide
IFN	Interferon
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRES	Internal ribosome entry site
ISG	Interferon stimulated gene
JEV	Japanese encephalitis virus
JFH1	Japanese Fulminant Hepatitis 1
kDa	kilo Daltons
LAR-I	Luciferase Assay Reagent I
LB	Luria Bertani broth
LC-MS	Liquid chromatography mass spectrometry
LCS	Low complexity sequence
LD	Lipid droplet
LDL	Low density lipoprotein
LDL-R	Low density lipoprotein receptor
luLD	luminal lipid droplet
LVP	lipoviroparticle
MAPK	Mitogen activated protein kinase
MAVS	mitochondrial antiviral signalling protein
mCPBA	meta-chloroperbenzoic acid
miRNA	micro ribonucleic acid
mM	millimolar
MM	Molecular mechanics
MOPS	3-(N-morpholino) propanesulphonic acid
mRNA	messenger ribonucleic acid
MS	Mass spectrometry
MTT	Thiazolyl Blue Tetrazolium Bromide
N	Asn, asparagine

NANBH	Non-A, non-B hepatitis
NEM	N-ethyl maleimide
NI	Nucleoside analogue inhibitor
Nluc	Nanoluciferase
nM	nanomolar
NMR	Nuclear magnetic resonance
NNI	Non-nucleoside analogue inhibitor
NPHV	non-primate hepacivirus
NPT	Neomycin phosphotransferase
NS	Non-structural
NS	Not statistically significant
NS2	Non-structural 2
NS3	Non-structural 3
NS4A	Non-structural 4A
NS4B	Non-structural 4B
NS5A	Non-structural 5A
NS5B	Non-structural 5B
NTP	Nucleotide triphosphate
Nu	Nucleophile
OCLN	Occludin
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG-IFN	pegylated interferon
PFA	Paraformaldehyde
PI4K-α	Phosphatidylinositol 4-kinase III α
PKR	protein kinase R
PLB	Passive Lysis Buffer
pM	picomolar
PMSF	Phenyl-methane-sulphonylfluoride
PPI	Peptidyl prolyl isomerase
ppm	parts per million
PR	HIV-1 protease
PVDF	polyvinylidene fluoride

RBV	Ribavirin
RC	Replication complex
RdRp	RNA dependent RNA polymerase
Rf	Retention factor
RIG-I	Retinoic acid-inducible gene I
RLU	Relative Luciferase Units
RNA	ribonucleic acid
rNTP	ribonucleoside triphosphate
RT-PCR	Reverse transcription polymerase chain reaction
S	Ser, serine
SAR	Structure activity relationship
SDS	Sodium dodecyl sulphate
SGR	Sub-genomic replicon
SL	Stem loop
SMV	Single membrane vesicle
SOC	Standard of care
SR-BI	Scavenger receptor-BI
ssRNA	single stranded ribonucleic acid
SVR	sustained virological response
T	Thr, threonine
TBS	Tris buffered saline
TC-PTP	T-cell protein tyrosine phosphatase
TEA	Triethylamine
TEMED	N,N,N',N'-tetramethylethylenediamine
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TLCK	Tosyl-L lysine chloromethyl ketone
TLR-3	Toll-like receptor 3
TMD	Trans-membrane domain
TMS	Tetramethylsilane
TNF-α	Tumour necrosis factor α
tNS2	truncated non-structural 2
TPCK	Tosyl phenylalanine chloromethyl ketone
TRIF	toll/IL-1 receptor

UTR	Untranslated region
vHTS	virtual high-throughput screening
VLDL	Very low density lipoprotein
WHO	World Health Organisation
Wt	wildtype
YFV	yellow fever virus

Chapter 1 - Introduction

1.1 Viral hepatitis

The term hepatitis refers to the inflammation of the liver and is most commonly attributed to viral hepatitis, wherein the causative agent is one of the hepatitis A-E human viruses. Prior to the identification of hepatitis C virus (HCV) in 1989 (Choo et al., 1989), cases of viral hepatitis were primarily attributed to hepatitis A virus (HAV) or hepatitis B virus (HBV). It was later realised that the majority of cases of viral hepatitis were caused by an alternative pathogen unreactive to HAV and HBV diagnostic tests (Feinstone et al., 1975, Knodell et al., 1975). This pathogen, termed non-A, non-B hepatitis (NANBH) could be transmitted to chimpanzees intravenously and was initially proposed to compose several causative agents (Bradley et al., 1983, Bradley et al., 1985). However, production of a complementary DNA library (cDNA) from the plasma of infected chimpanzees revealed a positive sense, single-stranded RNA molecule with a single ~3000 amino acid open reading frame (ORF) that showed significant homology to the non-structural (NS) proteins of *Flaviviridae* family members (Choo et al., 1989, Choo et al., 1990). Thus development of NANBH was attributed to a single infectious agent termed hepatitis C virus (HCV).

1.2 Hepatitis C virus

1.2.1 HCV classification, epidemiology and genetic heterogeneity

Hepatitis C virus is the prototype member of the *Hepacivirus* genus within the *Flaviviridae* family, which also comprises the genera *Flavivirus*, *Pestivirus* and *Pegivirus* (Figure 1.1A). Despite sequence homology and functional similarities between the non-structural proteins (Murray et al., 2008), HCV contains an altered genome organisation from that of members of the *Flavivirus* genus such as yellow fever virus (YFV), Dengue virus (DENV) and Japanese encephalitis virus (JEV). HCV is more closely related to the *Pegivirus* genus and the *Pestivirus* genus which typically infect mammals such as cattle and swine, as with the *Pestivirus* bovine viral diarrhoea virus (BVDV). Additional animal viruses are more closely related to HCV such as George Barker virus-B (GBV-B) which infects tamarin monkeys and more recently the non-primate hepacivirus (NPHV), an equine virus, along with HCV-like and *pegivirus*-like viruses isolated from rodents (Burbelo et al., 2012, Kapoor et al., 2013).

Within the *Hepacivirus* genus HCV can be divided into six distinct genotypes (and a recently postulated seventh genotype (Murphy et al., 2007)) which typically differ by 30-35% at the nucleotide level across the entire genome. These genotypes are further divided into numerous subtypes (a, b, c etc.), with genome sequence variation of 20-25%. While all the HCV

genotypes share significant similarities in aspects of their genome replication and ability to persistently infect the liver of their host, their distribution worldwide varies, as does their prevalence within the population of certain risk groups (Simmonds, 2004).

The most prevalent HCV genotypes, namely genotype 1b, 1a and to a lesser extent 3a, are present worldwide, but account for the majority of infections in the Americas and Europe. Genotype 2 also has a worldwide distribution, albeit less overall prevalence, but is endemic to regions of West Africa. Central Africa and the Middle East are endemic for genotype 4, while much of East Asia is associated with genotypes 3 and 6 (Simmonds, 2004, WHO, 2009) (Figure 1.1B).

While certain genotypes are geographically constrained, they can also be associated with higher rates of infection within the population. An example of this is Egypt, where the overwhelming majority of infections are genotype 4 and up to 20% of the population are thought to be infected, compared to the global average of 3% (Frank et al., 2000).

Despite high infection rates of some of the genotypes constrained to certain geographical locations, the majority of anti-viral strategies have attempted to target strains prevalent across the globe; primarily genotype 1 but also genotype 3 and to a lesser degree genotype 2. These strains make up a significant proportion of the 170-180 million people infected worldwide.
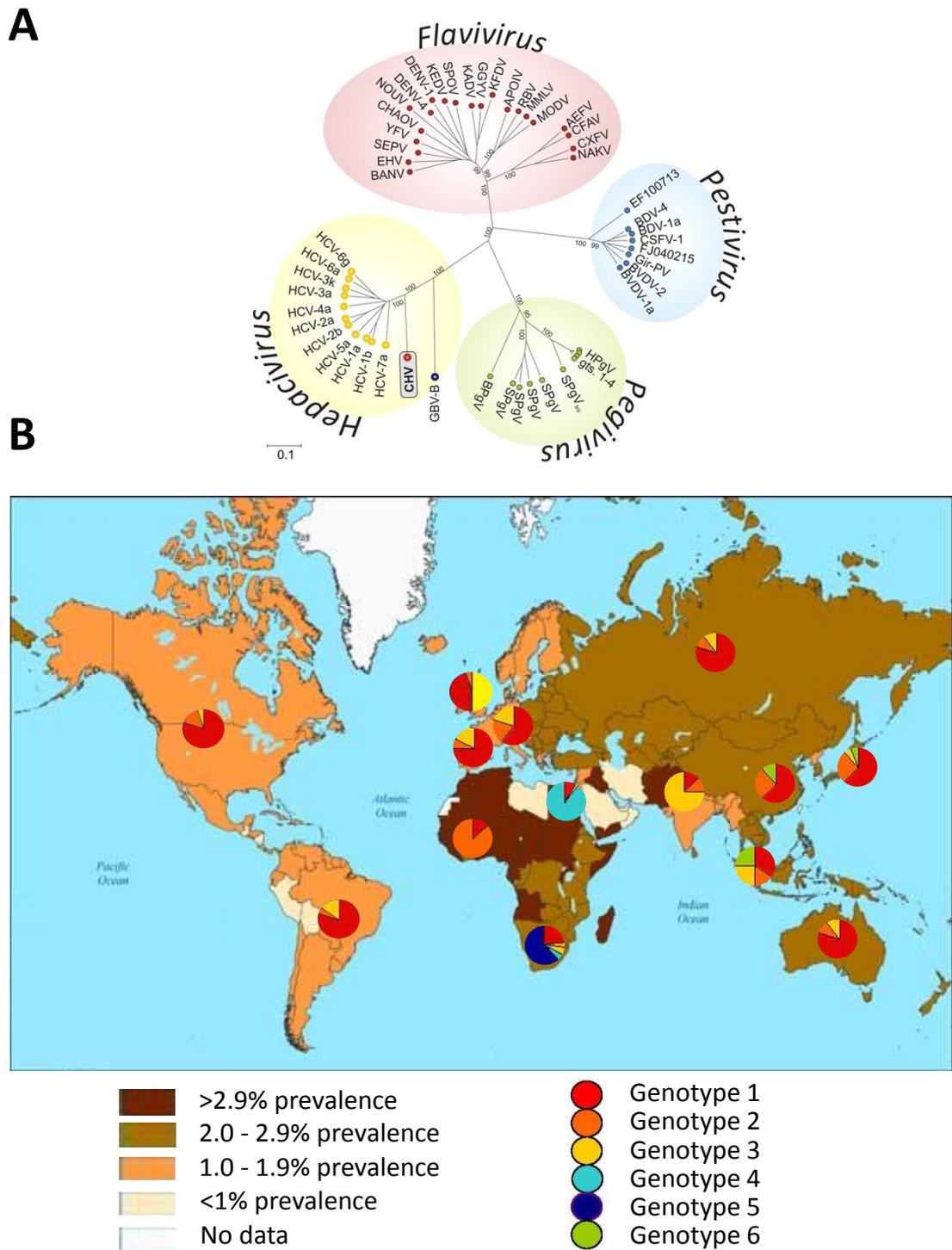
Figure 1.1: HCV epidemiology and distribution.

A) Genera within the Flaviviridae family organised by amino acid sequence identity within the RNA dependent RNA polymerase, adapted from (Kapoor et al., 2011). B) Global prevalence of HCV as percentage of population infected and genotype distribution, adapted from (WHO, 2009).

1.2.3 HCV transmission

HCV is a blood-borne pathogen. Prior to the introduction of donor screening in 1991 the major route of transmission was blood transfusions (Lauer and Walker, 2001). In developed nations blood transfusions are now routinely screened for indicators of liver disease; elevated serum alanine aminotransferase (ALT) and HCV antibodies. As such the majority of new infections are now due to intravenous drug-use, which is estimated to account for up to 80% of infections in some developed nations. The prevalence of infection in such countries has been reported as 64-94% of long-term injectors (> 6 years), while it has also been reported that HCV is more permissive to transmission through this route than other blood-borne viruses (Shepard et al., 2005).

In developing nations, a larger proportion of infections can often be attributed to unsafe healthcare related practices such as therapeutic injections and blood transfusions.

Other routes of transmission occur much more infrequently, with transmission due to occupation, such as with health care workers, and from religious or cultural practices including tattooing, body-piercing, circumcision and acupuncture. Perinatal transmission of HCV has been reported, but appears as low as 2.7-8.4%, while there is also a small risk of sexual transmission (Shepard et al., 2005).

1.2.4 Disease progression

The majority of HCV infections will progress from an acute infection to a chronic infection, whereby virus persistence can be detected after 6 months either through detection of HCV RNA in the blood or increased levels of ALTs.

70-80% of acute infections are asymptomatic and often undiagnosed, but where symptoms are observed they tend to present after 3-12 weeks, shortly after the typical peak in HCV RNA detected in the serum. Symptoms include weakness, jaundice, malaise and anorexia. Following the peak in serum HCV RNA of 10^5-10^7 IU/ml, only 15-25% of patients will clear the virus, wherein HCV RNA becomes undetectable and ALT levels normalise. The majority of cases (75-85%) will progress to chronic infection, though several factors can influence this including ethnicity, gender and age at infection (Chen and Morgan, 2006).

Disease progression from chronic infection involves development of liver fibrosis leading to cirrhosis. End stage liver cirrhosis and associated liver failure is a significant cause of mortality from HCV infection, as is the associated risk of developing hepatocellular carcinoma (HCC)

(Levrero, 2006). The risk of progressing through each disease state varies significantly both with risk factors and with time, but typically 10-20% of chronic infections will lead to cirrhosis over a period of 20 years (Chen and Morgan, 2006). This figure falls closer to 20% after 30 years (Lauer and Walker, 2001). Additional risk factors, in combination with time, include gender, age at time of infection and comorbid conditions, but are most significantly influenced by co-infection with human immunodeficiency virus (HIV) or HBV and by excessive alcohol consumption (Figure 1.2).

As with acute infection, the progression of a chronic infection to cirrhosis is usually asymptomatic, meaning many cases of HCV infection are not detected until the presentation of symptoms at end-stage cirrhosis or HCC. The development of HCC is almost exclusively linked to liver cirrhosis. Once the liver is cirrhotic the risk of developing HCC is 1-4% per year.

The presentation of disease and development of cirrhosis also varies between the HCV genotypes. For example, while genotype 3 is associated with increased levels of spontaneous clearance at the acute phase, it also presents with faster progression of liver fibrosis, accompanied with liver steatosis. By contrast a higher proportion of genotype 4 infections progress to a chronic infection (Nunez and Soriano, 2005).
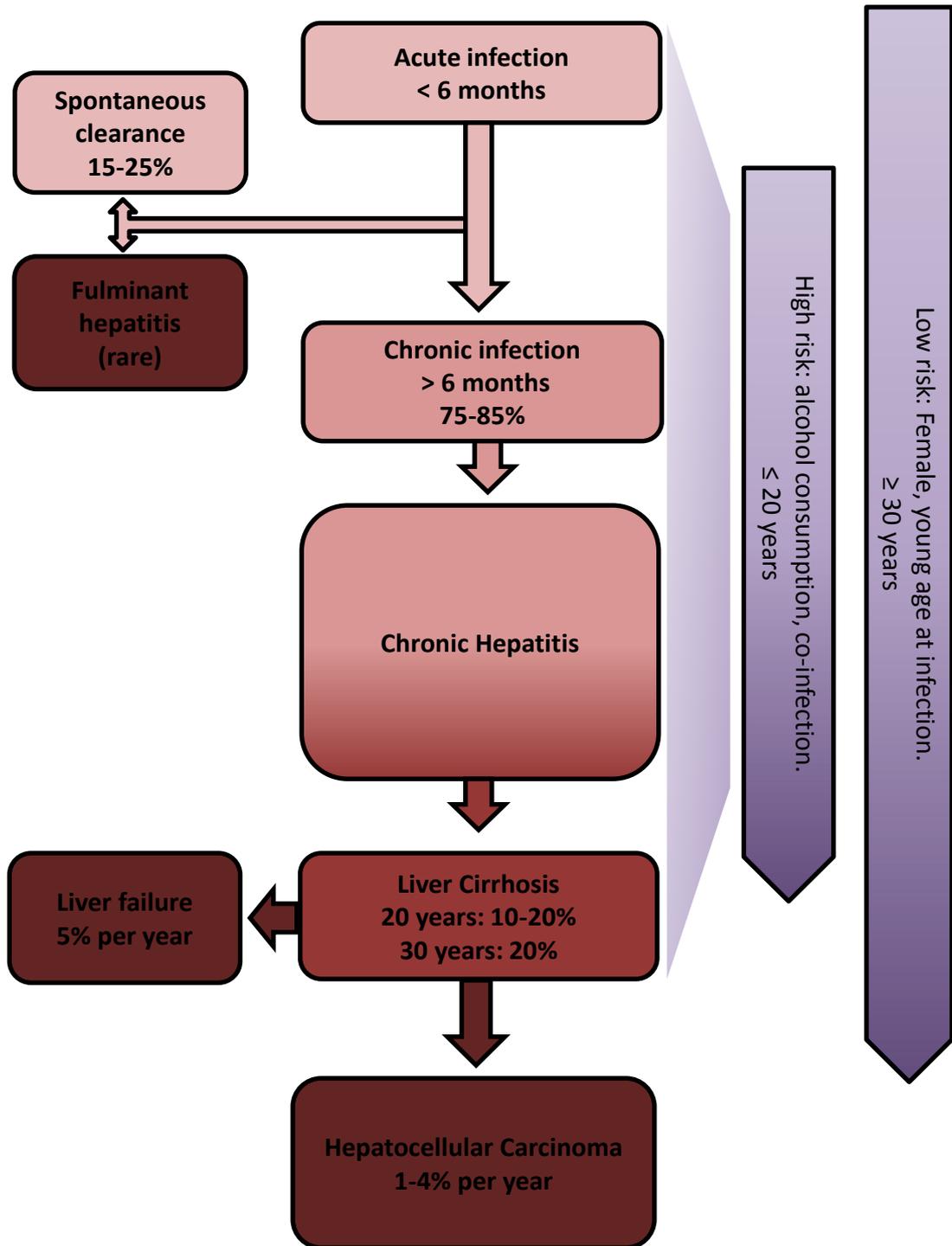
Figure 1.2: HCV disease progression.

Natural progression of HCV infection towards disease states associated with highest mortality (darker boxes). Known environmental risk factors for the development of liver cirrhosis from an acute infection are shown in purple, adapted from (Chen and Morgan, 2006) and (Lauer and Walker, 2001).

1.3 Treatment of HCV

Prior to recent advances in therapies, the standard of care (SOC) treatment for HCV infection was pegylated interferon-alpha (PEG-IFN) and ribavirin (RBV). The drive towards HCV anti-virals has been principally due to the extensive problems with this therapy.

The use of interferon (IFN) to treat viral hepatitis pre-dates the identification of HCV (Hoofnagle et al., 1986), with IFN monotherapy giving a sustained virological response (SVR), defined as the long-lasting absence of HCV RNA from the serum, in 5-20% of cases. SVR rates were improved to 40-50% with the addition of RBV, while the enhanced pharmacokinetics achieved by conjugating IFN to polyethylene glycol (PEG-IFN) further improved this to 54-63%.

However, the success rate of achieving SVR varies greatly between HCV genotypes. While genotypes 2 and 3 respond well to PEG-IFN-RBV therapy, with SVR achieved in 75% of cases, the more prevalent genotype 1 infections are only successful in 40-50% of cases. Likelihood of achieving SVR often depends on treatment duration. Typical PEG-IFN-RBV treatment lasts 24 weeks. This can be reduced to as low as 12 weeks for infections of genotypes 2 and 3, but often has to be extended for genotype 1 infections to 48 weeks or longer. When treatment is extended beyond 72 weeks patients frequently discontinue therapy before SVR is achieved, due mainly to the numerous serious side effects associated with IFN and RBV therapy which include haemolytic anaemia , fatigue, insomnia, depression, dermatitis, headaches, nausea, weight loss, diarrhoea, fever and muscle pain amongst others (Manns et al., 2006).

The success of a PEG-IFN-RBV therapy can be gauged at 12-24 weeks, at which point if a decline in serum HCV RNA of greater than 2 $\log_{10}$ is not observed SVR will not be achieved and therapy is discontinued. This yields a large number of patients classed as 'non-responders' to IFN therapy.

Patients who present with acute HCV infection can be treated with PEG-IFN alone, preventing the development of a chronic infection in 90% of cases. However, due to the aforementioned side effects of IFN treatment, treatment is often withheld until approximately 12 weeks to see if the patient spontaneously resolves the infection. While IFN therapy is fairly successful at clearing acute HCV infection, the asymptomatic nature of the majority of these cases means that they are usually not diagnosed. In fact many HCV infections are not diagnosed until late-stage fibrosis or cirrhosis has developed, a relevant factor as the chances of achieving SVR with PEG-IFN-RBV therapy are further reduced in cases of liver cirrhosis or HCC (Lauer and Walker, 2001).

Treatment of HCV was improved in 2012 with the approval of the first direct acting anti-virals (DAAs); Boceprevir (Merck) and Telaprevir (Janssen-Vertex). Both of these drugs target the virally encoded NS3/4A protease to directly block the virus lifecycle (Salam and Akimitsu, 2013). To date these anti-virals are only licensed for genotype 1 infections in combination with PEG-IFN and RBV, with clinical trials suggesting these therapies can improve SVR rates to 55-75% for Boceprevir triple therapy and 61-67% for Telaprevir triple therapy (Kanda et al., 2010).

More recently second generation protease inhibitors have progressed through clinical trials, such as Simeprevir (Janssen Pharmaceuticals) and Faldaprevir (Boehringer-Ingelheim), as has a promising nucleoside analogue targeting the viral polymerase, Sofosbuvir (Gilead Sciences). Clinical trial data for these inhibitors are summarised below.

The majority of trials are not yet of sufficient size to draw conclusions between genotype 1b and genotype 1a infections, but against treatment naive genotype 1 patients a 48 week course of the second generation protease inhibitor Faldeprevir in combination with PEG-IFN and RBV gave SVR in 73-92% of cases. It was noted that the chances of viral relapse after treatment were significantly higher if treatment duration was shorter.

Simeprevir, also a second generation protease inhibitor, has also been trialled in combination with PEG-IFN and RBV for 12 weeks followed by an additional 24 weeks treatment with PEG-IFN-RBV. Again SVR rates were improved in treatment naive genotype 1 populations to 80% SVR, but only 58% SVR was achieved in cases of cirrhosis (Barreiro et al., 2013, Miller et al., 2014).

The nucleoside analogue polymerase inhibitor Sofosbuvir has been trialled alongside PEG-IFN and RBV for 12 weeks followed by a 12 week course of RBV against a primarily treatment naive genotype 1 population, achieving 89% SVR. This was again reduced, to 80% SVR, when only cirrhotic patients were treated.

Patients who are IFN non-responders have proved even more difficult to treat. Faldeprevir, PEG-IFN and RBV only achieved 21-35% SVR and even an extended 48 week treatment of Simeprevir and RBV only produced 59% SVR. However, as evidence that combinations of inhibitors may enhance treatment, an IFN-free therapy of Simeprevir, Sofosbuvir and RBV for up to 24 weeks gave 93% SVR against an IFN non-responder cohort.

Sofosbuvir has also been trialled against genotype 3 infections. In treatment naive patients the success of a course of Sofosbuvir and RBV depended on treatment duration (12 weeks; 61% SVR, 24 weeks; 94% SVR). Achievement of SVR was reduced in cases of cirrhosis (12 weeks; 34% SVR) and IFN non-responders (12 weeks; 30%).

Genotype 2 infections have traditionally been associated with higher SVR rates than infections of genotypes 1 and 3 using traditional PEG-IFN-RBV therapy. A trial of Sofosbuvir and RBV for 12 weeks improved SVR rates to 97% (PEG-IFN-RBV control arm 78% SVR), but was again dramatically reduced, to 48% SVR, in cases of cirrhosis. While 12 week therapy gave 86% SVR against IFN non-responders, this could be improved to 94% SVR if treatment was extended to 16 weeks (Barreiro et al., 2013, Miller et al., 2014).

While some clinical trial data is available for infections with other HCV genotypes the cohorts tend to be smaller and the true degree of success of these inhibitors against genotypes 4-6 remains to be seen.

Despite the use of two inhibitors with different viral targets, protease and polymerase, the above clinical trial data presents similar results. These inhibitors as part of PEG-IFN-RBV based triple therapies achieve reasonably high SVR rates against treatment naïve patients, but are associated with less success in cases of cirrhosis and IFN non-responders. While the use of Sofosbuvir and Simeprevir together demonstrate that combining drug candidates may help improve SVR rates, all of these trials still involve the use of at least one of either PEG-IFN or RBV. Key future aims, in addition to improving SVR from therapies, are to negate the need for either PEG-IFN or RBV due to the adverse associated side effects, and to reduce the duration of treatment, as even in the absence of PEG-IFN and RBV several of the new drug molecules present with side effects.

With regards to removing PEG-IFN and RBV, the current clinical data on combination therapies and the success of similar therapies for the treatment of other viruses, such as HIV (Pavlos and Phillips, 2012), suggest the answer to this is additional inhibitors of the HCV lifecycle. In this regard one other class of HCV DAA is being progressed, which target the NS5A protein through an as yet ill-defined mechanism of action.

The most advanced of the NS5A inhibitors is Daclatasvir (Bristol Myers Squibb). A phase II trial of Daclatasvir, Sofosbuvir and RBV gave 100% SVR for genotype 1 infections and 91% SVR for genotypes 2 and 3. Similar results are reported for trials of the NS5A inhibitor Ledipasvir (Gilead) giving 100% SVR alongside Sofosbuvir and RBV. Yet a trial of Daclatasvir and

Asunaprevir (Bristol Myers Squibb), a second generation protease inhibitor, in genotype 1b IFN non-responders achieved SVR in only 64-91% of cases (Kohler et al., 2014), demonstrating that even as a double therapy these inhibitors are not always sufficient to clear the virus in infections which are more difficult to treat.

The busy pipeline of HCV anti-virals offers great promise for a robust treatment with a high success rate of clearing the virus in an IFN and RBV free therapy. Many classes of DAA now achieve pan-genotypic activity but the ability to treat HCV infection is still greatly diminished in IFN non-responders and, significantly, in the large patient populations where infection is not diagnosed before the development of liver cirrhosis. The current evidence supports the use of combinations of DAAs to improve the chances of clearing the virus and reduce the risk of viral rebound in an IFN and RBV free setting. Additional DAAs would greatly improve the treatment options and the likelihood of successful combination therapy. While the list of HCV DAAs progressing through clinical trials is extensive, all of these drugs target one of the three current anti-viral targets; protease, polymerase or NS5A. Supplementation of combination therapies would benefit most from additional anti-viral agents with a novel molecular target or mechanism of action. The only viral protein possessing catalytic activity which is not currently being utilised for anti-viral development is NS2, representing an entirely novel field of therapeutic agents.

Table 1.1 summarises the current drug development pipeline for HCV. The precise mechanism of action of each class of inhibitor is described in more detail in Section 1.5.

Table 1.1: Key HCV anti-virals and current progress.

Drug (Company)	Target	Current stage of development
Telaprevir (Janssen - Vertex)	1st generation NS3/4A protease	Approved
Boceprevir (Merck)	1st generation NS3/4A protease	Approved
Simeprevir (Janssen)	2nd generation NS3/4A protease	Phase III
Faldeprevir (Boehringer-Ingelheim)	2nd generation NS3/4A protease	Phase III
Asunaprevir (Bristol Myers Squibb)	2nd generation NS3/4A protease	Phase III
Danoprevir (Genentech)	2nd generation NS3/4A protease	Phase II
Vaniprevir (Merck)	2nd generation NS3/4A protease	Phase II
Narloprevir (Merck)	2nd generation NS3/4A protease	Phase II
Sofosbuvir (Gilead)	NS5B polymerase	Phase III
Mericitabine (Genetech)	NS5B polymerase	Phase III
Daclatasvir (Bristol Myers Squibb)	NS5A	Phase III
Ledipasvir (Gilead)	NS5A	Phase III
ABT-267 (AbbVie)	NS5A	Phase III
Alisporivir (Novartis)	Cyclophilin A PPIase	Phase III (on hold)

1.4 Molecular virology and lifecycle of HCV

1.4.1 Genome organisation

The HCV genome comprises a positive sense single stranded RNA molecule approximately 9,600 nucleotides in length coding for a single open reading frame (ORF). The coding region is flanked by both a 5' and a 3' un-translated region (UTR) (Figure 1.3). The 5' UTR is 341 nucleotides in length and forms stable RNA secondary structures in the form of four stem loops, termed stem-loop (SL) I-IV. SLII, III and IV and a small section at the 5' of the core ORF form an internal ribosome entry site (IRES) which binds directly to the 40s subunit of the ribosome to allow the initiation of translation in the absence of a 5' cap (Hashem et al., 2013). The 5' UTR also contains two tandem sites which bind to micro-RNA (miRNA) 122. The binding of miRNA-122 regulates replication and translation and helps to explain the tissue-specific tropism of HCV, as miRNA-122 is only present in liver hepatocytes (Shimakami et al., 2012).
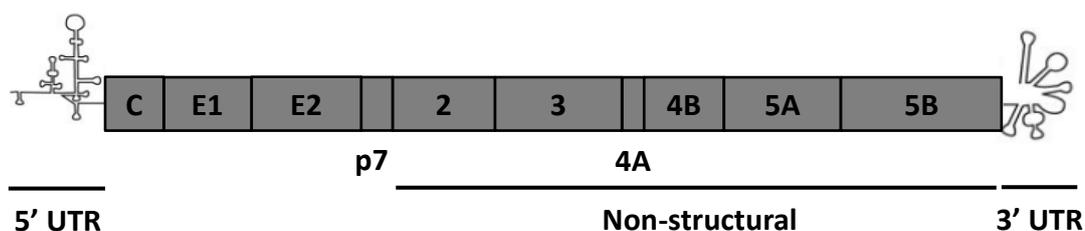
Figure 1.3: Genome organisation of HCV

The 3' UTR comprises a short variable region, a poly U/UC tract and a 3' X-tail. The variable region displays low sequence conservation between genotypes but has high conservation between strains of the same genotype. Similarly, the length of the poly U/UC stretch varies between strains, unlike the 3' X-tail which contains three highly conserved SLs important for the synthesis of negative sense RNA as a replication intermediate.

The viral genome also encodes several *cis*-acting regulatory elements again important for regulation of replication and translation. SLs within the coding region of NS5B interact via both kissing loop and pseudoknot interactions to SLs within the 3' UTR X-tail. It is likely the RNA genome contains multiple other secondary structure elements regulating aspects of the lifecycle. For example additional SLs have been identified in the core coding region, which play an as yet ill-defined role in replication (Vassilaki et al., 2008).

1.4.2 Virus particle composition

Analysis of the HCV virion has been hampered by its heterogeneous nature. HCV particles from cell culture systems are typically 50-80 nm in diameter but have been observed between 40-100 nm and more irregular in structure than other members of the *Flaviviridae* family (Catanese et al., 2013). The infectious HCV RNA circulating in the blood exists as a nucleocapsid particle encased in a host-derived membrane containing the viral envelope glycoproteins E1 and E2. This virion is known to be associated with host lipoprotein-containing particles, most notably apolipoprotein B or E (ApoB / ApoE) in low density lipoprotein (LDL) or very low density lipoproteins (VLDL) particles. These lipid particles are naturally secreted from hepatocytes. As such infectious RNA particles can be referred to as lipoviroparticles (LVPs) (Nielsen et al., 2006, Vieyres and Pietschmann, 2013). The association of the virion with (V)LDL particles and extensive glycosylation of the viral envelope proteins most likely play important roles in preventing immune detection (Bartenschlager et al., 2011).

1.4.3 Viral entry

HCV entry is a complex multistage process. Current models suggest initial interaction with the hepatocyte is mediated by attachment molecules via the (V)LDL associated with the virion, such as heparin sulphate glycosaminoglycans (GAGs) and the low density lipoprotein receptor (LDL-R) (Nielsen et al., 2006, Zhu et al., 2014). Following this a specific binding event occurs between the virion glycoproteins, most notably E2, with numerous entry receptors. CD81 is the most fully characterised of these receptor interactions, with a large 89 residue extracellular loop binding E2 (Pileri et al., 1998, Tang and Grise, 2009) but considerable evidence implicates other receptors such as scavenger receptor-BI (SR-BI) (Scarselli et al., 2002). Additional receptors identified as requirements for HCV entry include Claudin-1 (CLDN-1) (Evans et al., 2007), CLDN-6, CLDN-9 (Meertens et al., 2008) and occludin (OCLN) (Ploss et al., 2009). These receptors are not accessible to sinusoidal blood as they are located at tight junctions, leading to the proposal of a two-stage mechanism whereby an initial specific interaction occurs between the virion and CD81/ SR-BI which mediate a translocation to the tight junction receptors to produce a co-receptor complex at the late stage of the entry process (Harris et al., 2008). Internalisation occurs by clathrin-mediated endocytosis (Blanchard et al., 2006), most likely still in complex with a co-receptor complex as CD81 and CLDN-1 co-endocytosis has been documented (Farquhar et al., 2012). The process of internalisation requires the co-receptor complex of CD81 and CLDN-1, whose interactions can

be modulated through epidermal growth factor receptor (EGFR) dependent signalling pathways. As such the receptor tyrosine kinase EGFR plays a role as an additional entry factor (Zhu et al., 2014).

Following clathrin-mediated endocytosis, fusion of the viral and host membranes allows release of the viral nucleocapsid into the cytosol. While still a poorly understood process this is thought to require a conformational rearrangement of the viral envelope proteins in a pH dependent manner (Tscherne et al., 2006).

An alternative route of infection of hepatocytes is from direct cell-to-cell transmission. Such events have been observed in cell culture models of infectious HCV and can occur in the presence of antibodies capable of neutralising cell-free HCV virions, providing a clear advantage for immune evasion. Cell-to-cell transmission appears to require an overlapping set of cellular receptors. While both CD81-dependent and CD81-independent routes have been observed, both routes required CLDN1 (Timpe et al., 2008).

A

ApoB, ApoE

Host membrane

Capsid (core)

RNA genome

Envelope glycoproteins (E1, E2)

50-80 nm

B
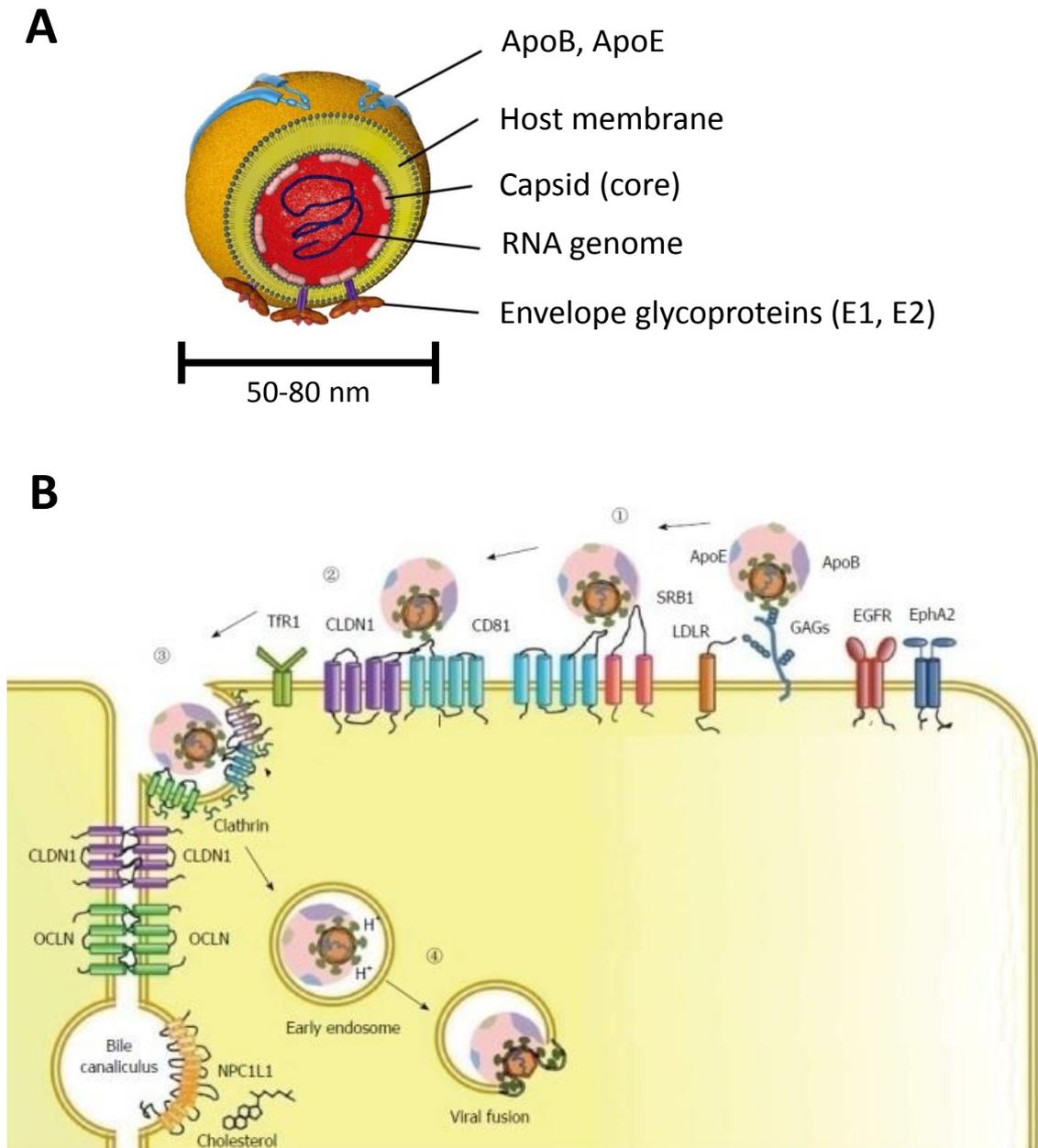
Figure 1.4: HCV entry.

A) Schematic of the HCV virion with key components labelled. Adapted from (Feneant et al., 2014). B) Summary of current model of HCV entry. (1) Attachment molecules GAGs and LDLR facilitate initial association leading to specific interaction with entry receptors CD81 and SRB-1. (2) Co-receptor complex forms with tight-junction associated receptors CLDN1 and OCLN. (3) Co-receptor and virion internalised by clathrin mediated endocytosis. (4) Acidification of early endosomes instigates membrane fusion and genomic release. Adapted from (Zhu et al., 2014).

1.4.4 Polyprotein processing

The released cytosolic genome is of a positive sense and thus acts directly as a mRNA. It is translated in a cap-independent manner through recruitment of the 40s ribosome subunit by the IRES and a kinetically slow conversion to an active 80s ribosome complex. Translation occurs at the rough endoplasmic reticulum (ER). Translation of the primary ORF produces a large polyprotein of approximately 3000 amino acids which is processed co- and post-translationally to form the mature viral proteins.

An additional ORF has been noted due to a +1/-2 ribosomal shift in the core coding region, producing a protein of roughly 160 residues termed Frameshift protein (F) or alternative reading frame protein (Xu et al., 2001). F protein is not required for the virus lifecycle in cell culture and is not conserved between genotypes, indicating a lack of function (Cristina et al., 2005). While antibodies reactive to F have been isolated from HCV-infected patients any potential role for F remains unclear (Morice et al., 2009, Xu et al., 2001).

The HCV ORF encodes 10 mature proteins; core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B, which can be divided into two sections; the structural proteins involved in virion assembly and release and the non-structural proteins which primarily instigate genome replication. These regions can be alternatively characterised into three sections with different functions. The N-terminal third of the polyprotein (core-E2) encodes structural proteins which directly form the infectious particle alongside host membranes, lipoprotein particles and RNA genome. These structural proteins are followed by p7 and non-structural (NS) 2, which facilitate the processes of virus assembly and release. The C-terminal portion of the polyprotein encoding NS3-5B is sufficient for genome replication in cell culture (Lohmann et al., 1999).

During translation the polyprotein is targeted to the ER by a signal sequence N-terminal of E1. Signal sequences are employed by numerous host proteins to instigate secretion by inserting proteins into the ER membrane, with the signal sequence subsequently removed by signal peptidase or signal peptide peptidase. These host cellular proteases perform the cleavage events required for the maturation of core, E1, E2, p7 and the mature N-terminus of NS2. Following insertion of the polyprotein into the ER lumen mediated by a signal peptide at the C-terminus of core, host signal peptidase cleaves at the C-terminus of the signal peptide to liberate the authentic N-terminus of E1. Further signal peptidase activity processes the junctions at E1-E2, E2-p7 and p7-NS2. Core remains ER bound by its C-terminus (which acts as the signal peptide for E1) until an additional cleavage event by the signal peptide peptidase,

either at the N-terminus of or within the signal sequence, releases the mature core protein (Figure 1.5A). After separation from the membrane anchor core associates with lipid droplets (LDs) (McLauchlan et al., 2002).

The signal peptidase and signal peptide peptidase cleavage events occur co-translationally or immediately post translation to leave mature core, membrane embedded E1 and E2 and transmembrane p7. The final membrane associated portion of polyprotein contains NS2-5B. The six non-structural proteins within this fragment are processed by virally encoded proteases. Such a mechanism may allow a greater degree of control over production of the NS proteins that perform genome replication. Some of these proteins are postulated to dimerise or oligomerise for functional reasons and as such it has been proposed that the use of virally encoded proteases allows the regulation of the onset of replication until sufficient levels of viral proteins have accumulated, as with other RNA viruses which employ self-encoded proteases to regulate different stages of their lifecycle (Hellen et al., 1989).

The first of the virally encoded proteases is an autoprotease activity encoded by the C-terminus of NS2 but also requiring the NS3 N-terminus. Therefore the initial step of polyprotein processing and continuation of the virus lifecycle is wholly dependent on the function of the NS2 autoprotease, which cleaves at the NS2-NS3 junction to produce mature NS2 and the native N-terminus of NS3. The NS2 autoprotease is described in more detail in Section 1.8. The NS3 N-terminus encodes a second protease activity which initially cleaves itself at the NS3-NS4A boundary. NS4A acts as a co-factor for the NS3 protease to allow proteolysis at the NS4A-NS4B, NS4B-NS5A and NS5A-NS5B junctions (Figure 1.5A) to produce the last of the 10 mature viral proteins which can begin the process of genome replication, assembly of new virus particles and ultimately release of infectious virions.

1.4.5 Genome replication

The process of genome replication is performed by the NS5B RNA-dependent RNA-polymerase (RdRp), which synthesises positive sense RNA from a negative sense intermediate, but replication also requires NS3, NS4A, NS4B and NS5A. These proteins are tightly associated to each other and to ER membranes and concentrated in a remodelled lipid bilayer termed 'membranous web' (Figure 1.5B). The remodelling of host membranes is a common characteristic of positive sense RNA viruses (Miller and Krijnse-Locker, 2008). In the case of HCV this remodelling is caused by NS4B, though NS5A has also been implicated via the host

factor phosphatidylinositol 4-kinase III α (PI4K-α) (Berger and Randall, 2009). ER-derived membrane is principally remodelled into single membrane vesicles (SMVs) or double membrane vesicles (DMVs) which house the viral non-structural proteins NS3-5B alongside positive and negative sense copies of the RNA genome. These membrane-encased concentrations of viral non-structural proteins are referred to as a 'replication complex' (RC) and are additionally thought to include numerous host factors (Berger and Randall, 2009). The formation of membranous web creates a favourable environment for the process of genome replication by allowing free exchange of nucleotide triphosphates (NTPs) for incorporation into newly synthesised RNA molecules, and release of freshly synthesised RNA into the cytosol, as well as facilitating concentration of the necessary components for replication. DMVs also provide several degrees of protection. Concentrating viral genomes within these DMVs is thought to aid protection from the host innate immune response, particularly RNA sensing mechanisms, while DMVs also provide a physical barrier to host proteases and nucleases (Gu and Rice, 2013).
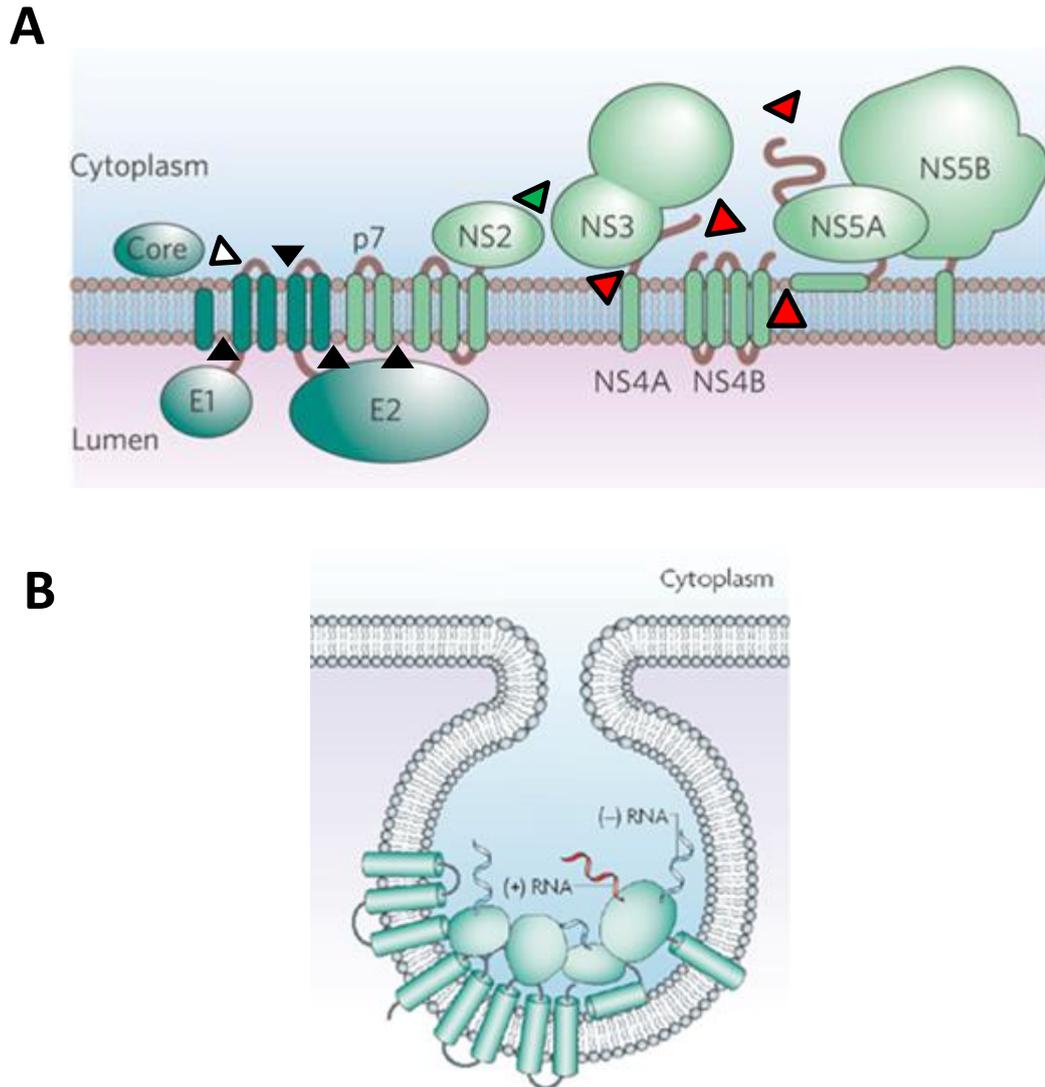
Figure 1.5: HCV polyprotein processing and replication.

A) Processing of the 10 HCV encoded proteins from the polyprotein. The mature N-termini of E1, E2, p7 and NS2 are produced by signal peptidase (black arrowheads). Mature core is produced following cleavage from the C-terminal signal sequence by signal peptide peptidase (white arrowhead). Processing at the NS2-NS3 boundary is performed by the NS2 autoprotease (green arrowhead). Proteolysis of the remaining non-structural proteins NS4A, NS4B, NS5A and NS5B is performed by the NS3 protease (red arrowheads). Figure adapted from (Lindenbach and Rice, 2005). B) Non-structural proteins form replication complexes in remodelled membranous structures termed 'membranous web'. Figure adapted from (Miller and Krijnse-Locker, 2008).

1.4.6 Virus assembly and release

The process of assembly of progeny virus particles can be broadly divided into two events; the formation of an RNA-core nucleocapsid in an early stage and a late stage whereby this nucleocapsid is enveloped.

The process by which RNA associates with mature, LD localised core in order to form a nucleocapsid is poorly understood. One current model hypothesises that viral RNA is delivered to the LDs allowing tight regulation between replication and assembly (Bartenschlager et al., 2011). This function is believed to be performed by NS5A which is known to bind viral RNA (Foster et al., 2010a, Hwang et al., 2010) and has roles in both replication and release (Appel et al., 2008). Additionally NS5A co-localises with core at LDs in cell culture systems for virus propagation (Miyanari et al., 2007).

An alternative model involves assembly of nucleocapsids at the ER, with core delivered from LDs to the ER membrane to interact with RNA-associated NS5A. This model may be more coherent with current understanding of the late stages of viral assembly and egress, which is tightly linked to the host VLDL pathway and involves the transfer of nucleocapsid from cytosolic LDs to luminal LDs (luLDs) (Bartenschlager et al., 2011). luLDs are produced at lipid-rich microdomains within the ER and through fusion with apoE-pre-VLDL particles eventually form mature VLDLs. Thus nucleocapsid-containing luLDs will mature through this pathway to form LVPs before egress through the Golgi secretory pathway.

Little is understood about the maturation of nucleocapsid loaded luLDs, particularly how the envelope glycoproteins E1 and E2 become associated with maturing LVPs. One potential mechanism can be drawn from the observation of NS2-mediated interactions with NS3, NS5A, p7 and the envelope glycoproteins at LDs (Jirasko et al., 2010, Popescu et al., 2011, Stapleford and Lindenbach, 2011). Hence NS2 may act as a link between glycoproteins and non-structural proteins involved in assembly.

Release of infectious virus is acid sensitive and therefore requires oligomerisation of p7 to form a proton conducting ion channel which allows maturation of the virion without exposure to acidic conditions in intracellular compartments (Wozniak et al., 2010).

1.5 HCV encoded proteins

1.5.1 Core

While core is often described as a multifunctional protein, its principle role is as a key component of the nucleocapsid to protect the virus genome. Mature core is 21 kDa and can be divided into two domains. Domain 1, comprising the N-terminal ~117 residues, is hydrophilic and plays important roles in RNA binding and oligomerisation. The first 82 amino acids of this domain, an intrinsically disordered region, are sufficient to form nucleocapsid like particles in the presence of structured RNA *in vitro* (Fromentin et al., 2007). Domain 1 is thought to function as a dimer. Domain 2 by contrast is hydrophobic and responsible for core association with LDs or the ER. This domain comprises two amphipathic α-helices separated by a hydrophobic loop, with this helix-turn-helix conformation known to be required for LD localisation (Gawlik and Gallay, 2014).

As a key component for the process of virus assembly, core interacts not only with genomic RNA but directly or indirectly with p7 and E1, NS5A, NS2 and NS3. In addition core has been shown to modulate a range of host processes; cell signalling, apoptosis and gene transcription (reviewed in (McLauchlan, 2000)).

Core possesses the highest pan-genotype sequence conservation of all the HCV encoded proteins and as such has been postulated as a potential target for anti-virals. Small molecules which block core dimerisation appear to exert an anti-viral effect but remain in early pre-clinical stages (Ni et al., 2011).

1.5.2 E1 and E2

E1 and E2 are the viral envelope glycoproteins which form an E1-E2 heterodimer. Both are type I trans-membrane glycoproteins which are present at the ER, with a short C-terminal trans-membrane domain (TMD) and an N-terminal ectodomain. E1 and E2 have complicated folding patterns requiring numerous host chaperones and are post-translationally modified by N-linked glycosylation. The number of glycosylation sites varies between genotypes but E1 can contain up to 5 glycans and E2 can have up to 11 sites of glycosylation (Vieyres et al., 2014). Glycosylation is important not only for correct folding of E1 and E2 but for the formation of E1-E2 heterodimers, which are incorporated in the host membranes which envelope the nucleocapsid during assembly and egress (see Section 1.4.6). In line with this E2 co-localises with core at LDs (Counihan et al., 2011). Further to this the extensive glycosylation of the

envelope proteins, particularly E2, mask antigenic sites within the LVP to aid immune evasion (Helle et al., 2007).

E2 contains three hypervariable regions (HVRs) which form the targets for HCV neutralising antibodies within the LVP. As such these regions are under a constant selective pressure and are very tolerant to modifications, an aspect further aiding the ability of HCV to evade the innate immune system and develop a persistent chronic infection (Vieyres et al., 2014). These HVRs, in particular HVR-1, mask important envelope epitopes such as those in E2 involved in CD81 binding (Bankwitz et al., 2010) and therefore regulate the binding event. Within the E1-E2 heterodimer the E2 ectodomain interacts with CD81 (Pileri et al., 1998), SR-B1 (Scarselli et al., 2002) and CLDN-1 (Evans et al., 2007) while E1 binds LDLR (Mazumdar et al., 2011) to facilitate entry.

The roles of E1-E2 in virus uncoating upon entry are ill-defined. Comparison to other *Flaviviridae* family members would predict one of these glycoproteins to acts as a class II fusion protein to mediate uncoating. Differing models suggest that either E1 or E2 may fulfil this function as a truncated class II fusion protein (Garry and Dash, 2003, Krey et al., 2010), though more recent evidence based on BVDV E2 homologs indicate a novel form of fusion protein (El Omari et al., 2013, Li et al., 2013). The fusion protein function of one or both of E1-E2 appears to be modulated by a pH-dependent conformational change requiring interaction with CD81 (Sharma et al., 2011).

1.5.3 p7

p7 is a 63 amino acid hydrophobic protein which oligomerises to form hexameric or heptameric cation-selective ion channels (Griffin et al., 2003, Premkumar et al., 2004). Each monomer comprises two TMDs linked by a conserved cytosolic loop. The formation of oligomeric p7 ion channels allows proton conductance across the membranes of organelles such as secretory vesicles in the egress and release pathway which would otherwise be acidified. Thus p7 protects the maturing virion from exposure to acidic conditions (Wozniak et al., 2010).

p7 plays additional roles in the early stages of virus assembly. Deletion of p7 from virus in cell culture blocks an early stage of assembly which occurs prior to formation of intracellular particles (Jones et al., 2007). The signal peptidase mediated cleavage of E2-p7 and p7-NS2 boundaries has been noted as inefficient, with suggestions that the ability to regulate the levels of mature forms of these proteins may be beneficial to the virus. p7 and its upstream

signal sequence (in E2) have been shown to facilitate correct localisation of NS2 adjacent to RCs in the ER which may be important for post-autoprotease roles of NS2 in virus assembly and release (Tedbury et al., 2011). Further to this p7 has been documented to play a role in re-localisation of core from lipid droplets to putative assembly sites in the ER, potentially through a direct interaction of p7 with core which is modulated by interactions with NS2 (Boson et al., 2011). The effects of p7 on both NS2 and core localisation are independent of a functional ion channel activity.

The finding that amantadine, a viroporin inhibitor active against the Influenza M2 viroporin, blocks p7 ion channel activity has led to progression of p7 as an anti-viral target. Effects of adamantine in cell culture and in clinical trials have been modest, but the application of structural NMR techniques and a subsequent model of the p7 oligomeric channel have allowed the identification of improved inhibitors, a process facilitated by virtual high-throughput screening (vHTS) and rational design (Foster et al., 2014).

1.5.4 Non-structural 2

NS2, along with the other non-structural proteins as a NS2-5B polyprotein, is directed to and embedded into the ER membrane by its signal sequence within p7 prior to signal peptidase mediated processing of the p7-NS2 junction. However, expression of NS2 alone leads to localisation at organelle membranes (Yamaga and Ou, 2002). NS2 is 23 kDa and contains two domains. A 92 residue hydrophobic N-terminal trans-membrane domain spans the ER and orientates a 124 residue soluble C-terminal domain in the cytoplasm. This C-terminal domain encodes a cysteine protease which processes the NS2-NS3 boundary, termed the NS2 autoprotease. Functions and current understanding of the C-terminal protease domain of NS2 are discussed in Section 1.8.

The NS2 N-terminal domain is thought to form three trans-membrane alpha helices, each of which has been independently observed in solvent by solution NMR (Jirasko et al., 2008, Jirasko et al., 2010). While the N-terminal TMD is clearly important for anchoring NS2 to ER membranes, the C-terminal catalytic domain has recently been shown to associate with membranes independently, and it is likely that a hydrophobic face of this soluble domain contributes an amphipathic membrane association (Lange et al., 2014). Membrane association of NS2 is essential in the virus lifecycle, at least as a requirement for the NS2 autoprotease activity (Lange et al., 2014), but functional membrane association does not appear to depend on the p7 encoded signal sequence, as demonstrated by viability of virus systems in cell culture where NS2 is genetically separated from p7 with an IRES (Jones et al., 2007).

NS2 plays additional roles in the virus lifecycle, most notably during assembly, following maturation from the NS2-5B polyprotein. Structural evidence suggests that mature NS2 is no longer active as a protease (Lorenz et al., 2006), hence effects on later stages of the lifecycle are thought to be mediated through protein-protein interactions. Numerous studies have documented NS2 interacting either directly or indirectly with the majority of virally encoded proteins. Confirmed interactions with p7 and the E1-E2 heterodimer (most likely through E2 via a signal peptidase host factor complex (Suzuki et al., 2013)) appear to be mediated by the N-terminal TMD (Jirasko et al., 2010, Stapleford and Lindenbach, 2011). A validated interaction with NS3 also requires the NS2 N-terminal domain but this effect could be mediated through interactions with p7.

Interestingly several groups have reported improved systems for culturing HCV through NS2-based chimeras. Lindenbach *et al.* first reported a chimeric virus construct termed J6JFH, with the core-NS2 proteins derived from the J6CF isolate and NS3-5B from the JFH1 isolate (both genotype 2a; these isolates will be discussed in more detail in Section 1.7) (Lindenbach et al., 2005). Infectious titres were improved a further 100-1000-fold by fusing J6CF and JFH1 genomes at a site within NS2 between the first and second N-terminal trans-membrane domains (designated Jc1) (Pietschmann et al., 2006). While all documented interactions with structural proteins occur through the N-terminus of NS2, the C-terminal catalytic domain (but not the catalytic activity) is required for virus assembly (Jones et al., 2007), suggesting the interactions formed through this domain also contribute to late stages of the virus lifecycle. As such NS2 has been postulated as a link between non-structural RC components and sites of virus assembly (Ma et al., 2011).

NS2 has been isolated from high molecular weight complexes which also contain E1, E2 and NS3 (Stapleford and Lindenbach, 2011). The glycosylation pattern of the envelope proteins in these complexes suggest they have not yet trafficked to the Golgi, indicating an early stage of assembly. NS2 has been confirmed to play an early role in assembly, similar to and potentially linked to that of p7 (Jones et al., 2007). It would therefore appear likely that NS2 interactions with E2, p7 and NS3 could occur at LDs prior to egress to the Golgi. Substantial evidence is building for a role of NS3 in release (Ma et al., 2008, Shimakami et al., 2011). An interaction between NS3 and the NS2 C-terminal domain (or an indirect association of NS3 with NS2 via p7) may help explain how NS3 could be targeted to sites of assembly and why the catalytic domain of NS2 is required for virus release. In agreement with this genetically separating NS2

and NS3 with an IRES led to a reduction in NS2-NS3 association and a reduction in virus titre (Stapleford and Lindenbach, 2011).

Investigations into NS2 interacting partners also suggested that NS2 is capable of multimerisation. Subsequently several groups have provided biochemical or structural evidence for the formation of NS2 homodimers (reviewed in Section 1.8.4). Perhaps the strongest evidence comes from a crystal structure of the post-cleaved NS2 catalytic domain which implicates the formation of the autoprotease active site with dimerisation (Lorenz et al., 2006). Whether dimerisation is required for the numerous protein-protein interactions of NS2 remains inconclusive.

Optimisation of the HCV sub-genomic replicon (SGR), as described in Section 1.7, clearly demonstrated that NS2 is not required for genome replication (Lohmann et al., 1999), a finding further confirmed by mutagenesis in infectious virus (Jones et al., 2007). In fact the presence of NS2 in HCV SGRs is associated with impaired replication kinetics (Lohmann et al., 1999, Tedbury et al., 2011). The cause of this is unknown but could be due to slower kinetics of processing at the NS2-NS3 boundary, which is not required in SGRs comprising NS3-5B, or alternatively because of NS2 sequestering viral components at sites of assembly rather than replication through the aforementioned protein-protein interactions.

Finally, NS2 has been implicated in the modulation of host signalling. Ectopic expression of NS2 has been shown to affect transcription from a variety of promoters by causing ER stress (von dem Bussche et al., 2010), suggesting post-cleaved NS2 may modulate protein levels of the host cell via its N-terminus. In addition NS2, along with core, E2 and NS5A, has been shown to bind to components of the mitochondrial intrinsic apoptotic pathway and block apoptosis (Welbourn and Pause, 2006), an area that may be of future interest given the emerging evidence supporting direct oncogenic potential of HCV towards the development of HCC. How these functions, which were mostly observed through expression of NS2 alone, relate to the complete virus lifecycle remains unknown, but the reported short half-life of NS2 within hepatoma cells may aid regulation of these functions. Both NS2 and unprocessed NS2-3 (in the context of abolished autoprotease activity) are rapidly degraded, with NS2-3 precursors subject to ubiquitination on NS2 encoded lysine residues for targeting to the proteasome (Welbourn et al., 2009). For genotype 1a NS2, degradation has been linked to serine 168 in NS2 (Franck et al., 2005). This residue sits in a putative casein kinase II consensus sequence suggesting phosphorylation could regulate turnover, though these results have not been replicated for other genotypes. Mutation of NS2 serine 168 to alanine has been shown to alter

the co-localisation of NS2 and NS5A (Tedbury et al., 2011) and loss of particle assembly due to S168 mutation can be recovered by mutations within NS5A (Yi et al., 2009). While potentially regulating a NS2-NS5A interaction, the precise role of serine 168 in the virus lifecycle has yet to be determined.

Several groups have also reported the observation of an anti-NS2 reactive 17 kDa species from virus lysates which is probably a truncation product, termed tNS2 (Boson et al., 2011, Jirasko et al., 2008, Jirasko et al., 2010). While studies on the protein-protein interactions of NS2 did observe tNS2, it was not isolated from the same p7-, E1-E2- and NS3-containing fractions assumed to be LDs or sites of assembly (Stapleford and Lindenbach, 2011). Whether tNS2 has any functional role is unknown, with some reports suggesting it may represent an aberrant initiation product as it appears upregulated in constructs with genetically separated p7 and NS2 (Ma et al., 2011).

As summarised in Section 1.8.5 despite concerted efforts no small molecule inhibitors of the NS2 autoprotease have been identified. Other than recent reports that DNA aptamers targeting the N-terminus of NS2 block the virus lifecycle (Gao et al., 2014) NS2 has not been pursued as an anti-viral target. Given the number of essential roles NS2 contributes to the virus lifecycle it is an attractive target for novel therapeutic interventions.

1.5.5 Non-structural 3

NS3 is 70 kDa and can be divided into two domains, each with an enzyme activity. The N-terminal 180 residues of NS3 forms a serine protease, while amino acids 181-631 function as an NTPase-dependent RNA helicase.

The N-terminal serine protease functions as the second virally encoded protease for polyprotein processing. Following production of the mature NS3 N-terminus by the NS2-encoded autoprotease activity, NS3, within the NS3-5B polyprotein, cleaves itself at the NS3-NS4A boundary. This reaction occurs in *cis* in that it cannot be trans-complemented by addition of exogenous NS3 protease from an alternate polyprotein. The remaining proteolytic cleavages occur in a specific order. Once the NS3-NS4A junction is processed, NS4A remains associated with NS3 as an essential co-factor in a non-covalent complex. NS4A contributes both to the NS3 protein fold and to the complex's membrane association. The NS3-4A protease can be complemented in *trans* and cleaves first the NS5A-NS5B boundary to release mature NS5B. The resulting NS4A-NS4B-NS5A polyprotein is next processed between NS4A

and NS4B, before the final proteolysis event occurs at the junction of NS4B-NS5A (reviewed in (Morikawa et al., 2011)).

Reconstitution of the NS3-4A protease activity *in vitro* preceded the ability to replicate the HCV lifecycle in cell culture and allowed detailed investigation into the workings of NS3 (Vishnuvardhan et al., 1997). The crystal structure of the NS3 protease domain has been solved to high resolution, revealing a chymotrypsin-like fold stabilised by the coordination of a zinc atom by C97, C99, C145 and H149 (Figure 1.6) (Penin et al., 2004, Romano et al., 2012). Two β-barrel subdomains are formed, with a β-sheet from NS4A contributing to one. The catalytic triad; H57, D81 and S139, function as a classic serine protease, with the acidic residue stabilising a conformation of the histidine capable of deprotonating the catalytic serine and thus allowing nucleophilic attack of the carbon atom within the scissile bond (Figure 1.9C).

Peptide substrates spanning P4-P2' fit into the active site cavity and will be processed, but maximal efficiency is observed with peptides spanning P6-P4' with the consensus sequence D/E-X-X-X-X-C/T | S/A-X-X-X where (|) indicates the scissile bond. An early observation in *in vitro* systems was that N-terminal products from the proteolysis of peptides were capable of inhibiting further proteolysis (Llinas-Brunet et al., 1998). This led to the development of peptide-based NS3 protease inhibitors which now include both linear tetrapeptide α-ketoamides (Han et al., 2000) and second generation macrocyclic inhibitors such as Simeprevir (Rosenquist et al., 2014). The first NS3 protease inhibitors approved for use in the clinic, Telaprevir and Boceprevir, are both the former and bind covalently but reversibly to the NS3 active site (Figure 1.6) (Flores et al., 2009).

The NS3-4A protease plays additional roles in persistence of HCV within host cells by processing components of innate immunity. A proportion of NS3-4A traffics to mitochondrial membranes where it cleaves the mitochondrial antiviral signalling protein (MAVS), toll/IL-1 receptor homology domain-containing adaptor inducing IFN-β (TRIF) and T-cell protein tyrosine phosphatase (TC-PTP). These proteins act as signalling molecules from innate immunity sensors for viral RNA, such as the retinoic acid-inducible I (RIG-I) or toll-like receptor 3 (TLR-3) signalling pathways. The adaptors cleaved by NS3-4A would otherwise induce IFN regulatory factor-3 to upregulate IFN stimulating genes (ISGs) and produce IFN-β, thus helping the host cell and neighbouring cells elicit an 'anti-viral' state (reviewed in (Morikawa et al., 2011)). Thus inhibitors of the NS3-4A protease may also boost the innate immune response to viral infection, further enhancing the anti-viral effect.

The C-terminal NS3 helicase domain is part of the DExH/D-box family of helicases which use ATP hydrolysis to drive the unwinding of double stranded or highly structured RNA (Kwong et al., 2005). Again, the enzyme activity can be reconstituted *in vitro* (Tai et al., 1996) and high resolution crystal structures are available (Appleby et al., 2011, Penin et al., 2004). The NS3 helicase is essential for replication, presumably for the unwinding of RNA duplexes within RCs (Salam and Akimitsu, 2013). The beneficial reasons for encoding these two enzyme activities on one protein, which remains unprocessed, in order to coordinate replication are only recently starting to become apparent (Beran et al., 2009).

Inhibitors of the NS3 helicase have been reported (Salam and Akimitsu, 2013) but challenges regarding selectivity over cellular helicases have hampered progression from pre-clinical development.
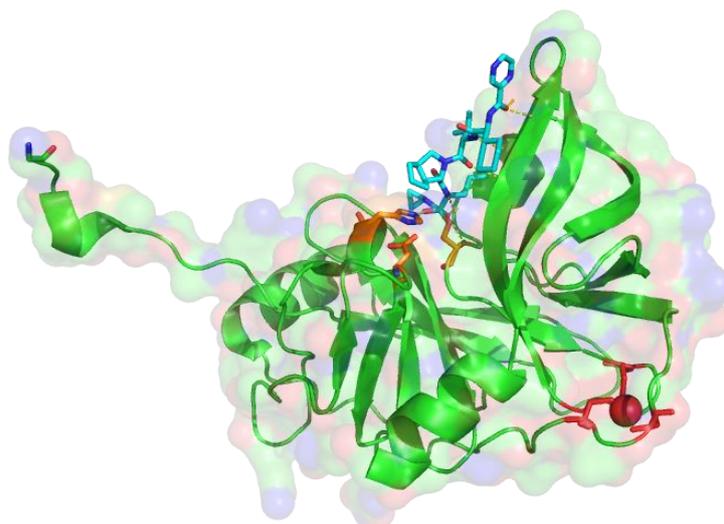
Figure 1.6: NS3 protease – Telaprevir co-crystal structure.

NS3-protease domain with catalytic residues highlighted with orange carbons. The catalytic serine forms a covalent interaction with the protease inhibitor Telaprevir (blue carbons). Additional polar contacts are shown. Structural zinc and co-ordinating residues are shown in red. Surface representation is shown with oxygen atoms in red and nitrogen atoms in blue. Image produced in Pymol using PDB 3SV6 (Romano et al., 2012).

1.5.6 Non-structural 4A

NS4A is a 54 residue co-factor for the NS3 protease which directly contributes a β-sheet to a β-barrel as part of the NS3 fold and serves to anchor NS3 to ER membranes. The first 21 amino acids form a trans-membrane α-helix. While the NS3 protease domain encodes an amphipathic helix for membrane association, through NS3-NS4A association the NS4A TMD is thought to ensure membrane localisation of the NS3 protein (Brass et al., 2008). Amino acids 21-32 form the β-strand contributing to the NS3 fold by interacting with the second β-barrel. The C-terminal residues (amino acids 40-54) are thought to interact with additional HCV proteins and play essential functions both in replication and release (Phan et al., 2011).

1.5.7 Non-structural 4B

NS4B is a highly hydrophobic 27 kDa protein, the central portion of which forms four TMDs (Lundin et al., 2003), while additional amphipathic helices are present at the cytosolic N- and C-termini (Gouttenoire et al., 2009a, Gouttenoire et al., 2009b). The primary function of NS4B is the remodelling of host membranes so as to produce the membranous web upon which RCs form, a process thought to require NS4B oligomerisation (Egger et al., 2002, Paul et al., 2011). However, NS4B has also been documented to play roles in both replication and assembly, most likely through protein-protein interactions (Aligo et al., 2009, Han et al., 2013).

Several groups have reported small molecule inhibitors of NS4B which prevent oligomerisation and show anti-viral effects in SGR systems (Dufner-Beattie et al., 2014, Einav et al., 2008, Rai and Deval, 2011). Compounds targeting NS4B remain at a pre-clinical stage.

1.5.8 Non-structural 5A

NS5A is a RNA-binding phosphoprotein with essential roles throughout the lifecycle. It consists of three domains linked by low-complexity sequences (LCS) and associates with membranes through an N-terminal amphipathic α-helix causing localisation both at ER membranes and LDs. Domain 1 includes a well conserved zinc-dependent structured region that forms two β-sheets, as demonstrated through two independent crystal structures (Love et al., 2009, Tellinghuisen et al., 2004, Tellinghuisen et al., 2005). Interestingly, both of these structures reveal NS5A domain 1 as a dimer, though the position of each monomer relative to the other differs. Considerable evidence has confirmed that NS5A, like several other HCV encoded proteins, functionally self-associates to dimers (or oligomers) (Bhattacharya et al., 2012, Lee et

al., 2011, Lim et al., 2012). The structured domain 1 is considered essential for multiple stages of the virus lifecycle.

Domains 2 and 3 of NS5A are predicted to be less structurally ordered, though these flexible regions may transiently possess significant structure (Feuerstein et al., 2012). Mutational analysis has implicated the C-terminus of domain 2 with a role in genome replication. By contrast, deletion of domain 3 has no effect on replication, but blocks virus assembly (Appel et al., 2008, Ross-Thriepland et al., 2013, Tellinghuisen et al., 2008).

All three domains of NS5A are able to independently and specifically bind viral RNA, in particular RNA from the UTRs (Foster et al., 2010a, Huang et al., 2005, Hwang et al., 2010). This, in addition to the observation of separate co-localisations with NS proteins at the ER and core at LDs, has led to the proposal that NS5A binds and shuttles RNA genomes from the RCs to sites of virus assembly so as to deliver the genome to core (Appel et al., 2008, Targett-Adams et al., 2011). Such a role may allow NS5A to act as a switch between the processes of genome replication and particle assembly, a theory all the more tempting due to the observation of two clearly distinguishable phosphorylation states of NS5A. These states can be resolved by denaturing gel electrophoresis into basally phosphorylated (termed p56 due to apparent molecular weight) and hyper-phosphorylated (p58) species. While both the kinases involved in phosphorylation and the sites of phosphorylation are beginning to become apparent (Huang et al., 2007, Ross-Thriepland and Harris, 2014), the precise roles of each species in the virus lifecycle and any potential effect mediating a 'switch' in the lifecycle have yet to be conclusively demonstrated.

NS5A mediates numerous additional interactions with host factors to perturb signalling pathways and produce a more favourable cellular environment for virus propagation. NS5A is able to dampen the IFN response, in part through a 40 residue region in LCS-I and domain 2 termed the interferon sensitivity determining region, by down regulating protein kinase R (PKR), which would otherwise be activated by viral double-stranded RNA leading to reduced translation. NS5A also down-regulates the mitogen activated protein kinase (MAPK) pathway through extracellular signal-regulated kinase (ERK) signalling, an implication of which is the stalling of the cell cycle in the G_1 phase. Apoptosis is blocked by NS5A through binding and sequestering of p53, down-regulating the TNF-α response and perturbing phosphatidylinositol 3-kinase signalling; while additional NS5A-mediated potential oncogenic effects include production of reactive oxygen species. The effects of these and other NS5A-host interactions are reviewed in (Macdonald and Harris, 2004).

Despite the lack of any documented enzyme activity, NS5A has emerged as a promising antiviral target for treatment of HCV. NS5A inhibitors were first identified through SGR-based high-throughput screening (HTS) and were linked to NS5A by the observation that activity was based on symmetry and the identification of resistance mutants at the dimer interface of the N-terminus of NS5A domain 1 (Gao et al., 2010). Tagged versions of initial inhibitors were subsequently observed to bind solely to NS5A, but only in the context of active replication. Studies demonstrating effects of NS5A inhibitors on the phosphorylation state and localisation of NS5A have suggested they may block conformational changes or other events linked to certain stages of the virus lifecycle, though the mechanism of action remains unknown (Lee et al., 2011, Qiu et al., 2011, Targett-Adams et al., 2011). The remarkable *in vitro* activity of current NS5A inhibitors, some of which have pan-genotype 50% effective concentrations (EC$_{50}$s) in the low pM range, creates a potency paradox whereby significantly fewer molecules of inhibitor than molecules of NS5A are sufficient to block all virus replication within a cell. One proposed explanation is an effect on all NS5A molecules within an oligomeric NS5A complex induced by a single binding event, almost as a 'prion protein' type effect. A more feasible explanation, though not mutually exclusive, may be that NS5A inhibitors only target the small proportion of NS5A functionally involved in RCs. Multiple pieces of evidence support the idea that NS5A inhibitors block the formation of new RCs and as such NS5A inhibitors are also referred to as 'replicase inhibitors' (Lee et al., 2011, Targett-Adams et al., 2011). More recent evidence suggests NS5A inhibitors exert additional effects against virus release (McGivern et al., 2014). Despite their impressive potency *in vitro*, NS5A inhibitors display a low barrier to resistance and as such are likely to be ineffective as a monotherapy.

1.5.9 Non-structural 5B

NS5B is an RNA-dependent RNA-polymerase (RdRp) and the central component of RCs (Behrens et al., 1996), responsible firstly for synthesising a negative strand intermediate from the positive sense genome and subsequently for producing positive strand replicates from the negative sense intermediate. NS5B is 65 kDa and membrane associated through a C-terminal 21 residue TMD (Ivashkina et al., 2002, Schmidt-Mende et al., 2001). The N-terminal ectodomain is cytosolic and performs RdRp functions, with structural data showing a classic 'right-hand' RNA polymerase fold with 'finger', 'palm' and 'thumb' subdomains (Lesburg et al., 1999). The cytosolic region contains a highly conserved GDD motif required for binding divalent cations. Mutation of this motif to GND ablates this ability and prevents RdRp activity (Yamashita et al., 1998). Additionally, NS5B requires functional oligomerisation for RdRp activity (Qin et al., 2002).

Synthesis of the negative strand intermediate is instigated by NS5B binding to the 3' UTR of the positive sense genome via the X-tail (Cheng et al., 1999). Due to the lack of proof-reading activity both the synthesis of negative strand intermediates and of subsequent positive sense copies is highly error prone. This error prone RdRp activity accounts for the high mutagenic rate of HCV which contributes not only to the genetic diversity between HCV genotypes but to the existence of quasi-species within a single host. Quasi-species have important implications for viral resistance to inhibitors as, rather than the selective pressure of an anti-viral having to produce a resistance mutant *de novo*, such a mutation may already exist within one such quasi-species allowing an immediate selective advantage. The error prone nature of NS5B and as a by-product the HCV quasi-species allows for rapid mutagenic change of the dominant HCV genome within the host and hence HCV can quickly react to environmental changes, a key reason for the requirement of combination therapy to successfully treat HCV.

As the central component of RCs NS5B interacts directly or indirectly with numerous other NS proteins. A direct interaction with NS5A has been identified and may regulate the polymerase activity of NS5B depending on the phosphorylation status of NS5A (Ivanov et al., 2009, Shirota et al., 2002) providing further evidence of the ability of HCV to self-regulate various stages of the lifecycle.

NS5B is the subject of two forms of anti-virals; nucleoside analogue inhibitors (NIs) and non-nucleoside analogue inhibitors (NNIs).

NNIs bind to an allosteric site between the 'finger' and 'thumb' domains of the polymerase and block intramolecular contacts involved in coordinated conformational flexibility required during synthesis of RNA (Di Marco et al., 2005). NIs potentially offer more promise as HCV anti-virals due to the high degree of conservation of the NS5B active site which allows pan-genotype activity with a relatively low barrier to resistance. NIs are administered as a pro-drug and require modification by three successive cellular kinases to form the active nucleoside analogue triphosphate which can be incorporated by the RdRp as a chain-terminator. Applying NIs as an anti-viral drug poses a series of problems; these molecules tend to have poor cellular permeability, are not active until the tri-phosphate stage, are influenced by cellular rNTP levels due to the nature of their competitive inhibition and risk incorporation into host cellular polymerases leading to off-target effects and toxicity. The success of the most advanced NI Sofosbuvir was driven principally by toxicology and pharmacology analysis, such as the use of a mono-phosphate pro-drug to aid absorption, distribution, metabolism and elimination (ADME) properties and increase levels of the intracellular triphosphate form (Coats et al., 2014). While

Sofosbuvir now appears set to form the backbone of all-oral HCV combination therapies, the inherent risks of NIs as drug molecules, as evidenced through numerous failed clinical trials of other NIs, should not be forgotten.

1.6 Current anti-viral strategies for HCV

The most developed therapeutic options for the treatment of HCV in IFN-RBV free therapy are the direct acting anti-virals (DAAs) previously described; namely NS3-4A protease inhibitors, NS5A 'replicase' inhibitors and NS5B NIs. Many of these inhibitors in combination with IFN-RBV or in IFN-RBV-free combinations with each other are expected to be approved shortly (clinical trial data for lead compounds in each series was summarised in Section 1.3). While some NS5B NNIs are also in clinical trials, additional DAAs comprising core inhibitors, p7 viroporin inhibitors, NS3 helicase inhibitors and NS4B inhibitors have, for the most part, remained in pre-clinical stages.

Inhibitors of HCV entry have been reported, again at a pre-clinical stage. These compounds exert an anti-viral effect at an undefined entry event post-attachment. While resistant isolates link these compounds to the C-terminal TMD of E2, it remains possible that they could function through a cellular target (Baldick et al., 2010, Coburn et al., 2012).

Additional anti-HCV strategies have attempted to target essential host factors with small molecule inhibitors. This strategy offers the advantage of reduced likelihood of viral resistance and will often exert a pan-genotype effect.

One such host factor is cyclophilin A (CypA), a chaperone protein which acts as a prolyl-peptidyl isomerase (PPI) to rectify misfolded *cis* peptide bonds prior to proline residues. CypA was shown to be required for HCV replication as a by-product of the anti-viral activity of the CypA antagonist cyclosporine A (CsA) (Nakagawa et al., 2004, Watashi et al., 2003); a cyclic undecapeptide natural product already in clinical use as an immunosuppressive (Matsuda and Koyasu, 2000). The anti-viral effect of CsA is through inhibition of the PPIase activity of CypA and as such non-immunosuppressive analogues of CsA which still block PPIase activity have been progressed as anti-virals targeting HCV (Paeshuyse et al., 2006), though the lead compound Alisporovir (DEB025) is currently on hold in Phase III due to problems with toxicity.

While the majority of evidence now links the mechanism of action of CsA to CypA rather than alternative Cyp family members, the effect on the HCV lifecycle is not fully understood

(Chatterji et al., 2009, Yang et al., 2008). Some reports have suggested a NS5B-Cyp interaction, though the common consensus is that CypA binds to domain 2 and LCS-II of NS5A where the PPIase activity is required for correct folding of two proline-rich regions. Resistance to CsA maps within this region, such as the NS5A mutant D316E (JFH1 isolate numbering), with biochemical and nuclear magnetic resonance (NMR) data demonstrating resistance does not affect NS5A-CypA binding, rather naturally favours the *trans* state of surrounding proline residues (Coelmont et al., 2010, Fernandes et al., 2007, Grise et al., 2012, Yang et al., 2010). As these regions of NS5A are primarily disordered, the requirement for CypA in terms of the HCV lifecycle remains elusive, although *in vitro* evidence suggests CypA induced isomerisation may aid the ability of NS5A domain 2 to bind RNA (Foster et al., 2011), while indications of competitive binding for the same sites on NS5A by CypA and NS5B are intriguing (Rosnoblet et al., 2012). The anti-viral effect of CsA and its analogues are further complicated by enhanced sensitivity in the presence of NS2, as described in Section 1.8.5.

A prophylactic or therapeutic vaccine to HCV would be attractive not only due to the problem of DAA resistance but the significant costs of current DAA therapies. Evidence does exist for impaired HCV infection following previous spontaneous HCV clearance, suggesting cross-protective immunity can occur naturally. However the lack of a suitable animal model makes vaccine development difficult, while the high HCV mutagenic rate and genetic diversity suggest there is a potential for escape from the immune responses induced by a vaccine and make it hard to produce a pan-genotype cross-protective vaccine (Feinstone et al., 2012).

1.7 HCV model systems

For the first decade of HCV research the virus lifecycle could not be reproduced in cell culture. For this reason early drug development targeted enzyme activities that could be reconstituted *in vitro* into suitable high-throughput screens; namely the NS3 protease, the NS3 helicase and the NS5B RdRp. The ability to study and screen non-discriminately for inhibitors of HCV genome replication was only realised upon production of the first sub-genomic replicon (SGR) capable of reproducing the HCV replication cycle. This was achieved through a bi-cistronic construct wherein the first ORF, driven by the HCV IRES from the 5' UTR, encodes for a reporter gene and the second ORF, driven through an encephalomyocarditis virus (EMCV) IRES, contains the NS3-5B replication cassette (in addition to the non-coding 3' UTR) (Figure 1.7). Important for successful establishment of autonomous RNA replication of the SGR was the use

of a selectable marker, neomycin phosphotransferase (NPT) and of a certain hepatoma cell line; Huh7 (Nakabayashi et al., 1982). The use of an NPT reporter for resistance to the otherwise cytotoxic G418 (Geneticin) drove the selection of culture adaptive mutations (CAMs) which, at least in early genotype 1-derived SGRs, were paramount for achieving detectable levels of replication (Lohmann and Bartenschlager, 2014, Lohmann et al., 1999).

Since the development of the SGR several groups have documented a detrimental effect on replicative ability when the second ORF includes the NS2 protein so as to span NS2-5B (Figure 1.7) (Lohmann et al., 1999, Tedbury et al., 2011). The reasons for this are unknown but are speculated in Section 1.5.4, while data presented in Chapter 3 will further examine this effect.

The identification of a genotype 2a strain of HCV from a patient with fulminant hepatitis, termed JFH1, significantly improved levels of replication without the need for CAMs (Kato et al., 2003) such that the incorporation of a firefly luciferase reporter (ffLuc) in place of NPT allowed rapid and accurate quantification of replication kinetics (Targett-Adams and McLauchlan, 2005). To date the majority of HTS performed against HCV replication has used JFH1 or culture-adapted genotype 1 SGRs spanning NS3-5B with a variety of reporters (Kwong, 2014).

In 2005 the entire HCV lifecycle was successfully reproduced in cell culture using full JFH1 molecular clones (Zhong et al., 2005), thus allowing studies on virus entry which had previously relied on pseudotyping (HCVpp) and exploration of virus assembly and egress. Systems for studying the cell culture based HCV lifecycle (HCVcc) were enhanced further though NS2-based chimeras J6JFH and Jc1, as described in Section 1.5.4 (Lindenbach et al., 2005, Pietschmann et al., 2006), while the p7-NS2 junction also proved amenable to reporter insertion (Jones et al., 2007) (Figure 1.7).

While Huh7 cells are by far the most permissive to the HCV lifecycle in cell culture, certain sub-populations of this heterogeneous cell line appear more permissive than others. Clearing selectable SGRs from Huh7 cells with IFN or a DAA therefore yields a population of cells more permissive to the HCV lifecycle. This has been used to produce cleared-SGR cell lines with enhanced SGR replication kinetics, termed Huh7.5 cells (Blight et al., 2002). The enhanced replication kinetics observed in Huh7.5 cells have been partially attributed to defective RIG-I, leading to a reduced immunogenic response to the presence of cytoplasmic viral RNA (Sumpter et al., 2005). However, the re-introduction of functional RIG-I still yielded a cell line

more permissive than naïve Huh7 cells (Binder et al., 2007), suggesting these cells may also benefit from upregulated pro-viral factors or additional down-regulated anti-viral factors.

The range of genotypes amenable to study through SGRs or chimeric HCVcc has significantly expanded since the development of the original construct, and continues to do so. Use of the HCVcc system in primary hepatocytes has also been reported providing a system with greater physiological relevance (Ploss et al., 2010, Podevin et al., 2010). Animal models for studying HCV remain limited to chimpanzees and the use of the related virus GBV-B in tamarins. The shortage of suitable animal models negatively impacts the later stages of HCV drug discovery, though the use of humanised livers in mice has recently expanded these options (Dorner et al., 2013, von Schaewen and Ploss, 2014).

Current studies utilising SGRs usually employ either NPT or ffLuc reporters. The SGR is placed under the control of a T7 promoter on a vector allowing transcripts to be produced *in vitro* from linearised DNA. For SGRs with a NPT reporter (SGR-neo), transcripts can be electroporated into cells which are then subjected to a selective pressure with G418. The resulting replicon-harbouring cell lines can be maintained indefinitely under G418 selection and are thus termed 'stable' SGRs. Transcripts of SGRs with a ffLuc reporter (SGR-luc) can be electroporated into cells allowing rapid and sensitive quantification of replication through luciferase quantification up to 72-96 hours. However, as they lack a selectable marker these constructs cannot be maintained as stable cell lines and transcripts must be prepared and electroporated at the outset of each assay. As such SGR-luc constructs are known as 'transient' SGRs. For the purposes of screening this problem has been overcome using a ffLuc-NPT fusion protein within the first ORF (SGR-feo). These constructs allow selection and maintenance of stable cell lines that constitutively produce luciferase as a measure of replication (Wyles et al., 2009, Wyles et al., 2007) and can therefore act both as 'transient' and 'stable' SGRs (Figure 1.7).

As HCVcc requires the RNA to be packaged into infectious virions, minimising the size of the reporter gene (and hence the size of the RNA that is packaged) is thought to improve the number of successfully assembled infectious virions. Elevated levels of reporter signal both for replication kinetics and from infection of naïve cells (as a measure of released infectious virus) have been observed using a comparatively small Nanoluciferase (Nluc) reporter (Promega) (Yutaka Amako, unpublished data). As such Jc1-Nluc virus offers the highest reporter signal for investigation of the full HCV lifecycle.
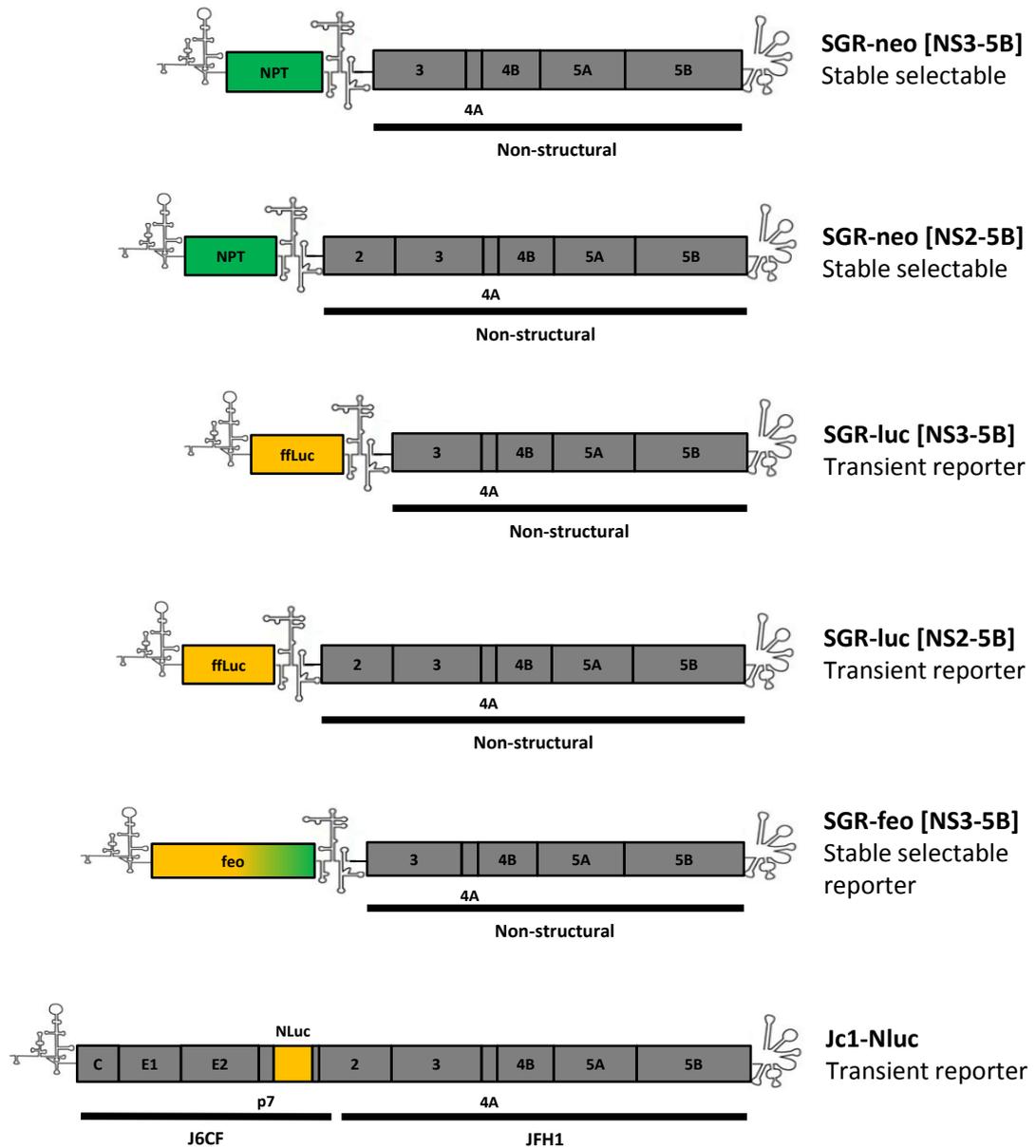
Figure 1.7: Current SGR and virus constructs relevant to this study.

SGRs with neomycin phosphotransferase (NPT; SGR-neo) are amenable to the selection of stable cell lines using G418. SGRs with firefly luciferase (ffLuc; SGR-luc) provide a rapid and sensitive reporter but require transient electroporation into cells. SGR-neo and SGR-luc have been produced spanning either NS3-5B or NS2-5B. SGRs with NPT-ffLuc fusion proteins (SGR-feo) combine the benefits of selecting stable cell lines with a rapid and sensitive reporter, but to date are only available spanning NS3-5B. Jc1 is a chimeric strain which has enhanced lifecycle kinetics in cell culture, with structural proteins and the NS2 N-terminus derived from J6CF and the remaining portion of NS2 and NS3-5B derived from JFH1. Jc1 Nluc contains a Nanoluciferase reporter between p7 and NS2.

1.8 The NS2-3 autoprotease

1.8.1 The NS2 cysteine protease

Initial investigations into the processing of the HCV polyprotein identified roles for host signal peptidases and the NS3 protease, but the mechanism of processing at the NS2-NS3 junction remained unknown. In 1993 two groups reported the identification of a second virally-encoded protease which was independent of NS3 protease activity and proposed to be a zinc-dependent metalloproteinase (Hijikata et al., 1993). The protease activity was found to be encoded within the C-terminus of NS2 but requiring the NS3 N-terminal protease domain. Mutation of residues histidine 952 or cysteine 993 within NS2 to alanine abolished this activity, leading to the proposal of a cysteine protease, while mutation of the NS3 protease active site residues did not affect NS2-NS3 processing (Grakoui et al., 1993). Furthermore, co-transfection of H952A and C993A mutants suggested *trans* cleavage could occur, albeit inefficiently. Mutagenic studies of the substrate cleavage sequence of the NS2 protease revealed that it could tolerate most substitutions other than those expected to cause a significant conformational change, such as the introduction of proline at P1 or P1' positions or mutating both P1 and P1' to alanine (Reed et al., 1995).

Several groups subsequently set out to characterise this second viral protease. Radiolabelled *in vitro* translation revealed that while cleavage at the E2-p7-NS2 boundaries required addition of microsomal membranes, supporting a requirement for host factors such as signal peptidase, processing of the NS2-NS3 junction required only the addition of detergents, confirming that host factors were not a requirement for activity (Pieroni et al., 1997). To attempt to characterise the nature of the protease a range of protease inhibitors were tested for the ability to block proteolysis. A range of serine, cysteine and aspartyl protease inhibitors had no effect on the NS2 protease (Table 1.2). This observation is unsurprising as the NS2 C-terminal domain shows no homology to any of these protease families. Phenyl-methane-sulfonylfluoride (PMSF), a covalent serine protease inhibitor, also had no effect, but an irreversible cysteine protease inhibitor; tosyl phenylalanine chloromethyl ketone (TPCK) did block processing at a concentration of 500 μM. Additional alkylating agents tosyl-L lysine chloromethyl ketone (TLCK), N-ethyl maleimide (NEM) and Iodoacetic acid (IAA) or Iodoacetamide (IAM) showed similar activity (Table 1.2). In contrast an irreversible cysteine protease inhibitor based around an epoxide, E64 (Figure 4.1C), showed no activity, although this compound was only tested at 10 μM. Additional observations noted that the protease activity was enhanced by the addition of DTT, suggesting the requirement of reducing

conditions and the correct redox state of the catalytic cysteine, while the reaction could also be inhibited with zinc chelators.

Similar results were reported by successfully reconstituting protease activity with bacterially expressed protein. The truncation of NS2 to remove the N-terminal hydrophobic domain and the C-terminus of NS3 (leaving the NS2 C-terminal domain and the NS3 N-terminal domain; polyprotein residues 907-1206) increased the solubility of bacterially expressed protein and further defined the minimal unit required for protease activity. Overexpressed protein required purification under denaturing conditions and dilution into a glycerol and detergent rich buffer to instigate proteolysis (Pallaoro et al., 2001). The resulting protease reaction was also sensitive to TPCK, IAM and NEM inhibition (at 500 µM) as well as EDTA (2 mM) (Table 1.2). However unlike with *in vitro* observations previously made with the NS3 protease, activity was not inhibited by peptides that are the result of proteolysis. N-terminal product, C-terminal product and complete peptide substrates showed no activity in the µM range. Similar peptides conjugated to mechanism-based protease inhibitors also showed no activity. Mechanism-based inhibitors, also termed suicide inhibitors, act as a substrate analogue to form a non-reversible covalent complex through the catalytic function of the active site residues using reactive electrophilic warheads. However such warheads in the form of aldehydes and hydroxamic acid moieties as well as classic metalloprotease inhibitors conjugated to N-terminal pentapeptides (P5-P1) did not show activity against the NS2 autoprotease (Pallaoro et al., 2001).

In contrast a similar assay utilised by Thibeault *et. al* did reveal weak activity of substrate peptide sequences against *in vitro* NS2-mediated autoproteolysis. While a range of protease inhibitors were again reported to be inactive, some peptide sequences did show activity, the most potent of which was an N-terminal decapeptide with an EC_{50} of 90 µM (Thibeault et al., 2001) (Table 1.2). These results differ from those reported by Pallaoro *et. al* and whether N-terminal proteolysis products inhibit the NS2 protease remains controversial. Peptides derived from NS4A showed greater activity, with EC_{50}s as low as 600 nM (Thibeault et al., 2001). However these peptides are likely to exert their action through a conformational effect on NS3 rather than directly against the NS2 encoded protease, further confirming a role of the NS3 protease domain (though not the NS3 protease activity) on the NS2 autoprotease.

Table 1.2: Current chemicals and peptides tested against the NS2-3 protease *in vitro*.

N.D. = not determined.

Chemical (peptide sequence)	Action	(Pieroni et al., 1997) % inhibition (concentration)	(Pallaoro et al., 2001) % inhibition (concentration)	(Thibeault et al., 2001) % inhibition (concentration)
Antipain	Trypsin and papain inhibitor	0% (0.15 mg/ml)	N.D.	N.D.
Aprotinin	Trypsin inhibitor	0% (0.5 mg/ml)	N.D.	0% (1 mg/ml)
Pepstatin	Aspartyl protease inhibitor	0% (0.5 mg/ml)	N.D.	0% (0.01 mg/ml)
Pefabloc	Serine protease inhibitor	N.D.	N.D.	0% (1 mg/ml)
Leupeptin	Serine/ cysteine protease inhibitor	N.D.	N.D.	0% (0.1 mg/ml)
E64	Epoxide-based cysteine protease inhibitor	0% (10 µM)	N.D.	0% (0.2 mg/ml)
Captopril	Metalloprotease inhibitor	N.D.	N.D.	0% (1 mM)
PMSF	Irreversible serine protease inhibitor	0% (3 mM)	N.D.	N.D.
TPCK	Irreversible cysteine protease inhibitor	50% (500 µM)	35% (500 µM)	50% (500 µM)
TLCK	Irreversible trypsin-like serine protease inhibitor	5% (500 µM)	N.D.	5% (500 µM)
NEM	Thiol alkylating agent	80% (4 mM)	70% (500 µM)	80% (4 mM)
IAA/ IAM	Alkylating agent	80% (500 µM)	25% (500 µM)	100% (1 mM)
EDTA	Zn^{2+} chelator	60% (2 mM)	60% (2 mM)	100% (2 mM)
EGTA	Ca^{2+} chelator	20% (2 mM)	N.D.	N.D.
1,10-phenanthroline	Zn^{2+} chelator	70% (2 mM)	N.D.	80% (1 mM)
(GWRLL)	P5-P1 peptide	N.D.	0% (100 µM)	N.D.
(SFEGQGWRLL)	P10-P1 peptide	N.D.	N.D.	50% (90 µM)
(APITAY)	P1'-P6' peptide	N.D.	N.D.	0% (1 mM)

1.8.2 The role of NS3 and zinc on the NS2 autoprotease

Several studies have confirmed that NS2-NS3 proteolysis can be blocked by zinc chelators (Table 1.2). Mutagenesis of residues C1123, C1125 and C1171 within the N-terminus of NS3 all abolished NS2 autoprotease activity (Hijikata et al., 1993, Tedbury and Harris, 2007), while structural data of the NS3 protease domain showed that these residues co-ordinate zinc (Figure 1.6), alongside a water-mediated interaction with H1175. Based on the observations that zinc chelators blocked NS2-NS3 processing, that water-mediated co-ordination is often indicative of catalytic zinc and that the NS3 co-ordinated zinc is within 11 Å of the site of NS2-NS3 proteolysis at the NS3 N-terminus, a role for the NS3 co-ordinated zinc in catalysis was postulated (Wu et al., 1998). However mutagenic analysis still implicated residues H952, E972 and C993 of NS2 as the catalytic triad of a cysteine protease.

The NS2 autoprotease was confirmed as a cysteine protease upon crystallisation of the post-cleaved NS2 catalytic domain, termed NS2^{pro} (Lorenz et al., 2006). This structure revealed a classic cysteine protease catalytic triad with a glutamic acid and a histidine positioned for coordination and de-protonation of a catalytic cysteine from an alternative monomer (Figure 1.9D). As such, NS2^{pro} appears to form two composite cysteine protease active sites at the dimer interface. Interestingly however, this crystal structure did not contain zinc.

A comparative analysis of the role of zinc within both the NS2 protease and the NS3 protease suggested that processing at the NS2-NS3 boundary was more sensitive to zinc chelation, indicative of a tighter zinc association for the NS3 protease activity, perhaps through a conformational change which occurs post NS2-NS3 cleavage (Tedbury and Harris, 2007). In agreement with this proposal, mutation of H1175 in NS3 reduced NS3 protease activity with little effect on NS2-NS3 processing. An additional cysteine within NS2, C922, was identified as a requirement for NS2 autoprotease activity but was not necessary for the NS3 protease. As such it was suggested that NS2-encoded C992 and NS3-encoded C1123, C1125 and C1171 coordinate zinc in the NS2-NS3 pre-cursor and are required for NS2 autoprotease activity. Following processing from NS2, a structural rearrangement allows NS3 to coordinate zinc through C1123, C1125 and C1171 as well as H1175 (Figure 1.8A). In both cases the role of zinc appears to be entirely structural. The addition of NS4A-derived peptides could favour the conformation required for NS3 protease activity, thus inhibiting NS2-NS3 autoproteolysis (Thibeault et al., 2001).

The apparent requirement of the NS3 protease domain for processing of the NS2-NS3 precursor begs the question of whether the NS2 C-terminal domain is a stand-alone protease,

as indicated by the post-cleaved NS2^{pro} crystal structure. Previous analysis of the minimal subunit required for NS2-NS3 processing agreed with a requirement for zinc coordination as at least the first 181 residues of NS3 were required.

Following this, expression of a tagged NS2 construct with only the first 40 residues of NS3 followed by a GST tag demonstrated detectable levels of proteolysis, while proteolysis could be further stimulated by the addition of the remainder of NS3 (residues 41-213) or the remainder of the NS3 'zinc-binding' domain (residues 41-180) (Schregel et al., 2009). Furthermore, NS2 followed by only two residues of NS3; P1' and P2' residues alanine and proline, was capable of proteolysis in the same system, though in the absence of NS3 residues 3-40 *in cis*, the addition of NS3 residues 41-213 had no enhancing effect. These findings demonstrated that the protease activity responsible for NS2-NS3 processing is entirely encoded within the NS2 C-terminal domain, with the NS3 protease domain only responsible for enhancing proteolysis. This 'enhancing' effect of NS3 was further characterised, with NS2 requiring the first 60 amino acids of NS3 *in cis* in order to be fully stimulated by the NS3 'zinc-binding' domain *in trans*. The mechanism of this enhancing role of NS3 remains ill-defined. It is possible NS3 contributes to some degree to the NS2 autoprotease active site, or that its role is mainly for optimal orientation of the substrate cleavage sequence. In any case, the NS2 autoprotease can be regarded, to some degree, as a stand-alone protease, with the enzyme activity entirely encoded within the NS2 catalytic C-terminal domain present in the post-cleaved crystal structure.

1.8.3 Relevance of the post-cleaved NS2 crystal structure

As previously stated, the crystal structure of the post-cleaved NS2 catalytic domain suggested a functionally relevant dimer. The structure of each monomer contains two sub-domains, an N-terminal sub-domain comprised of two anti-parallel α-helices linked by extended loops and turns and a C-terminal sub-domain forming an anti-parallel β-sheet (Figure 1.8D). Within the dimer, the N-terminal sub-domain of one monomer interacts with the C-terminal sub-domain of the other; with the extended linkers joining the two sub-domains crossing in the central part of the ensemble (Figure 1.8B). In this way, a β-strand from the linker of one monomer forms part of the β-sheet in the C-terminal sub-domain of the alternate monomer.

Essential residues H952 and E972 (NS2^{pro} residues H143 and E163) lie on a loop within the extended linker, while the catalytic C993 (residue C184 in the crystal structure) is positioned on a loop between β-strands in the C-terminal sub-domain (Figure 1.8D). It is apparent that these three residues are coordinated so as to produce a cysteine protease with serine protease

geometry (Figure 1.9). However, H143 and E163 within the linker domain appear to coordinate C184 from the alternate monomer (Figure 1.8B, C). The post-cleaved NS2^{pro} structure therefore represents a completely novel protein fold in addition to the first documented protease that forms two composite active sites at the dimer interface (Figure 1.8B) (Lorenz et al., 2006).

The N-terminal domain helices of this structure contain a high proportion of hydrophobic residues. As a crystal this region contained several molecules of detergent, suggesting a peripheral membrane association. It was subsequently confirmed that a GFP-NS2^{pro} fusion protein localises to ER or mitochondria derived membrane, that deletion of the N-terminal α-helices of the NS2 catalytic domain abolishes this effect, and that these helices alone are sufficient to target GFP to membranes (Lange et al., 2014). This effect was further mapped to a positively charged residue at position 131 or 134 within the second N-terminal α-helix of NS2^{pro} (Figure 1.8B, D). Mutagenesis at this site ablated GFP-NS2^{pro} membrane localisation. This mutation also blocked NS2-NS3 polyprotein processing (though in this context, when the NS2 N-terminal TMD is present NS2 remained membrane bound). In contrast, this residue did not appear to significantly impact other, post-cleavage roles of NS2 on the virus lifecycle.

The C-terminus of the post-cleaved NS2 crystal structure is the P1 residue of the substrate cleavage sequence. This residue, L217, remains within the active site cavity, forming several polar contacts to the side chains of H143 and C184 and to the backbone nitrogen of C184 (Figure 1.8C). It has been proposed that this residue, and others from the C-terminal β-strand of NS2 which form part of the substrate sequence, may have an inhibitory effect by preventing further protease activity following NS2-NS3 proteolysis (Lorenz et al., 2006). A similar inhibitory conformation of a C-terminal residue has been observed in the Sindbis virus capsid serine protease active site (Tong et al., 1993) (PDB 2SNV).
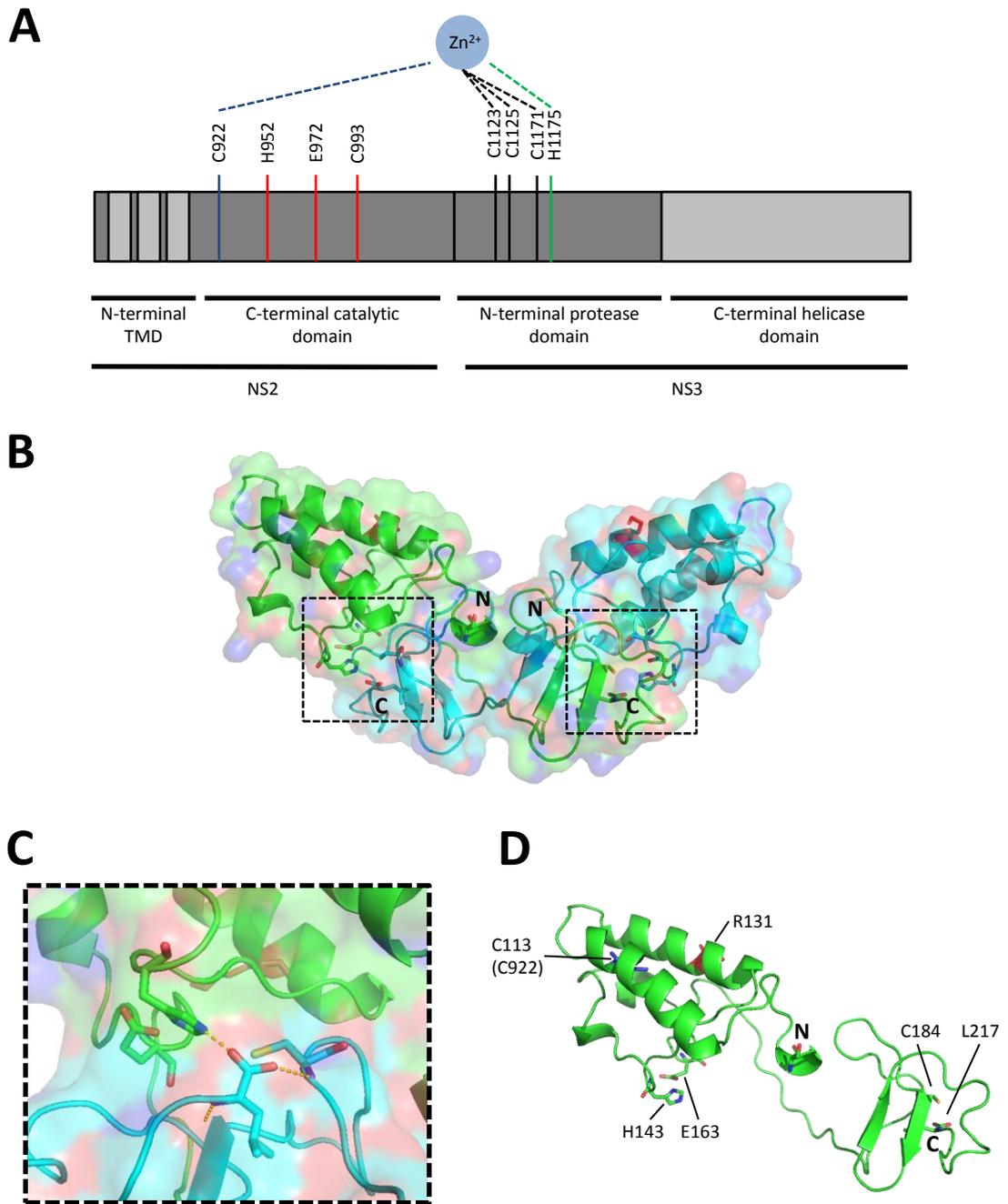
Figure 1.8: Structural features of the NS2 autoprotease.

A) Schematic of the NS2 and NS3 proteins. NS2 catalytic residues are shown (red bars). NS3 residues C1123, C1125 and C1171 (black bars) coordinate zinc. For NS2 protease activity, C922 in NS2 is the fourth zinc coordinating residue (blue lines). For NS3 protease activity following NS2-NS3 processing H1175 in NS3 is the fourth zinc coordinating residue (green lines). B) Crystal structure of the NS2^pro catalytic domain homodimer. One monomer is shown with green carbons, the other with blue carbons. Residue 131 required for membrane association is highlighted in red. Catalytic residues forming a composite active site at the dimer interface are

highlighted in dotted boxes. N- and C-termini are shown. Surface representation is shown with oxygen atoms in red and nitrogen atoms in blue. C) C-terminal L217 forms contacts with catalytic residues in the active site. E972 (E163) and H952 (H143) from one monomer chain and C993 (C184) from another can be seen forming the catalytic triad. D) Monomeric form of the NS2^{pro} catalytic domain crystal structure with the same residues highlighted. Zinc coordinating C992 (C113) is shown in blue. For an alternative orientation of the NS2^{pro} monomer see Figure 3.1. Images made in Pymol from PDB 2HD0.

As described in Section 1.8.2, the catalytic domain of NS2 present within this structure is sufficient for basal protease activity. The post-cleaved NS2 crystal structure, solved to a resolution of 2.28 Å, is therefore likely to represent a reasonable model for the proteolytic activity of NS2, with the caveat that following NS2-NS3 processing, the new C-terminal residue of NS2, L217, appears to remain coordinated within the active site, most likely to exert an inhibitory effect. Despite this, the active site of NS2 in the crystal structure, in particular the catalytic triad, appears to be present in a physiologically relevant context. Comparison of the catalytic triad (Figure 1.9F) to well-refined structures of classic serine and cysteine proteases Chymotrypsin (Figure 1.9A) and Papain (Figure 1.9B) reveal similar spatial orientation between these residues with respect to a cysteine protease with serine protease geometry, as in the Poliovirus 3C capsid structure (Figure 1.9E). Similar conformations are also observed in structures of serine or cysteine proteases covalently bound to mechanism-based inhibitors, such as the HCV NS3 serine protease in complex with Telaprevir (Figure 1.9C) and the cysteine protease Calpain in complex with ZLAK3001; an α-ketoamide based inhibitor (Figure 1.9D), indicative of a physiologically relevant orientation.

Distances between heteroatoms of the acidic residue responsible for correct coordination of histidine are reported for all of the structures from Figure 1.9 in Table 1.3, as are distances between the histidine nitrogen and the catalytic alcohol or thiol which it deprotonates. The similarities between these distance constraints again emphasises that the orientation of catalytic residues in the NS2^{pro} crystal structure are in an enzymatically active conformation. Polar contacts of L217 within the active site to the residues responsible for coordination of a proteolysis transition state further support this model (Figure 1.8C). The stabilisation of a transition state within the proteolysed peptide bond is an essential feature of these proteases,

and is achieved through additional polar contacts formed by the 'oxyanion hole'. In this respect, a contact between L217 and the backbone nitrogen of C184 (Figure 1.8C) adds further evidence that the conformation of the active site in the structure remains relevant to an active protease state.

The structure described herein does not contain the NS3 N-terminal domain required to further enhance NS2 autoprotease activity. A contribution of this region of the polyprotein to the active site, or even to residues forming the oxyanion hole, cannot be comprehensively ruled out. However, given the evidence for a catalytically active conformation of the NS2^{pro} structure and of basal NS2 protease activity in the absence of NS3, this region may just be influencing the NS2 autoprotease through optimal orientation of the substrate sequence. Nonetheless, interpretation of the crystal structure of the post-cleaved NS2 catalytic domain as a model of the NS2 autoprotease should take into account the potential for NS3-encoded contributions which are not present.

Structural data of the NS2-NS3 precursor would greatly aid understanding of the catalytic mechanism of the NS2 autoprotease and conformational changes occurring in NS2 and NS3 following processing. Such a structure would be more representative of the active form of the NS2 autoprotease and may be of more use for structure-based drug design. With this aim in mind, active site mutants which prevent NS2-NS3 processing have been analysed *in vitro* using tryptophan fluorescence studies and circular dichroism, revealing all of these mutants not only remove key components of the catalytic triad but also have significant effects on the overall protein fold of the NS2-NS3 precursor. As such structural studies into active site mutants would not yield an enzymatically active conformation. Structural studies of the pre-cleaved NS2-NS3 subunit may prove more successful through the use of a specific small molecule inhibitor to prevent autoproteolysis without disrupting the protein fold (Foster et al., 2010b).

A) Chymotrypsin — Serine protease

B) Papain — Cysteine protease

C) HCV NS3 protease - Telaprevir — Serine protease

D) Calpain – α-ketoamide — Cysteine protease

E) Poliovirus 3C capsid — Cysteine protease with Serine protease geometry

F) HCV NS2 protease — Cysteine protease with Serine protease geometry
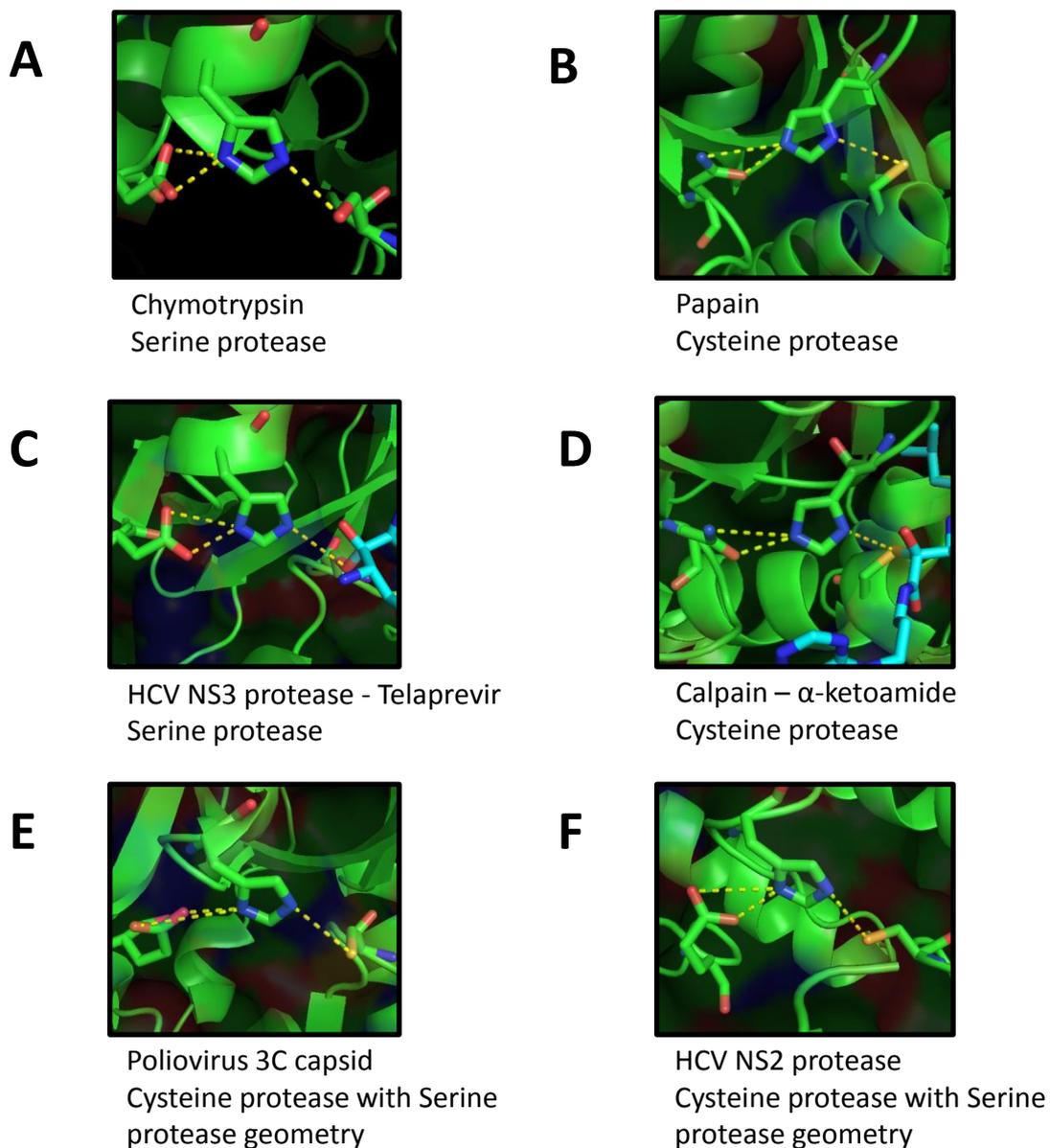
Figure 1.9: Catalytic residues of the post-cleaved NS2 protease structure in an enzymatically active conformation.

A) Catalytic triad of the serine protease Chymotrypsin. (PDB 4H4F). B) Catalytic triad of the cysteine protease Papain. (PDB 1PPN). C) Catalytic triad of the HCV serine protease NS3 in complex with mechanism-based inhibitor Telaprevir. (PDB 3SV6). D) Catalytic triad of the cysteine protease Calpain in complex with an α-ketoamide mechanism-based inhibitor. (PDB 2R9C). E) Catalytic triad of the Poliovirus 3C capsid protease, a cysteine protease with serine protease active site geometry. (PDB 1L1N). F) Catalytic triad of the HCV post-cleaved NS2 autoprotease. (PDB 2HD0).

Table 1.3: Distance constraints of protease catalytic triads.

Distances for indicated bonds of each protease structure from Figure 1.9. Distances were derived from PDB codes indicated in Figure 1.9 and calculated in Pymol.
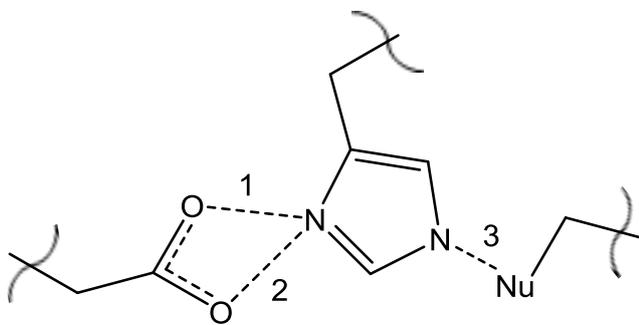
Protease structure	Distance 1 (Å)	Distance 2 (Å)	Distance 3 (Å)
Chymotrypsin	3.2	2.7	3.7
Papain	3.9	2.6	3.8
NS3-Telaprevir	3.4	2.6	3.9
Calpain-α-ketoamide	4.1	2.7	3.6
Poliovirus 3C capsid	4.7	2.7	3.6
NS2 autoprotease	3.9	2.7	3.7

1.8.4 Evidence of NS2 homo-dimerisation

Based on the formation of a composite active site within the crystal structure, activity of the NS2 autoprotease would be expected to require NS2 homo-dimerisation. Biochemical analysis by analytical ultracentrifugation and cross-linking experiments with the bacterially expressed NS2^{pro} used for crystallisation confirmed a dimeric species in solution. To confirm this requirement within the context of cellular NS2-NS3 processing, separate HCV polyproteins containing H143A or C184A mutants respectively were co-expressed, with detectable levels of NS2-NS3 processing observed. As these catalytic residues are drawn from different monomers within the composite active site, a dimer formed from a H143A monomer and a C184A monomer should have one active site with two ablated catalytic residues and one functional active site. Both of these mutants alone were unable to process NS2-NS3, so the observation of processed NS2 and NS3 when the constructs were co-expressed is evidence that the two defective constructs can form a dimer and that functional NS2 proteolysis can be performed through a composite active site formed by the two constructs (Lorenz et al., 2006).

Additional evidence supports the theory that NS2 dimerisation may be required for protease function. Previous analysis of the NS2 autoprotease *in vitro* had noted that the reaction was

concentration dependent (Pallaoro et al., 2001), while similar *in vitro* studies have confirmed the observation of dimers formed from a construct spanning the NS2 catalytic domain and the NS3 N-terminal protease domain using cross-linking reagents (Foster et al., 2010b). Many of the residues located at the dimer interface within the crystal structure are conserved across genotypes. An additional study, which confirmed that cellular overexpression of NS2 and of the NS2^{pro} C-terminal domain produced dimers, found that mutagenesis of conserved residues at the dimer interface reduced levels of NS2 dimerisation and impacted HCV replication and hence the release of infectious virus (Gorzin et al., 2012). While the above data support the ability of NS2 to form dimers, demonstrating this at physiologically relevant concentrations has proved challenging.

1.8.5 NS2 as a target for anti-virals

Inhibitors of the NS3 protease are already approved for use in the treatment of certain HCV genotypes. Targeting a virally-encoded protease in order to block polyprotein processing is therefore an established route to exert an anti-viral effect and produce therapeutic effects in the clinic.

While the NS2 autoprotease is responsible only for processing at the NS2-NS3 junction, numerous mutagenic studies have demonstrated that when NS2 autoprotease activity is abolished, virus replication cannot proceed and infectious virus is not produced (Jones et al., 2007, Kolykhalov et al., 2000, Welbourn et al., 2005). As replication can occur in the absence of NS2 with SGRs comprising NS3-5B, the block on replication can be attributed to either an inhibitory effect of NS2 or a lack of mature NS3. While an inactive NS2 autoprotease appeared to have no effect on NS3 protease activity in one *in vitro* study (Tedbury and Harris, 2007), an alternative *in vitro* study observed a 4-5 fold decrease in NS3 mediated catalysis when unprocessed from NS2, despite similar binding kinetics for the NS3 protease substrate peptide (Welbourn et al., 2005). This effect was proposed to be due to a reduced ability of NS3 to bind to NS4A whilst still in complex with NS2. These results suggest that efficient NS3-mediated processing requires maturation from the NS2-NS3 precursor in order for NS3 to bind NS4A, in line with earlier observations that NS3-NS4A interactions require the NS3 N-terminus (Neddermann et al., 1999). However, analysis of polyprotein processing in SGRs with an inactive NS2 autoprotease revealed the NS3 protease was still active to some degree. Therefore, inactivation of the NS2 autoprotease so as to prevent maturation of NS3 does not appear to directly affect the NS3 protease active site, but may significantly impair its activity,

most likely by decreasing NS3 affinity for the NS4A cofactor (Welbourn et al., 2005). Such impairment might be of greater relevance *in vivo* than in cell culture systems.

Additionally, an inhibitory effect of the unprocessed NS2-NS3 precursor has been proposed. As previously stated, unprocessed NS2-NS3 is rapidly degraded through a proteasome dependent mechanism (Section 1.5.4). When mutant SGRs were analysed for protein levels at a range of timepoints, wildtype SGR containing NS2-5B showed a gradual accumulation of NS3. The same SGR with defective NS5B RdRp (GND mutant) did not show an increase in NS3 as no replication occurs, but NS3 translated from the initial input RNA was stable for 24-48 hours. By contrast, the same SGR with a defective NS2 autoprotease (H952A mutant), which also could not replicate, displayed a loss of the NS2-NS3 precursor after 4-8 hours, an effect that could be reversed by pharmacological inhibition of the proteasome (Welbourn et al., 2005). An inactive NS2 autoprotease preventing maturation of NS3 may lead to rapid, proteasome-mediated degradation of the NS2-NS3 precursor. This could have important implications for the ability of the virus to accumulate sufficient NS proteins to allow replication to occur and may have an additional effect on NS3-mediated processing of host factors that create a more favourable environment for the HCV lifecycle, such as processing of components of innate immune signalling pathways, as discussed in Section 1.5.5. In line with this, a full HCV transcript devoid of NS2 autoprotease activity (containing H952A and C993A mutants) failed to establish infection when introduced into chimpanzees through intrahepatic injection. By comparison, wildtype HCV transcripts with a functional NS2 autoprotease produced elevated ALTs, HCV-specific seroconversion and viremia, demonstrating prolonged infection (Kolykhalov et al., 2000).

Based on these observations, several groups have postulated the NS2 autoprotease as an attractive anti-viral target (Cheng et al., 2011, Lorenz, 2010, Lorenz et al., 2006, Rice, 2011, Schregel et al., 2009, Welbourn et al., 2005, Welbourn and Pause, 2006). The NS2 autoprotease represents a stand-alone protease with an essential function in the HCV lifecycle and no homology to any human proteases. Initial concerns regarding the limited window of opportunity in which to drug an autoprotease with rapid processing kinetics have been dampened by the growing evidence that NS2 may require an accumulation of polyprotein, as evidenced by the suggestion that NS2 may have to dimerise to form composite active sites (Lorenz, 2010).

The development of cell-based high-throughput screens targeting the NS2 autoprotease to attempt to identify NS2-specific anti-virals has been reported. One assay involved co-

expression of a reporter and substrate plasmid in Cos-1 cells. The reporter plasmid encodes chloramphenicol acetyltransferase downstream of a human T-cell leukemia virus type-1 (HTLV-1) long terminal repeat promoter. The substrate plasmid contained NS2 and its cleavage sequence followed by the Tax1 protein of HTLV. NS2-mediated cleavage liberated Tax1 so as to promote reporter expression (Hirowatari et al., 1995). An alternative assay selected stable cell lines in Jurkat cells expressing a β-lactamase reporter, followed by NS2-NS3 and a destabilising ubiquitin-like domain so that NS2-mediated processing stabilised the reporter (Whitney et al., 2002). However, neither of these groups has subsequently reported the identification of specific inhibitors targeting the NS2 protease activity.

A panel of protease inhibitors have already been tested against NS2 protease activity *in vitro* and were inactive (Table 1.2), consistent with the lack of homology of NS2 to these protease families. Some covalent inhibitors such as thiol alkylating agents blocked NS2-mediated proteolysis. An epoxide based cysteine protease inhibitor E64 was inactive, whilst a covalent serine protease inhibitor TLCK did block function. These findings could be explained by the active site observed in the post-cleaved NS2 crystal structure, which reveals a cysteine protease catalytic triad with serine protease geometry. Whether covalent protease inhibitors could prove useful in the development of an NS2 autoprotease inhibitor will be examined in Chapter 4.

In vitro studies have also analysed the effect of substrate peptides, with inconclusive results. Different groups have reported no activity or relatively weak activity of peptide substrates against the NS2 autoprotease. Whether peptide-mimetics represent a viable route to an inhibitor, similar to the success observed with NS3 protease inhibitors, remains to be seen, though early indications are that this will prove more challenging due to the nature of the protease and the assays required.

Some inhibitors of NS2-mediated functions have been reported based on host factor interactions. *In vitro* translation studies with rabbit reticulate lysates identified the requirement of ATP for efficient processing. This was linked to the requirement of ATP hydrolysis for the chaperone heat-shock protein 90 (Hsp90). Immunoprecipiation of Hsp90 in this *in vitro* system demonstrated a direct interaction with the NS2-3 precursor, which could be abrogated by Hsp90 inhibitors geldanamycin, herbimycin A and radicicol. Furthermore, these inhibitors blocked NS2 autoprotease activity *in vitro* and were more potent when added during translation, suggesting a role in ensuring correct folding of the NS2 autoprotease (Waxman et al., 2001). Cell-based assays where the stability of a reporter is mediated by the NS2 protease

activity have also observed an effect of Hsp90 inhibitors (Waxman et al., 2001, Whitney et al., 2002). The relevance of this effect in the context of HCV replication is examined in Chapter 3.

Additionally, correct folding of the NS2 autoprotease has been linked to CypA. As described in Section 1.6, CypA antagonists potently block HCV replication in cell culture. While the mode of action is somewhat ambiguous it is thought to be mediated through preventing CypA isomerisation of proline residues in NS5A domain 2 and LCS-II. However, infectious JFH1 virus has been noted as significantly more sensitive to inhibition by CsA (and analogues thereof) than JFH1 derived SGRs containing NS3-5B (Ciesek et al., 2009, Kaul et al., 2009). This shift in sensitivity was narrowed down to the NS2 gene, as a JFH1 SGR comprising NS2-5B was 5-10-fold more sensitive to CsA inhibition than the same SGR comprising NS3-5B (Ciesek et al., 2009). This effect was ablated by genetically separating NS2 and NS3 so as to bypass the requirement for NS2 autoprotease activity. A role for CypA in the correct folding of NS2, which is essential when the NS2 autoprotease is required for replication, was postulated. In Chapter 3 data will be presented exploring the role of CypA on the folding and function of the NS2 autoprotease and the role of NS2 on virus sensitivity to CsA.

1.9 Rational drug design and virtual high throughput screening

In the absence of a suitable starting point for the identification of an inhibitor of the NS2 autoprotease such as a classic cysteine protease inhibitor or substrate peptide derivative, the development of a NS2-acting DAA requires identification of a novel, active molecule. Due to the technical difficulties in developing assays suitable for measuring levels of NS2-mediated proteolysis, a HTS programme for inhibitor identification would be challenging. However, *in vitro* assays for NS2 autoproteolysis may be amenable to a lower throughput form of screening. This can be achieved through the use of computer aided drug design (CADD). CADD is increasingly used in the drug discovery process to reduce the cost and workload of identifying "hit" molecules (Ou-Yang et al., 2012). As with HTS, the aim is to identify a novel lead molecule, usually with modest activity in the μM range, for progression through structure activity relationship (SAR) analysis into a lead compound. CADD has been documented to increase the hit rate of screening by filtering out molecules unlikely to show activity and focussing screening libraries on molecules predicted to have increased likelihood of activity. However, rather than producing a larger number of hit compounds from a screening programme, CADD is usually applied to allow identification of a similar number of active

molecules through a much more manageable screening library (Sliwoski et al., 2014). CADD requires existing information about the system in one of two forms; ligand-guided CADD requires a known active compound and searches for chemically and conformationally similar molecules. Structure-guided CADD requires structural information of the target protein but is applicable in the absence of a known active compound.

Structure-guided CADD can be applied to structural information of the target in one of two ways. HTS libraries can be filtered to yield a more desirable subset of molecules using *in silico* virtual high-throughput screening (vHTS) with the aim of ranking large screening libraries based on modelling the most energetically favourable binding pose of each molecule to the target structure. Alternatively, *de novo* rational drug design can be employed to produce entirely novel chemical scaffolds designed to produce favourable interactions with the target structure.

1.9.1 *De novo* design

De novo design usually takes into account favourable properties of a "hit" compound, broadly defined as "lead-like space" (Nadin et al., 2012). These properties (e.g. LogP -1 to 3 and Mw 200-350) are usually accounted for to some degree within HTS libraries and hence are present within the virtual libraries used for vHTS. Despite this, *de novo* design allows for almost unlimited possibilities and hence molecules produced most likely require synthesis before they can be tested. Even when accounting for synthetic attainability during design, synthesising a reasonable sized library of *de novo* designed compounds for biological testing is relatively costly and time-consuming.

The *de novo* design software SPROUT has previously been reported for identification of novel active ligands (Heikkila et al., 2006, Mok et al., 2013). SPROUT is comprised of a number of modules capable of first defining a binding site by molecular recognition properties (e.g. defining hydrogen bond donors and acceptors and hydrophobic regions) before orientating rigid templates so as to best match these molecular properties. The rigid templates, anchored to their respective target sites, are then joined with a 'skeleton' to create a complete molecule. Numerous output solutions are produced as each rigid template may be joined through several skeleton scaffolds (Gillet et al., 1994). This method allows definition of certain key sites within the target receptor and allows these sites to be preferentially targeted manually.

To overcome the fact that molecules designed *de novo* in SPROUT are unlikely to be commercially available, features deemed key to the modelled interaction can be retained in compounds from vHTS libraries through the use of ligand-guided CADD using programs such as ROCS (Rapid Overlay of Chemical Structures) to rank chemical and structural similarity (Rush et al., 2005). However, results of ligand-guided CADD using a *de novo* designed molecule as the input, for which neither activity nor the binding pose are experimentally defined, should be treated with caution and hence may require screening of a library of predicted hits to identify true active compounds.

1.9.2 Virtual high-throughput screening

Algorithms for vHTS and *de novo* drug design respectively both rely on 'docking' as a form of scoring the interactions of a candidate ligand with the target structure. The process of docking can be divided into two requirements; the production of a feasible binding pose, and the correct scoring of both the best pose for each ligand, and this ligand pose relative to all other ligand poses. Producing a binding pose requires exploring conformational flexibility of the ligand. Ligand flexibility can be explored through systematic algorithms whereby rigid fragments of the ligand can be docked individually before being linked to score an overall pose, a method implemented within the vHTS software eHITS (Zsoldos et al., 2006). Ligand flexibility can also be explored through pre-generation of a defined number of conformers for each ligand prior to docking using OMEGA (Hawkins et al., 2010). Alternatively, a random approach can be taken to sampling the space available through repetitive cycles of random rotation or translation one bond at a time before scoring, as with Autodock (Totrov and Abagyan, 1997) and Glide (Friesner et al., 2004).

While producing the correct binding pose of a ligand is important for docking, it is not sufficient if the correct binding pose is not preferentially scored above incorrect poses, nor is it of use for vHTS if poses cannot be reliably scored between different ligands. Scoring of ligand poses has a number of limitations that provide the main drawback to CADD. Most commonly, a molecular mechanics (MM) force-field is employed to calculate the difference between the energy of the receptor-ligand complex and the ligand alone so as to account for unfavourable conformational changes within the ligand upon binding. Such a force-field, based around a Monte Carlo randomised algorithm, is employed in the scoring of Autodock poses for example. These energies are calculated from relatively simple interactions such as electrostatic, hydrophobic and van der Waals interactions. Importantly these scoring systems are based solely around enthalpic contributions. The entropic effects of displacing water molecules, both

from the ligand and the protein target, are not accounted for. This can be a major drawback as some ligand-protein interactions can be dominated by this entropy effect (Kitchen et al., 2004). Other scoring systems, such as those employed within eHITS, use knowledge-based and empirical scoring that fit experimental data and can be tailored to the system. For example, while the scoring system of eHITS still scores chemical properties such as hydrogen bonds, it also accounts for variability in the atomic coordinates of the crystal structure with respect to these interactions, and can also account for entropic effects to some degree (Zsoldos et al., 2006).

Difficulties in accurately scoring docking poses means that some molecules are scored incorrectly and hence the ranking of vHTS libraries does not directly translate into the observed degree of activity. Despite this numerous studies have demonstrated the ability of vHTS to significantly enrich screening libraries with "hit" compounds.

One example comes from a comparative study of HTS and vHTS for screening against angiogenin, an enzyme responsible for inducing angiogenesis. A HTS campaign involving 18,111 compounds produced 12 hits once verified and after removal of false positives, a hit rate of 0.066%. Of the 12 hit compounds, 5 were ranked in the top 2% from vHTS using structures from the same screening library (Jenkins et al., 2003). In real terms, HTS yielded one genuine hit compound for every 1,509 compounds screened. In contrast, screening of a vHTS-enriched library yielded five hits from 362 compounds, a hit rate of 1.38%. These hit rates are fairly typical for both HTS and vHTS. The use of CADD has been instrumental in the discovery of small molecule inhibitors in a range of systems (Schneider and Bohm, 2002) including HIV protease inhibitors and non-peptide based inhibitors of cysteine proteases as anti-infectives (Agarwal and Fishwick, 2010).

1.10 Project Aims

The NS2 autoprotease is a stand-alone protease essential for the HCV lifecycle with no homology to human proteases and as such represents an attractive drug target. To the best of our knowledge there are currently no small molecule inhibitors of the NS2 autoprotease in pre-clinical development. Despite concerted efforts, identification of NS2 inhibitors has been hampered by contradicting results and weak activities of classical protease inhibitors and substrate peptides, suggesting these approaches may not be successful in developing a DAA targeting NS2. Furthermore, screening for inhibitors of the NS2 autoprotease has been limited by the technical difficulties of conducting HTS in challenging *in vitro* assays and the lack of a cell-based HTS relevant to the HCV lifecycle. This study therefore set out to identify novel small molecule inhibitors directly acting against the NS2 autoprotease and assess the viability of inhibiting the NS2 autoprotease to exert an anti-viral effect.

To directly assess inhibitory activity against the NS2 autoprotease a method of quantifying NS2-mediated proteolysis *in vitro* is required. Such an assay must have sufficient sensitivity and reproducibility to accurately determine the potency of an inhibitor by determining EC_{50} following treatment at a range of concentrations, thus allowing comparison of the inhibitory activity from a series of compounds.

As *in vitro* assays quantifying NS2 autoprotease activity are low-throughput, the application of *in silico* CADD based on the post-cleaved NS2 crystal structure as a model for the NS2 autoprotease should improve hit rates and allow fewer compounds to be screened to identify an NS2 autoprotease inhibitor. Current data suggests the NS2 crystal structure to be relevant to the enzyme activity of the NS2-NS3 precursor, particularly with regard to the catalytic triad. Therefore CADD will be applied to a model of the NS2 autoprotease based on the post-cleaved NS2^{pro} crystal structure to enrich a relatively small screening library.

The majority of HTS programmes against HCV replication have used SGRs containing NS3-5B and as such would not identify potential inhibitors of NS2 autoprotease-dependent replication. To assess the ability of a small molecule inhibitor of the NS2 autoprotease to exert an anti-viral effect, compounds will be tested against a SGR comprising NS2-5B. These SGRs require optimisation and validation for use in a cell-based screening assay. Such an assay would benefit from a measure of selectivity towards NS2-dependent genome replication. SGRs containing NS2-5B will also be utilised to explore the potency of potential inhibitors against NS2-dependent genome replication as well as to investigate mode-of-action.

Finally, to confirm anti-viral effects of an NS2 autoprotease inhibitor, compounds will be tested for inhibitory activity against the complete HCV lifecycle using the HCVcc system. These systems are only available for particular HCV genotypes and their application will therefore be reliant on inhibitory activity against the NS2 autoprotease of genotype 2a (JFH1).

Chapter 2 - Materials and Methods

2.1 General Materials

2.1.1 Bacterial strains

Eschericia Coli (E. Coli) DH5α; Genotype F – φ80*lac*ZΔM15, *rec*A1, *end*A1, *hsd*R17 (r_k^-,m_k^+), *pho*A, *sup*E44, *thi*-1, *gyr*A96, *rel*A1, (Δ*lac*ZYA-argF), U169, λ- used for cloning were purchased from Life Technologies.

Chemically competent bacteria were produced using the Z-competent kit (Zymo Research).

2.1.2 Mammalian cell lines

Huh7 cells are a human hepatocellular carcinoma cell line (Nakabayashi et al., 1982) previously shown to support autonomous replication of HCV sub-genomic replicons (SGRs) (Lohmann et al., 1999).

Huh7.5 cells are a sub-population of Huh7s produced by clearing SGR from Huh7 cells with IFNα which support enhanced HCV replication (Blight et al., 2002) and were a kind gift from Charles Rice, The Rockefeller University, New York.

SGR-feo-Con1(BM4-5) [NS3-5B] and SGR-feo-JFH1 [NS3-5B] stable cell lines were produced as a polyclonal population in Huh7.5 cells by Yutaka Amako at the University of Leeds using constructs provided by David Wyles, University of California, San Diego.

2.1.3 Virus sequences

HCV sequences were obtained from the EU HCV database with the following Genbank accession numbers;

JFH1 – AB047639	HC-J6 – D00944	ED43 - Y11604
Con1 – AJ238799	H77 - AF011752	SA13 - AF064490
J4L6S – AF054247	NZL1 - D17763	EUHK2 - Y12083

FK5.1 is a culture adapted strain derived from Con1 with CAMs in NS3 (E176G and I254T) and CAMs in NS5A (N137D and S225P) (Krieger et al., 2001). Con1(BM4-5) is an alternative culture adapted strain derived from Con1 with a single deletion in NS5A (ΔS230) (Guo et al., 2001). The NS2 crystal structure (PDB 2HD0) is from HCV-J, a genotype 1b partial patient isolate which possesses greatest sequence identity to accession number FJ931730.

2.1.4 SGR and virus constructs

SGR-luc-JFH1 [NS3-5B] was used as reported (Targett-Adams and McLauchlan, 2005) as was a derivative thereof; mSGR-luc-JFH1 [NS3-5B] (Hughes et al., 2009). The JFH1 NS2 gene was previously introduced into SGR-neo-JFH1 [NS3-5B] (Kato et al., 2003) and SGR-luc-JFH1 [NS3-5B] (Targett-Adams and McLauchlan, 2005) to produce SGR-neo-JFH1 [NS2-5B] and SGR-luc-JFH1 [NS2-5B] (Tedbury et al., 2011). SGR-neo-FK5.1 [NS3-5B] and SGR-luc-FK5.1 [NS3-5B] are culture adapted SGRs derived from the Con1 strain containing CAMS in NS3 (E176G andT254I), and in NS5A (N137D and S225P) (Krieger et al., 2001) into which the Con1 NS2 gene has been introduced to yield SGR-neo-FK5.1 [NS2-5B] and SGR-luc-FK5.1 [NS2-5B] (Tedbury et al., 2011). SGR-feo-JFH1 [NS3-5B] and SGR-feo-Con1(BM4-5) [NS3-5B] were used as reported (Wyles et al., 2007). Con1 (BM4-5) is a culture adapted Con1 SGR with NS5A ΔS230. A replicon with NS5A-encoded CsA resistance mSGR-luc-JFH1 [NS3-5B] (D316E) has been previously reported (Ross-Thriepland et al., 2013).

Jc1-NLuc was produced at the University of Leeds by Yutaka Amako and is based on the Jc1 construct previously reported (Pietschmann et al., 2006).

2.1.5 Expression constructs

pNS2-3 derived from J4L6S (henceforth referred to as J4) and JFH1 isolates in the pET23a vector were used as reported (Foster et al., 2010b, Tedbury and Harris, 2007). pGST-CypA in pGEX59 has been previously reported (Foster et al., 2011).

2.1.6 Antibodies

Anti-NS5A serum (Sheep) (in house) was used at 1:5000.

Anti-FLAG M2 (Mouse) (F1803 – Sigma Aldrich) was used at 1:5000.

Anti-JFH1 NS2 serum (Rabbit) (Volker Lohmann, University of Heidelberg, (Jirasko et al., 2008)) was used at 1:500.

Anti-GAPDH (Mouse) (9484 – mAbcam) was used at 1:20,000.

Anti-cyclophilin A (Rabbit) (2175S – New England Biolabs) was used at 1:1000.

Anti-NS3 serum (Sheep) (in house) was used at 1:3000

Anti-neomycin phosphotransferase II (Rabbit) (06-747 -UpState) was used at 1:1000.

2.1.7 Chemicals

Imidazole, G418 disulphate salt, Thiazolyl Blue Tetrazolium Bromide (MTT formazan, 98%), 1,10-phenanthroline and cyclosporine A (CsA) were purchased from Sigma-Aldrich. Ethyldiamine tetraacetic acid (EDTA) was from VWR. L-cysteine was from AnalaR. 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate (CHAPS) was purchased from Glycon. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Generon. Guanidine hydrochloride (98%) was from Alfa Aesar. Dimethyl sulphoxide (DMSO, > 99.7%) was from Fisher. Ampicillin was from Melford Laboratories. Tosyl phenylalalnine chloromethane ketone (TPCK) was purchased from AppliChem. HCV DAA Telaprevir was purchased from MedChemExpress. HCV DAA Daclatasvir was from SelleckChem. Screening compound libraries were from ChemBridge 'Screening Compounds' library and purchased as dry powder.

2.1.8 Purification resins

HisTrap™ FF 1ml columns, Chelating Sepharose™ Fast Flow and Glutathione Sepharose™ 4B were purchased from GE Healthcare.

2.2 General Methods

2.2.1 Nucleic acid manipulation

2.2.1.1 Oligonucleotide primers

DNA oligonucleotides were purchased from Integrated DNA Technologies, resuspended in deionised water (dH$_2$O) to 100 µM and stored at -20 °C.

2.2.1.2 Site directed mutagenesis

QuikChange Mutagenesis (Agilent Technologies) was used for introduction of single amino acid mutations. Complementary primers 30-42 nucleotides in length housing required nucleotide changes centrally were used to perform a PCR with initial denaturation at 95 °C for 30 seconds then 5 cycles of; denaturation at 95 °C for 30 seconds; annealing at 42-52 °C for 1 minute; extension at 68 °C for 25 minutes. Following this a further 13 cycles were performed as before with an elevated annealing temperature of 55 °C. Parental DNA was then removed by DpnI digestion (20 U, 37 °C, 1 hour) before PCR products were transformed into competent DH5α bacteria as described in Section 2.2.1.9.

2.2.1.3 Agarose gel electrophoresis

DNA agarose gels contained 0.8% (w/v) agarose in 1 x TAE buffer (40 mM Tris, 0.11% (v/v) acetic acid, 1 mM EDTA) and 1:10,000 SYBR® Safe DNA Gel Stain (Invitrogen). Gels were run in 1 x TAE buffer at 8 V/cm. DNA samples were loaded in gel loading buffer (0.025% (w/v) Orange G, 5% (v/v) glycerol in 1 x TAE) and compared to Hyperladder I markers (Bioline). DNA gels were typically run at 100 V for 50 minutes.

2.2.1.4 DNA extraction from agarose gels

DNA bands were visualised with blue light to avoid UV irradiation. Bands were excised with minimal agarose and DNA purified using Zymoclean^{TM} Gel DNA Recovery Kit (Zymo Research) following the manufacturer's instructions.

2.2.1.5 Restriction Digestion

All restriction enzymes were from New England Biolabs (NEB). Reactions were carried out in 20 µl volume at 37 °C using recommended and relevant buffers and enzyme units. All reactions were performed for a minimum of 1 hour.

2.2.1.6 Phenol:Chloroform extraction

Purification of DNA was achieved by addition of equal volume phenol:chloroform:isoamyl alcohol (25:25:1), briefly vortexed and centrifuged 13,000 rpm 5 minutes. The upper aqueous layer was removed and equal volume chloroform added, vortexed and centrifuged 13,000 rpm 5 minutes. Again the aqueous layer containing DNA was removed and 2 x volume ethanol and 0.1 x volume 3 M sodium acetate were added before incubation at -20 °C and recovery of precipitated DNA by centrifugation at 13,000 rpm for 20 minutes. Recovered DNA pellet was washed in 70% ethanol and centrifuged 13,000 rpm 10 minutes before resuspension in an appropriate volume of dH_2O.

2.2.1.7 DNA ligation

Ligations between digested vector and DNA inserts were performed at 1:3 and 1:6 molar ratios using 1 unit of T4 DNA ligase (Invitrogen) at either 27 °C, 18 °C or 4 °C for 1-16 hours. Following ligation one tenth of the ligation reaction was transformed into competent DH5α bacteria alongside a vector only ligation control transformation.

2.2.1.8 Transformation of competent bacteria

Chemically competent bacteria were prepared by the manufacturer's instructions using Z-Competent™ E. Coli Transformation Buffer Set (Zymo Research) and stored at -80 °C. Typically, 10-50 ng of DNA was added to 50 µl competent bacteria on ice before spreading over Luria Bertani (LB)-agar plates supplemented with 100 µg/ml ampicillin. Transformations of DNA constructs containing HCV NS5A sequences were grown at 30 °C. Transformations of all other cultures were grown at 37 °C.

2.2.1.9 Purification of plasmid DNA

Plasmid DNA was produced from a single colony of transformed competent DH5α bacteria grown in a suitable volume of LB media with 100 µg/ml ampicillin for roughly 16 hours at 37 °C. Plasmid DNA containing HCV NS5A sequences was cultured at 30 °C. Plasmid DNA was purified on a small scale via alkaline lysis using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) or a larger scale using Plasmid Midi Kit (Qiagen) using the manufacturers protocol.

2.2.1.10 Oligonucelotide quantification

Concentration and purity of oligonucleotides was assessed by measuring absorbance at 280 nm and 260 nm using a NanoDrop 1000 Spectrophotometer (Thermo Scientific). 260/280 nm ratio was used to assess oligonucleotide purity.

2.2.1.11 DNA sequencing

All DNA constructs were confirmed by Sanger sequencing with appropriate primers using Beckman Coulter Genomics Sequencing service.

2.2.1.12 RNA manipulations

For handling of RNA all surfaces and equipment involved was first treated with RNase AWAY (Molecular BioProducts). Solutions used for RNA manipulation such as dH_2O or Phosphate Buffered Saline (PBS) were treated with 0.1% (v/v) diethyl pyrocarbonate (DEPC) and incubated shaking at 37 °C for 16 hours. DEPC was removed by autoclaving.

2.2.1.13 Preparation of DNA template

For *in vitro* transcription of RNA, DNA template plasmids were linearised in a 50 μl volume with appropriate restriction enzymes. Sub-genomic replicon (SGR) or virus constructs derived from the JFH1 isolate were linearised with XbaI at 37 °C for 1-16 hours. SGRs derived from Con1 or FK5.1 isolates were linearised with high fidelity ScaI for 1-4 hours at 37 °C. Restriction enzymes were inactivated at the suggested temperature for 20 minutes, cooled at room temperature for 10 minutes then on ice for a further 10 minutes. To remove overhang regions of the linearised DNA, reaction volume was increased to 100 μl with dH_2O and treated with 10 U Mung Bean Nuclease (NEB) at 30 °C for 45 minutes before inactivation by addition of 0.1% (w/v) SDS. Linearised DNA was purified by phenol/chloroform extraction as described in Section 2.2.1.7 and resuspended in DEPC treated dH_2O.

2.2.1.14 In vitro transcription of RNA

In vitro transcription was performed using 1 μg of linearised DNA template and the T7 RiboMax Express system kit (Promega). Transcription proceeded at 30 °C for 90-120 minutes before treatment with DNase at 37 °C for 15-30 minutes as per manufacturer's instructions. Following DNase treatment RNA was purified by phenol/ chloroform extraction as in Section 2.2.1.7 with the exception of the use of phenol:chloroform at pH 5.2. Purified RNA in the aqueous phase was precipitated by addition of 1 x volume isopropanol and 0.1 x volume 3 M sodium acetate pH 5.2 on ice for 5 minutes before centrifugation at 13,000 rpm for 10

minutes. RNA pellets were washed in 70% ethanol and resuspended in 18-20 µl DEPC-dH$_2$O. RNA was stored at -80 °C.

2.2.1.15 RNA agarose gel electrophoresis

RNA transcripts were analysed on a denaturing 3-(N-morpholino) propanesulphonic acid (MOPS) formaldehyde gel. 1% (w/v) agarose was added to 30 ml MOPS buffer (40 mM MOPS, 10 mM sodium actetate, 1 mM EDTA) and microwaved at 340 W until dissolved. Once cool, 6.5% (v/v) formaldehyde and 1:10^4 SYBR® Safe DNA Gel Stain were added. RNA (typically 1 µg) was boiled at 65 °C for 10 minutes in RNA loading buffer (47.5% (v/v) formamide, 9 mM EDTA, 0.0125% (w/v) SDS, 0.0125% (w/v) Xylene Cyanol, 0.0125% (w/v) Bromophenol Blue) before loading alongside 1 µg ssRNA Ladder (NEB). RNA gels were resolved at 80 V for 1 hour.

2.2.1.16 Reverse transcription polymerase chain reaction

Reverse transcription polymerase chain reaction (RT-PCR) was used to produce cDNA from extracted cellular RNA. Reaction mixtures comprised 0.5 mM deoxynucleotide triphosphates (dNTPs), 200 ng random hexamer primers (Thermo Scientific) or 1 µM gene-specific primers and 1 µg of RNA in 13 µl dH$_2$O and were incubated at 65 °C for 5 minutes and a further 5 minutes on ice. 10 mM DTT was then added and the reaction volume increased to 19 µl in First Strand Buffer (Invitrogen). Reactions were incubated at 25 °C for 2 minutes before addition of 200 U Superscript® II RNase H Reverse Transcriptase (Invitrogen). Reactions were then incubated at 25 °C for 10 minutes, 42 °C for 50 minutes and 70 °C for 15 minutes. cDNA was stored at -20 °C.

2.2.1.17 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was used to amplify cDNA from RNA reverse transcriptions using a 'touchdown' PCR cycle. PCR reaction mixtures comprised 0.3 mM deoxynucleotide triphosphates (dNTPs), 400 nM of each primer, 2 µl of cDNA reaction mixture (or 100 ng control vector DNA) and 1 U Vent® DNA polymerase in a reaction volume of 50 µl ThermoPol® Reaction Buffer (NEB). Reactions were denatured for 1 minute at 94 °C followed by 20 cycles of a 30 second denaturation step at 94 °C, 30 seconds at a variable annealing temperature starting at 65 °C and dropping 0.5 °C each cycle (65-55 °C) and a 1 minute extension step at 72 °C. Following this a further 15 cycles were performed at an annealing temperature of 55 °C. A final extension step at 72 °C for 5 minutes was performed.

2.2.2 Protein Biochemistry

2.2.2.1 SDS-PAGE analysis

Unless otherwise stated SDS-PAGE analysis was carried out using reducing Tris-Glycine 15% polyacrylamide gels (15% (v/v) acrylamide, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate (APS), 0.01% (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED), 376 mM Tris-HCl, pH 8.8) with a 6% stacking gel (6% (v/v) acrylamide, 0.1% (w/v) SDS, 0.1% (w/v) APS, 0.01% (v/v) TEMED, 376 mM Tris-HCl, pH 6.8). Gels were run in 1 x SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS), typically at 200 V for 1 hour. Protein samples were loaded in ¼ volume 4 x Laemmli buffer (8% (w/v) SDS, 30% (v/v) glycerol, 20 mM dithiothreitol (DTT), 0.02% (w/v) Bromophenol Blue, 200 mM Tris-HCl, pH 6.8) and boiled for 5 minutes at 95 °C before loading. Protein size was compared to ColorPlus Prestained Protein Marker, Broad Range (7-175 kDa) (NEB).

2.2.2.2 Coomassie Blue staining

To visualise all proteins present gels were stained with Coomassie Brilliant Blue R-250 stain (50% (v/v) methanol, 10% (v/v) acetic acid, 0.05% (w/v) Brilliant Blue R-250) for over 1 hour. Gels were de-stained in 45% (v/v) methanol, 10% (v/v) acetic acid.

2.2.2.3 Western blot analysis

Resolved gels were transferred to polyvinylidene fluoride (PVDF) Immobilon®-FL Transfer Membrane (Immobilon) using a TE77X semi-dry transfer unit (AA Hoefer) at 15 V, 0.5 A for 1 hour. Transfer membrane was activated in methanol and gels and transfer were soaked in Transfer buffer (20% (v/v) methanol, 25 mM Tris, 192 mM glycine). Transferred membranes were blocked in 50% (v/v) Odyssey® Blocking Buffer (LI-COR) in TBS. Primary and secondary antibodies were incubated for at least 1 hour at room temperature or overnight at 4 °C at the appropriate dilution in 25% (v/v) Odyssey® Blocking Buffer in TBS. Between steps unbound antibodies were washed off by 3-5 x 10 ml washes in TBS supplemented with 0.1% (v/v) Tween®-20. Following removal of all unbound antibodies, membranes were washed briefly in dH$_2$O then allowed to dry. Membranes were imaged using an Odyssey imager (LI-COR).

2.2.2.4 Acetone precipitation

To precipitate protein from denaturing conditions, 2 x volume cold acetone was added, briefly vortexed and incubated -20 °C for at least 1 hour. Samples were centrifuged at 13,000 rpm 30 minutes, washed in 70% (v/v) ethanol and resuspended in the initial volume of 1x Laemmli buffer.

2.2.2.5 Protein concentration

Protein samples were concentrated using Vivaspin 20 ml concentrators (Sartorius stedim biotech) with a 10,000 molecular weight cut-off. Concentrators were centrifuged at 4,000 rpm at 4 °C until reduced to the required volume.

2.2.2.6 Buffer dialysis

Protein solution buffer was altered by dialysis in 2 L of required buffer at 4 °C overnight using Snakeskin Dialysis Tubing (Thermo Scientific) with a 10 kDa cut-off.

2.2.2.7 Protein quantification

Protein concentration was quantified by comparison to a bovine serum albumin (BSA) standard curve using the Pierce®BCA Protein Assay Kit (Thermo Scientific). Protein samples were diluted in dH_2O appropriately to 50 µl volume and mixed with 50 µl complete BCA solution in a 96 well plate before incubating at 37 °C for 30-45 minutes. Absorbance was measured at 570 nm using an infinite f50 platereader (Tecan).

Alternatively, for quantification of protein under denaturing conditions, 5 µl protein samples (and BSA standards in the appropriate denaturant) were added to a 96 well plate before addition of 250 µl Bradford reagent (5% (v/v) methanol, 0.5% (w/v) Coomassie Blue G250, 8.5% (v/v) phosphoric acid (H_3PO_4)). Plates were incubated whilst shaking for 5 minutes at 27 °C before measuring absorbance at 570 nm.

2.2.3 Tissue culture techniques

2.2.3.1 Microscope

Cultured cells were visualised using an OPTIKA XDS-2 microscope.

2.2.3.2 Passaging of mammalian cells

Mammalian cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 100 IU penicillin ml^{-1}, 100 µg streptomycin ml^{-1} and 1% (v/v) non-essential amino acids. Cells were cultured in a humidified incubator at 37 °C with 5% CO_2. Cells were typically passaged in a T175 flask and before reaching confluency media was removed and cells were washed in PBS. The cell monolayer was then incubated for 5 minutes at 37 °C in 0.5 mg/ml trypsin to detach cells before inactivating trypsin through the addition of excess complete media. Cells were typically subcultured 1:10 into fresh flasks.

2.2.3.3 Freezing of cell stocks

A minimum of 1×10^6 cells were frozen in cell storage media (10% (v/v) DMSO, 30% (v/v) FBS in full media). Cell stocks were stored at -80 °C.

2.2.3.4 RNA electroporation of mammalian cells

Cells were counted by 1:1 dilution into HyClone® Trypan Blue Solution before live cells were counted. Cells were pelleted at 1500 x g for 5 minutes then washed twice in 10 ml ice cold DEPC-treated PBS. Cells were resuspended to a final concentration of 1×10^7 cells/ml in ice cold RNase free PBS. For SGR electroporations 2 μg RNA was added to 4×10^6 cells in a chilled 4 mm electroporation cuvette (Geneflow). For infectious virus construct electroporations 5 μg RNA was used. Electroporation was performed at 950 μF and 270 V with a BioRad Gene Pulser. Cells were gently resuspended in full media using a Pasteur pipette before seeding at the required density.

2.3 *In vitro* methods

2.3.1 *In vitro* NS2-3 autoprocessing assay

2.3.1.1 Bacterial strains

NS2-3 overexpression was performed in z-competent BL21 DE3 pLysS *E. Coli*.

2.3.1.2 Plasmid constructs

The NS2-3 coding sequence of JFH1 NS2-3 (JFH1 residues 906-1209) and J4 NS2-3 (J4L6S residues 904-1206) along with a 3' Flag-tag were previously cloned into pET23a (pNS2-3) by Phillip Tedbury (Tedbury et al., 2007). All constructs were confirmed by Sanger sequencing. Plasmid containing bacteria were grown under ampicillin selection.

2.3.1.3 NS2-3 overexpression

A 50 ml culture of BL21 pLysS transformed with pNS2-3 was grown overnight at 37 °C and used to inoculate 500 ml LB media supplemented with 1% (w/v) glucose and 50 μg/ml ampicillin. Cultures were grown whilst shaking at 37 °C to OD_{600} of 1.0. Cultures were cooled at 4 °C for 30 minutes before addition of 1 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG) to induce overexpression through the lac operon. Induction was performed for 3 hours at 37 °C before harvesting bacteria by centrifugation at 6000 x g for 10 minutes at 4 °C. Bacterial pellets were stored at -20 °C.

2.3.1.4 NS2-3 purification of inclusion bodies

Bacterial pellets were re-suspended in 5 ml/g Lysis Buffer (1% (v/v) Triton X-100, 20 mM MgSO$_4$, 5 mM DTT, 100 mM Tris-HCl, pH 8.0) supplemented with 1 µg/ml lysozyme and 20 µg/ml DNase I before incubating for 20 minutes on ice. Membranous components of lysed bacteria were pelleted by centrifugation at 22,000 x g for 1 hour at 4 °C. Pellets were washed twice in 10 ml/g ice cold Wash Buffer 1 (2% (v/v) Triton X-100, 2 M Urea, 5 mM EDTA, 5 mM DTT, 100 mM Tris-HCl, pH 8.0) by re-suspension and centrifugation at 22,000 x g for 30 minutes at 4 °C. Pellets were washed a final time in 5 ml/g ice cold Wash Buffer 2 (5 mM EDTA, 5 mM DTT, 100 mM Tris-HCl, pH 8.0) by re-suspension and centrifugation at 22,000 x g for 1 hour at 4 °C. Pelleted inclusion bodies were solubilised in ice cold Extraction Buffer (6 M guanidine hydrochloride (GdnHCl), 500 mM NaCl, 50 mM Tris-HCl, pH 8.0) and centrifuged at 40,000 x g for 1 hour at 4 °C. Resulting supernatant was filtered through a 0.22 µm filter before progressing to His-tag purification. Extracted inclusion bodies were stored at 4 °C for no longer than 16 hours.

2.3.1.5 NS2-3 His-tag purification

Filtered inclusion bodies in Extraction Buffer were loaded onto either 5 ml Chelating Sepharose^{TM} Fast Flow resin charged with Nickel following the manufacturer's instructions or onto a HisTrap^{TM} FF 1 ml column.

Resins were loaded by rotating loose resin with extracted inclusion bodies for 1 hour at 4 °C. Resin was then pelleted at 4,000 x g for 5 minutes at 4 °C before flowthrough supernatant was removed. Washes and elutions were performed by batch in the same manner.

Columns were charged using a Bio-Rad Econo Gradient Pump at a flow rate of 1 ml/min at 27 °C. Washes and elutions were performed in the same manner.

Following loading resins or columns were washed with 1-5 ml fractions of Extraction Buffer containing increasing amounts of imidazole. Imidazole gradients ran from 5 mM imidazole to 500 mM imidazole. Fractions were acetone precipitated, re-suspended in the same volume of 1x Laemmli loading buffer and analysed by Coomassie stain of 15% SDS-PAGE and anti-FLAG western blot. Fractions with sufficient protein quantity and purity were pooled and concentrated if necessary to 20 µM. NS2-3 was stored in 20 µl aliquots at -80 °C.

2.3.1.6 NS2-3 refolding reaction

NS2-3 at 20 µM in Extraction Buffer was diluted 1:100 into 100 µl Refolding Buffer #6. Refolding Buffer was optimised as in section 3.2.1 from the following solutions;

Refolding Buffer #1: 30% (v/v) glycerol, 250 mM NaCl, 3 mM L-cysteine, 0.5% (w/v) CHAPS, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 7.0

Refolding Buffer #2: 30% (v/v) glycerol, 250 mM NaCl, 3 mM glycine, 0.5% (w/v) CHAPS, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 7.0

Refolding Buffer #3: 30% (v/v) glycerol, 250 mM NaCl, 3 mM L-cysteine, 0.5% (v/v) Triton-X-100, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 7.0

Refolding Buffer #4: 30% (v/v) glycerol, 250 mM NaCl, 3 mM L-cysteine, 0.5% (w/v) CHAPS, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 8.0

Refolding Buffer #5: 30% (v/v) glycerol, 250 mM NaCl, 3 mM L-cysteine, 0.5% (w/v) CHAPS, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 7.0, 1 mM DTT

Refolding Buffer #6: 30% (v/v) glycerol, 250 mM NaCl, 3 mM L-cysteine, 0.5% (w/v) CHAPS, 50 μM ZnCl$_2$, 50 mM HEPES-NaOH pH 7.0, 10 mM DTT

0 hour (0h) samples were stopped immediately by the addition of 20 μl 5 x Laemmli loading buffer. Reactions were incubated at 27 °C for 16 hours unless otherwise stated before halting the reaction with Laemmli loading buffer. Samples were boiled for 5 minutes at 95 °C and analysed by 15% SDS-PAGE and anti-FLAG western blot.

2.3.1.7 Compound treatment of NS2-3 refolding reaction

Dose response curves were made as DMSO stock solutions at 133.3 x final concentration. 0.75 μl compound stock in DMSO was added to the 100 μl Refolding Buffer and mixed prior to addition of NS2-3. Refolding reactions were then treated as in Section 2.3.1.6.

2.3.1.8 Western blot analysis and quantification of autoprotease activity

Once reactions were halted samples were boiled unless otherwise stated and 15 μl was analysed by 15% SDS-PAGE as described in Section 2.2.2.1. Western blot was performed as in Section 2.2.2.3 using 1:5000 anti-FLAG M2 primary antibody and 1:10^4 IRDye 680RD Donkey anti-Mouse secondary antibody. Bands corresponding to NS2-3-FLAG and NS3-FLAG were quantified and normalised to DMSO control values. Data was modelled to calculate EC$_{50}$ by fitting a nonlinear regression using the FindECanything function of Prism 6 (GraphPad). Use of antibodies conjugated to infrared fluorescent dyes with the Odyssey Infrared Imaging system (LI-COR Biosciences) produces an accurate, sensitive quantification of protein levels across a linear range (Schutz-Geshwender, 2004).

2.3.2 pGST-CypA overexpression and purification

2.3.2.1 GST-CypA overexpression

50 ml cultures of wildtype pGST-CypA and H126Q pGST-CypA in BL21 DE3 cells were used to inoculate a 500 ml LB supplemented with 50 µg/ml ampicillin. Cultures were grown whilst shaking at 37 °C to OD_{600} of 0.8. Cultures were cooled at 4 °C for 30 minutes before addition of 1 mM IPTG for a 4 hour induction at 20 °C. Bacteria were harvested by centrifugation at 6000 x g for 10 minutes at 4 °C and pellets stored at -20 °C.

2.3.2.2 GST-CypA purification

Bacterial pellets were re-suspended in 10 ml per 500 ml culture phosphate buffered saline (PBS) supplemented with 1% (v/v) Triton-X-100, 2 µg/ml Anrafinin (Sigma-Aldrich), 0.2 mM Pefabloc® (Roche), 1 µg/ml Leupeptin (Sigma-Alrich) and 1 µg/ml Pepstatin A (Sigma-Alrich). Suspensions were sonicated on ice with 6 x 20 second pulses at 10 microns separated by 20 seconds. Soluble material was isolated by centrifugation at 16,000 x g for 30 minutes at 4 °C.

Soluble material was passed through a 0.22 µm filter and loaded onto 5 ml Glutathione Sepharose™ 4B resin by rotation at 4 °C for 1 hour. Resin was pre-equilibrated by 3 x 5 minute rotations at 4 °C in PBS followed by 5 minute rotation at 4 °C in PBS + 1% (v/v) Triton-X-100. Loaded resin was washed 2 x by rotation in 10 ml PBS + 1% (v/v) Triton-X-100 for 5 minutes at 4 °C, then 2 x 10 ml PBS for 5 minutes at 4 °C followed by 2 x 10 ml 50 mM Tris-HCl pH 8.0 for 5 minutes at 4 °C. GST-CypA was eluted from the resin by 2 x rotations in 5 ml 50 mM Tris-HCl pH 8.0 supplemented with 40 mM reduced glutathione (Acros Organics) for 10 minutes at 4 °C. Purified GST-CypA was stored at 4 °C.

2.4 Mammalian tissue culture methods

2.4.1 Compound treatment of mammalian cells

2.4.1.1 Seeding cells

Cells were trypsinised and diluted 1:1 with Hyclone® Trypan Blue Solution to stain dead cells. Live cells were counted using a Marienfeld Haemocytometer. Unless otherwise stated cell solutions were diluted to 2 x 10^5 cells/ml. 100 µl of this cell suspension was seeded in each well of a 96-well plate to give 2 x 10^4 cells/well. Cells were left a minimum of 4 hours before media was changed or compound was added.

2.4.1.2 Preparation of DMSO standard curves

All compound dose-response curves contained standard final DMSO, typically at 0.25% (v/v). To ensure standard levels of DMSO, concentration ranges were produced by serial dilution in DMSO at 400 x final concentration. 400 x stock concentration ranges were diluted 400-fold in media to produce the final concentrations through several dilution steps; typically 1:20 into an intermediate plate and 1:20 into the final plate. All reported EC_{50} and CC_{50} curves in cell-based systems were performed in duplicate with mean and standard deviation plotted. DMSO only control is plotted on the y axis.

2.4.1.3 Luciferase endpoint

Media was removed and wells were washed 3 x with 100 μl PBS before addition of 30 μl 1 x Passive Lysis Buffer (PLB; Promega). Plates were stored frozen at -20 °C. All luciferase dose response curves were read in a white 96-well microplate using a BMG Labtech Fluostar plate reader injecting 40 μl of LAR-I reagent (Promega) and recording light emission over 6 seconds. Emission data from 1.6 – 6.0 seconds was averaged and data was normalised to DMSO controls. Data was modelled to calculate EC_{50} by fitting a non-linear regression using the FindECanything function of Prism 6 (GraphPad).

2.4.1.4 Selection of stable cell lines

10 μg of RNA of a SGR-feo construct was electroporated into Huh7.5 cells as described in Section 2.2.3.4. Cells were seeded in 10 ml in a 10 cm dish for 24 hours before media was changed and supplemented with 500 μg/ml G418. During selection cells were subcultured as required, with discarded cells analysed for luciferase signal. Upon the formation of G418 resistant colonies, either individual colonies were picked to yield monoclonal populations or selection was continued before a polyclonal population was harvested.

2.4.1.5 RNA extraction

Relevant SGR-feo cell lines were seeded at a density of 4 x 10^4 cells/ml in either a 12-well plate or a 10 cm dish. Cells were washed 2 x in 10 ml PBS and incubated for 10-15 minutes in 1 ml TRIzol® Reagent (Invitrogen) gently shaking at 27 °C. TRIzol extractions were stored as 500 μl aliquots at -80 °C. RNA was purified from extractions by RNA chloroform extraction as in Section 2.2.1.14.

2.4.1.6 Western blot analysis

Cells were seeded at a density of 80 x 10^4 cells/ml for up to 48 hours before washing in PBS and lysing in a relevant volume of PLB. PLB lysates were stored at -20 °C. Lysates underwent

centrifugation at 4000 rpm for 2 minutes and insoluble material was discarded. Lysates were quantified by BCA assay against BSA standards as in Section 2.2.2.7. Required amounts of protein were analysed by SDS-PAGE and western blot analysis as in Sections 2.2.2.1 and 2.2.2.3.

2.4.2 Mammalian cell toxicity analysis

2.4.2.1 MTT assay

Cells were seeded in a clear 96 well plate as described in Section 2.4.1.1 and treated with required concentrations of the compound of interest as in Section 2.4.1.2. At the required timepoint, typically 48 hours, media was removed and replaced with 150 µl high glucose, no glutamine, no phenol red DMEM (Gibco®) before addition of 30 µl 6 mM Thiazolyl Blue Tetrazolium Bromide. 6 mM Thiazolyl Blue Tetrazolium Bromide was prepared by dissolving Thiazolyl Blue Tetrazolium Bromide to 12 mM in PBS, filtering through a 0.22 µm filter and dilution to 6 mM with high glucose, no glutamine, no phenol red DMEM. Cells were incubated in 1 mM Thiazolyl Blue Tetrazolium Bromide for ≥ 2 hours before media was removed and Thiazolyl Blue Tetrazolium Bromide crystals were dissolved in 100 µl DMSO by shaking for 10 minutes at 27 °C. Plates were then read at 570 nm using an infiniteF50 platereader (Tecan). Data was normalised to DMSO controls and modelled using a non-linear regression to fit a CC_{50} using Prism 6 (GraphPad).

2.4.2.2 ATPLite assay

ATPLite Luminescence ATP Detection Assay System was performed following the manufacturer's instructions (Perkin Elmer) with luminescence quantified using a BMG Labtech Fluostar plate reader. Data was normalised to DMSO controls and modelled using a non-linear regression to fit a CC_{50} using Prism 6 (GraphPad).

2.5 Biosafety level 3 (BSL3) mammalian tissue culture methods

2.5.1 Compound treatment of mammalian cells under BSL3

Infectious virus construct transcripts were electroporated into Huh7.5 cells as described in Section 2.2.3.4 and seeded to the same cell density in 96 well plate format in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS), 100 IU penicillin ml^{-1}, 100 µg streptomycin ml^{-1} and 1% (v/v) non-essential amino acids and 20 mM

Hepes Buffer (Lonza). Cells were transferred to BSL3 containment within 4 hours post electroporation (h.p.e.). Cells were treated with compound as in Section 2.4.1.2 with the exception of a final volume not exceeding 100 µl and luciferase endpoint taken as in Section 2.4.1.3.

2.5.2 BSL3 MTT mammalian cell toxicity assay

MTT assay was performed as in Section 2.4.2.1. After incubation in 1 mM Thiazolyl Blue Tetrazolium Bromide for ≥ 2 hours media was removed and virus inactivated with 100 µl/well 4% paraformaldehyde (PFA) for ≥ 10 minutes. PFA solution was gently removed and Thiazolyl Blue Tetrazolium Bromide crystals were dissolved in 100 µl DMSO by shaking for 10 minutes at 27 °C. Plates were then read at 570 nm using an infiniteF50 platereader (Tecan) and analysed as in section 2.4.2.1.

2.5.3 Infection of naïve mammalian cells

Following 48 hour treatment with a range of compound concentrations in 96 well format, 100 µl virus containing supernatant was removed and added to 2 x 10^4 Huh7.5 cells in 96 well format. Following 16 hour incubation, cells were washed in PBS and luciferase endpoint taken as in Section 2.4.1.3.

2.6 Chemistry Materials

Reagents and anhydrous solvents were purchased from commercial sources and used without further purification. Magnesium Sulphate was from Fisher Ltd. 1-Hydroxybenzotriazole, 1-(3-dimethylaminopropyl)-3-ethy-carbodiimide hydrochloride and meta-chloroperbenzoic acid (mCPBA) were from Sigma Aldrich. 4-Spiroindene-piperidine hydrochloride (SM-2) was from KeyOrganics. 1-Methyl-1-cyclohexanecarboxylic acid (SM-1) and cyclohexane carboxylic acid were from Sigma Aldrich.

2.7 Chemistry Methods

Thin-layer chromatography (TLC) was performed using aluminium pre-coated silica gel plates (Machery-Nagek, 0.2 mm thickness) (Merck Chemicals) using the indicated solvents, with retention factor (Rf) reported. Column chromatography was carried out using geduran silica

gel 60 (40-63 µm) (Merck Chemicals) using the same solvent as used for TLC. Melting points were determined using a Reichert Hot Stage apparatus.

Proton (^{1}H) and carbon (^{13}C) NMR spectra were recorded on a 300 / 75 MHz Bruker DPX300 Fourier transform spectrometer with an internal deuterium lock. Chemical shifts are reported in parts per million (ppm) with reference to an internal tetramethylsilane (TMS) standard. For proton spectra (δ_H), the integration (e.g. 1H), multiplicities (e.g. s, singlet, d, doublet, t, triplet, m, multiplet etc.) and coupling constants (J, measured in Hertz) are reported. For carbon spectra (δ_C) the assigned carbon for each peak is indicated. NMR samples were prepared in deuterated chloroform (CDCl$_3$) from Sigma-Aldrich.

Mass spectrometry was performed in DMSO using a VG Autospec mass spectrometer by electron spray ionisation at 70 eV. LC-MS was carried out using an Agilent Technologies 1200 series HPLC with UV detection from 190-750 nm coupled to a Bruker Ultra HCT mass spectrometer.

Synthetic procedures are described In Appendix I.

Chapter 3 - Development and validation of screening assays for identification of NS2-targeting compounds

3.1 Introduction

The NS2 autoprotease has long been postulated as a potential therapeutic target for the treatment of HCV infection. As understanding of the nature of the NS2-encoded protease has grown, this essential enzyme activity has appeared a more feasible target. The protease activity is required in the virus lifecycle, acts at the stage of polyprotein processing and appears to further regulate the role of the NS3 protease in this stage of the lifecycle (Jones et al., 2007, Kolykhalov et al., 2000, Welbourn et al., 2005). Inhibitors of polyprotein processing by NS3-mediated proteolysis have already demonstrated a beneficial therapeutic effect in the clinic as evidenced by the use of NS3 protease inhibitors in new therapies. The catalytic domain of NS2 has a novel fold, unusual cysteine protease active site geometry and no homology to any human proteases indicating that it may be amenable to a selective small molecule inhibitor (Lorenz, 2010, Lorenz et al., 2006). While a *cis*-acting unimolecular autoprotease reaction may be expected to process as soon as correctly folded, and hence display a narrow time-window in which to block the proteolysis event, indications that the NS2 autoprotease may be regulated by the amounts of polyprotein, function in a concentration dependent manner and require multimerisation suggest this window may be extended (Lorenz, 2010).

The success of NS3 protease inhibitors and NS5B NIs was initially driven through *in vitro* assays amenable to HTS. Despite similar efforts the conditions required to reconstitute proteolysis and the nature of the NS2 protease have hampered the application of *in vitro* assays for the NS2 autoprotease into a HTS system. *In vitro* systems for studying the NS2 autoprotease include *in vitro* translation, which can include addition of cellular factors through microsomal membranes, or the reconstitution of enzyme activity *in vitro* using bacterially expressed and purified recombinant proteins (Pallaoro et al., 2001, Pieroni et al., 1997, Tedbury and Harris, 2007, Thibeault et al., 2001). Alternatively tagged NS2-based expression constructs can be transfected into cell lines or expressed through vaccinia systems and analysed by immunoprecipitation and western blot analysis (Gorzin et al., 2012, Schregel et al., 2009).

While there are advantages to screening compounds in a cell-based system with regards to progression as a drug candidate, as these systems account for compounds exerting cellular toxicity or poor membrane permeability, they contribute a greater degree of complexity to both the process and analysis of initial screening. In addition *in vitro* assays generally allow screening of compounds at higher concentrations. Furthermore, the complexity of a cell-based

system, and to some degree *in vitro* systems requiring addition of cellular components, makes it more difficult to answer the fundamental question of whether a small molecule can block the process of NS2-NS3 processing regardless of its viability as a drug candidate. Therefore, to explore whether the proteolytic activity of the NS2 autoprotease can be inhibited by a small molecule, a system involving purified recombinant protein and reconstitution of enzyme activity *in vitro* was chosen.

The first section of this chapter (3.2.1) describes the purification of recombinant proteins encoding a fragment of the NS2-NS3 precursor sufficient for NS2 autoprotease-mediated processing (NS2-3). The optimisation of refolding this protein so as to initiate NS2-mediated proteolysis with respect to its use in an assay amenable to screening compounds is described. The activity of some previously reported inhibitors of the NS2 protease reaction are explored with respect to their use as a positive control.

The aforementioned benefits of a cell-based assay (ruling out compounds exerting cellular toxicity, instability or poor membrane permeability) can be achieved with a SGR based assay. Compounds can be tested against a SGR comprising NS2-5B to identify those that inhibit NS2-dependent HCV genome replication. As described in Figure 1.7, SGRs that contain NS2-5B are already available and have been used to prove that the NS2 autoprotease activity is required for genome replication both for the genotype 2a isolate JFH1 and a genotype 1b culture adapted isolate termed FK5.1 (Tedbury et al., 2011). SGRs comprising NS2-5B currently have either a NPT reporter, which is not amenable to a rapid and sensitive measure of replication, or luciferase, which cannot be maintained as a stable cell population and hence requires transient electroporation of RNA transcripts prior to each assay. The process of electroporation of SGR transcripts into cells not only increases the complexity of a screening assay, particularly with regards to HTS, but introduces additional variability into the system. Screening against NS2-5B SGRs would therefore benefit from a fusion reporter (feo) allowing a stable cell line to be maintained and providing sensitive quantification of genome replication. In optimising SGRs containing NS2-5B for screening, maximisation of the replication kinetics and hence the luciferase signal should be considered to obtain the greatest signal-to-noise ratio.

Section 3.2.2 of this chapter describes the cloning, selection and validation of SGR-feo [NS2-5B] stable cell lines with respect to their use in a cell based screen against NS2-dependent HCV replication.

The use of a positive control in cell-based screening assays is more problematic as no specific small molecule inhibitors of the NS2 autoprotease have been described to date. However, indications that host factor chaperones contribute to correct folding of NS2 and help modulate protease activity could allow the use of chaperone antagonists. Both Hsp90 and CypA have been postulated to play a role in the correct folding of NS2, with antagonists shown to block NS2-NS3 processing in cell culture and have an additional effect in SGRs containing NS2-5B respectively (Ciesek et al., 2009, Waxman et al., 2001).

Section 3.2.3 of this chapter explores the potential of these chaperone antagonists to serve as a positive control against NS2-dependent HCV replication. This section also examines in greater detail the potential effect of CypA on the NS2 protein with implications for the anti-viral effect of CsA and its analogues, and explores the anti-viral activity of Hsp90 inhibitors in a system relevant to HCV replication.

3.2 Results

3.2.1 NS2-NS3 in vitro autoprocessing assay

The cloning, overexpression, purification and optimisation of refolding of the NS2-3 recombinant protein has been previously reported (Foster et al., 2010b, Tedbury and Harris, 2007). As depicted in Figure 3.1, these constructs encode the C-terminal 'catalytic' domain of NS2 and the N-terminal protease domain of NS3. For the genotype 2a JFH1 isolate (EU HCV database accession number AB047639) this corresponds to residues 906-1209 of the HCV polyprotein, while the genotype 1b J4 isolate (EU HCV database accession number AF054247) encodes residues 904-1206. Constructs are preceded at the N-terminus with four lysine residues and a hexa-histidine tag and flanked at the C-terminus by a FLAG epitope (DYKDDDDK) for purification and detection purposes respectively.

Within the pNS2-3 expression construct, NS2-3-FLAG derived from the JFH1 isolate has a molecular weight of 35,030.49 Da (35 kDa). Successful reconstitution of the NS2 autoprotease activity will initiate proteolysis at the NS2-NS3 junction, leading to processing of the P5 - P5' substrate sequence GWKLLAPITA by cleavage of the scissile bond between the P1 leucine of NS2 and the P1' alanine of NS3. The His-tagged NS2 product has a predicted molecular weight of 15,005.62 Da (15 kDa). The FLAG-tagged NS3 product has a molecular weight of 20,042.89 Da (20 kDa) (Expasy Bioinformatics Resource Portal).

Similarly, NS2-3-FLAG derived from the J4 isolate of pNS2-3 has a molecular weight of 34,911.31 Da (35 kDa) and is processed to a His-NS2 product 15,049.67 Da (15 kDa) in molecular weight and a NS3-FLAG product 19,879.66 Da (20 kDa) in molecular weight. The J4 isolate of NS2-3 has an alternative P3 residue, with the P5 – P5' cleavage sequence GWRLLAPITA.

A model of the unprocessed NS2-3 protein based on separate crystal structures of the post-cleaved NS2 protease domain ((Lorenz et al., 2006), PDB 2HD0) and the protease domain of NS3 in complex with NS4A is shown in Figure 3.1B. NS2-3 is purified in denaturant, meaning structural elements may not be present until dilution into refolding buffer.
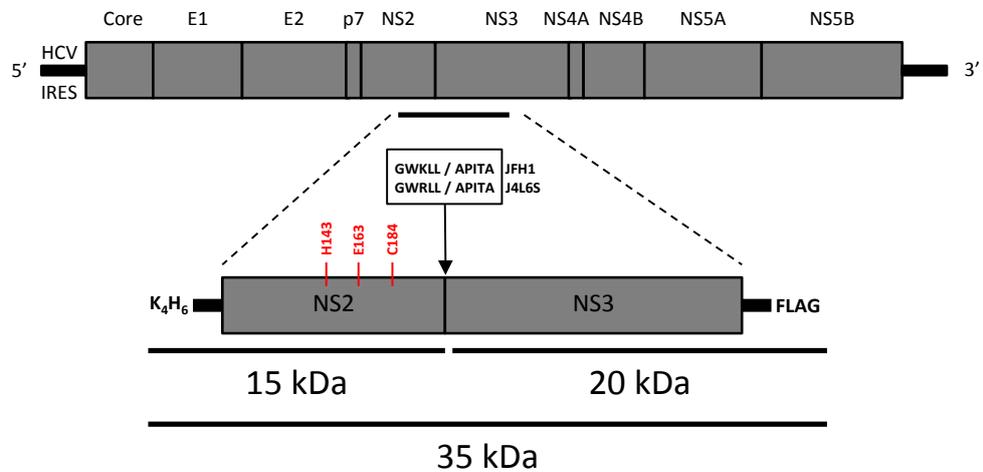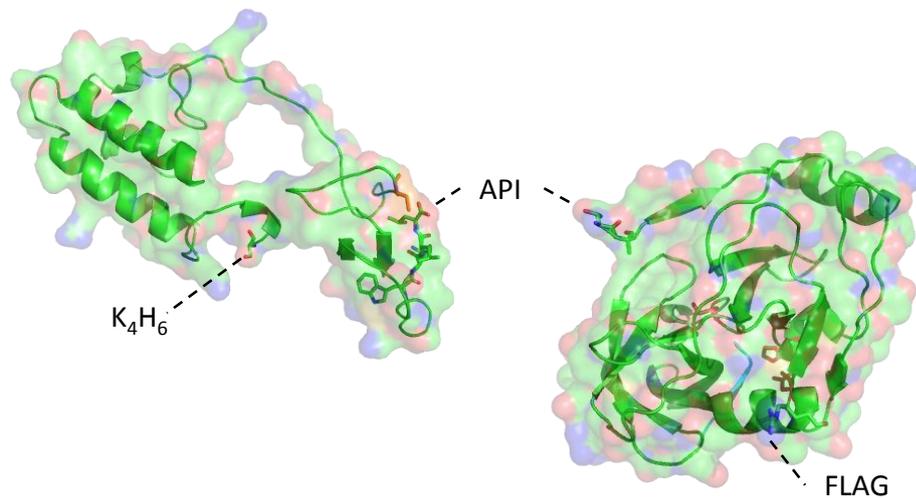
Figure 3.1: NS2-3 protein overexpression construct.

A) NS2-3 comprises the C-terminal catalytic domain of NS2 and the N-terminal protease domain of NS3. The N-terminus is flanked by four lysine residues and a hexa-histidine tag (K_4H_6) and the C-terminus by a FLAG epitope. Residues of the catalytic triad of the NS2 autoprotease are marked in red. The site of proteolysis is indicated by the arrow, with P5 - P5' residues for JFH1 and J4 isolates shown. B) Cartoon of NS2-3 based on crystal structures of NS2 (PDB 2HD0) and NS3 (PDB 3RC4). The catalytic cysteine of NS2 is shown in orange. Catalytic residues of the NS3 protease, highlighted in red, and the NS4B binding site, highlighted in blue, are on the reverse face of NS3. N-terminal and C-terminal tags and P5 – P5' residues are shown. P1', P2' and P3' residues at the N-terminus of NS3 are not present in the crystal structure. Surface representation is shown with oxygen atoms in red and nitrogen atoms in blue.

Overexpression and purification of NS2-3, as described in Section 2.3.1, followed a minor modification of previously reported methods in *E. coli* BL21 (DE3) pLysS cells. Following a 3 hour induction with IPTG at 37 °C cells were lysed and centrifuged at 22,000 x g to pellet inclusion bodies. Inclusion bodies were washed in a solution of 2 M Urea before solubilisation in 6 M guanidine hydrochloride (GdnHCl). Purified inclusion bodies contained predominantly un-processed 35 kDa NS2-3 along with the presence of an additional smaller species (Figure 3.2). In line with some previous observations this is most likely a NS2-3 truncation product (Foster et al., 2010b).

Purification of NS2-3 under denaturing conditions worked equally efficiently either by batch purification using a Ni^{2+} charged sepharose resin or commercially available pre-packed 1 ml Hi-Trap columns. Significant levels of NS2-3 were frequently observed in flowthrough fractions, but the high yields obtained from purification suggest the capacity of the column may have been reached. Washing with increasing concentrations of imidazole gradually eluted NS2-3, with the majority of protein eluted after 500 mM imidazole washes. Purified NS2-3 contained both the truncation product and higher molecular weight species. These may correspond to dimers or aggregates as a significant proportion can be removed by the addition of 10 mM DTT to increase levels of monomeric NS2-3 (Figure 3.2B).

NS2-3 purification fractions were selected based on optimal purity and concentrated or diluted as required to 20 µM (0.7 mg/ml). Refolding reactions were initiated by 1:100 dilution into detergent rich refolding buffer. Optimisation of the refolding reaction tested several different buffer compositions as outlined in Section 2.3.1.6. Buffers were tested using J4 NS2-3 following 16 hour refolding reactions by western blot analysis for the appearance of NS3-FLAG, one of the proteolysis products. Refolding reactions were compared to a reaction in Refolding Buffer #6 which was immediately halted (0h) by the addition of Laemmli SDS loading buffer (Figure 3.3).

Optimisation of the buffer confirmed previous observations that DTT was required, with greater cleavage efficiency observed with 10 mM DTT (Refolding Buffer #6) over 1 mM DTT (Refolding Buffer #5) (Figure 3.3). Successful refolding and autoproteolysis was also temperature optimised, with greatest autoprotease activity observed at room temperature (Figure 3.3B).

A
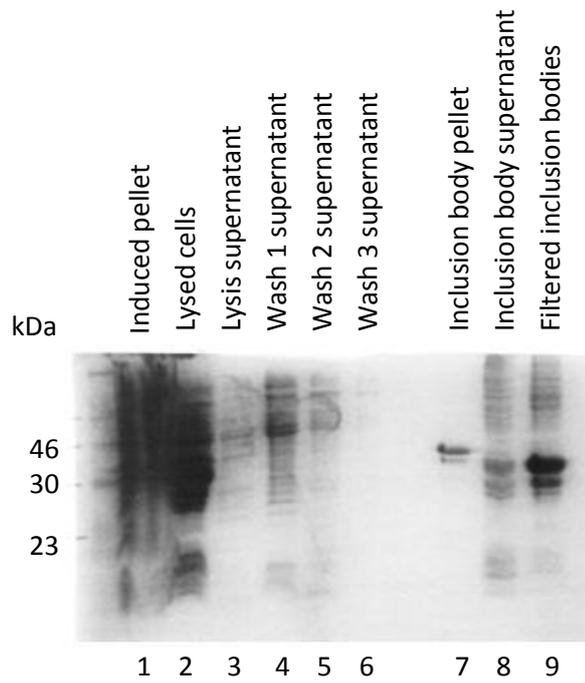
B
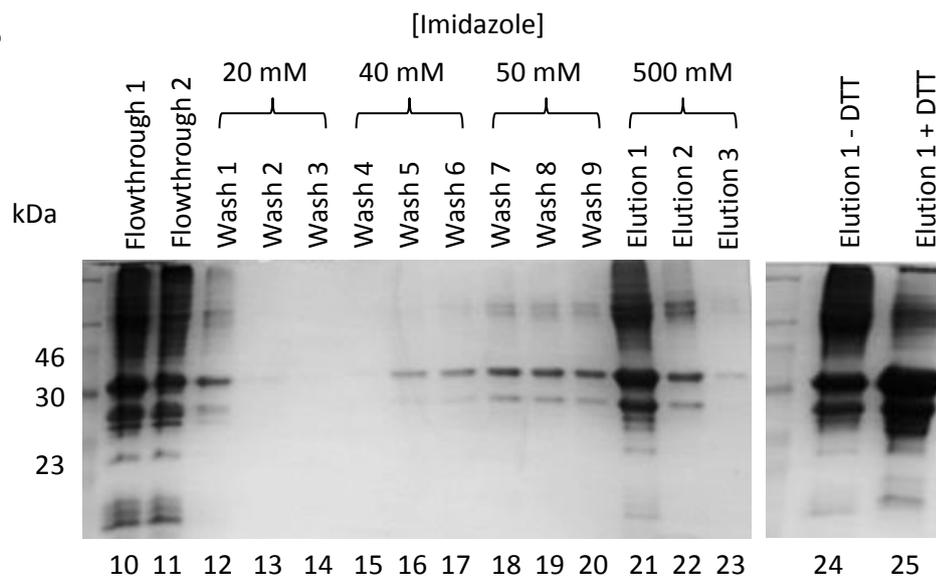
Figure 3.2: NS2-3 purification.

A) Key J4 NS2-3 overexpression and inclusion body isolation fractions analysed by Coomassie stain of SDS-PAGE. Induced cell pellets (lane 1) were lysed (lane 2) and inclusion bodies pelleted to remove soluble material (lane 3). Pelleted inclusion bodies were washed three times (lanes 4-6) before solubilising in 6 M GdnHCl and centrifugation to separate soluble inclusion bodies (lane 8) and remaining cellular debris (lane 7). Inclusion bodies were passed through a 0.45 micron filter (lane 9). B) Coomassie staining of SDS-PAGE analysis for NS2-3 purification fractions. Inclusion bodies were loaded onto Ni^{2+} charged resin or a Ni^{2+} affinity column (lanes 10-11). Bound NS2-3 was washed in GdnHCl containing increasing concentrations of imidazole (lanes 12-20) and eluted in 500 mM imidazole (lanes 21-23). Higher order species were observed in eluted protein fractions (lane 24) which could be removed or reduced to increase levels of monomeric NS2-3 by addition of 10 mM DTT (lane 25).
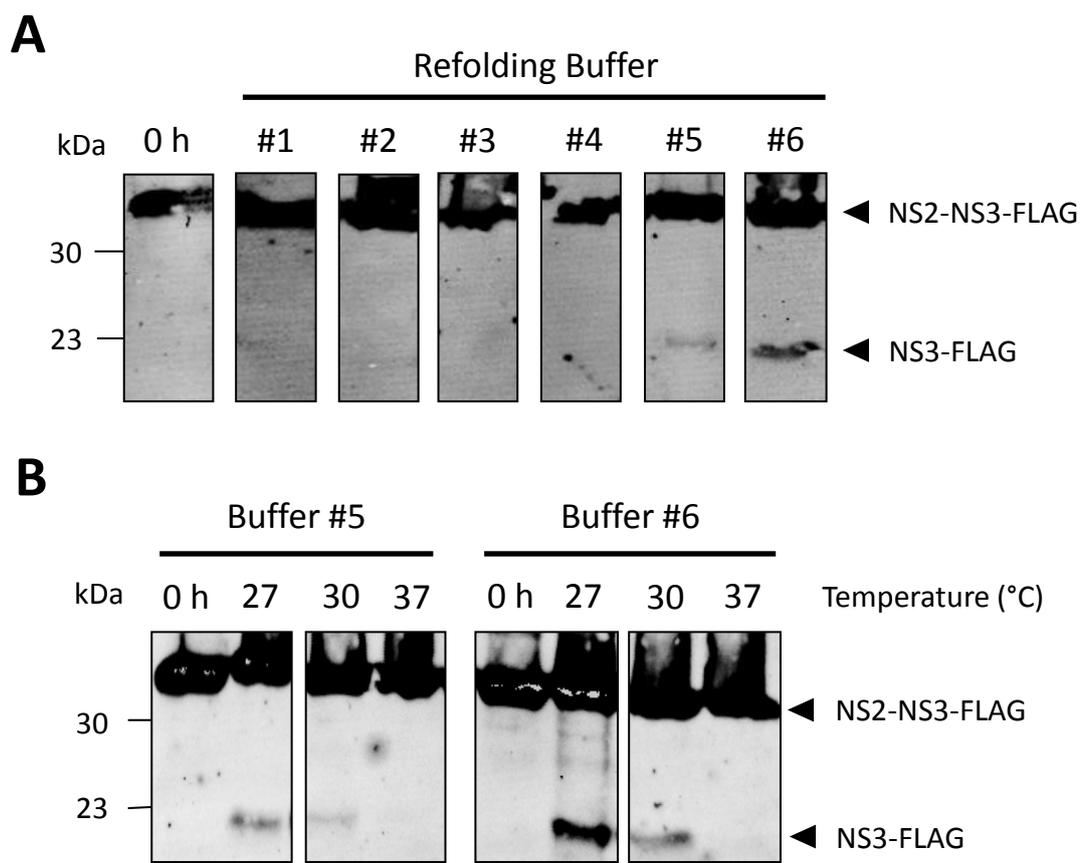
Figure 3.3: Optimisation of NS2-3 refolding reactions.

A) J4 NS2-3 refolding reactions were performed in a range of buffer compositions (Refolding Buffer #1-#6, outlined in Section 2.3.1.6). Reactions were resolved by SDS-PAGE and analysed by western blot using αFLAG antibody. The 20 kDa NS3-FLAG autoproteolysis product was only observed in Buffer #5 and Buffer #6. B) Reactions in Buffer #5 and Buffer #6 were temperature optimised and analysed by western blot, with both producing the greatest NS3-FLAG at 27 °C. Reduced levels of NS3-FLAG were observed at 30 °C and no significant production of NS3-FLAG over 0 hour (0h) control at 37 °C.

To confirm that the observation of NS3-FLAG in refolding reactions was due to the catalytic activity of the NS2 autoprotease, the NS2-3 tertiary structure was disrupted by the addition of zinc chelators. Previous observations have demonstrated that the catalytic activity of NS2 is dependent on zinc, most likely as a structural element coordinated by C922 in NS2 and residues 1123, 1125 and 1171 in NS3 (Figure 1.8A). In line with this, these studies demonstrated that *in vitro* NS2 autoprotease activity could be abrogated by the addition of zinc chelators (Table 1.2).

These observations were confirmed for the NS2-3 refolding reaction by addition of 10 mM EDTA or 10 mM 1,10-phenanthroline. Both zinc chelators completely abolished the production of detectable NS3-FLAG cleavage product after 16 hour refolding reactions, while untreated reactions produced NS3-FLAG (Figure 3.4). To demonstrate that this activity was through selective chelation of zinc atoms the treatment of refolding reactions with 10 mM EGTA, which selectively chelates calcium, did not have the same effect (Figure 3.4B). That the production of NS3-FLAG in refolding reactions can be inhibited by zinc chelators demonstrates that correctly folded NS2-3 is required and that the presence of NS3-FLAG represents catalytic activity of the NS2-3 pre-cursor.

To explore the kinetics of NS2-3 autoproteolysis in the refolding reaction a timecourse was performed. NS2-3 refolding reactions were stopped by the addition of Laemmli SDS loading buffer every 30 minutes for 4 hours and additionally at 8, 16, 24 and 48 hours. Levels of NS3-FLAG were quantified and compared to a 0 hour sample (Figure 3.5). NS3-FLAG significantly above background was observed after 2.5 hours. Levels of NS3-FLAG did not increase after 3.5 hours, suggesting the reaction was complete by this timepoint. Total protein appeared reduced after 48 hours incubation, with decreased levels of NS3-FLAG and NS2-NS3-FLAG suggesting protein may not be stable over this length of time.

Levels of NS3-FLAG up to 24 hours were plotted and a non-linear regression was fitted. These results demonstrate that the refolding reaction is complete by 3.5 hours (Figure 3.5B). As a consequence some further optimisation reactions were performed over a 4 hour refolding reaction. For consistency, all compound treatments were performed over a 16 hour timepoint.
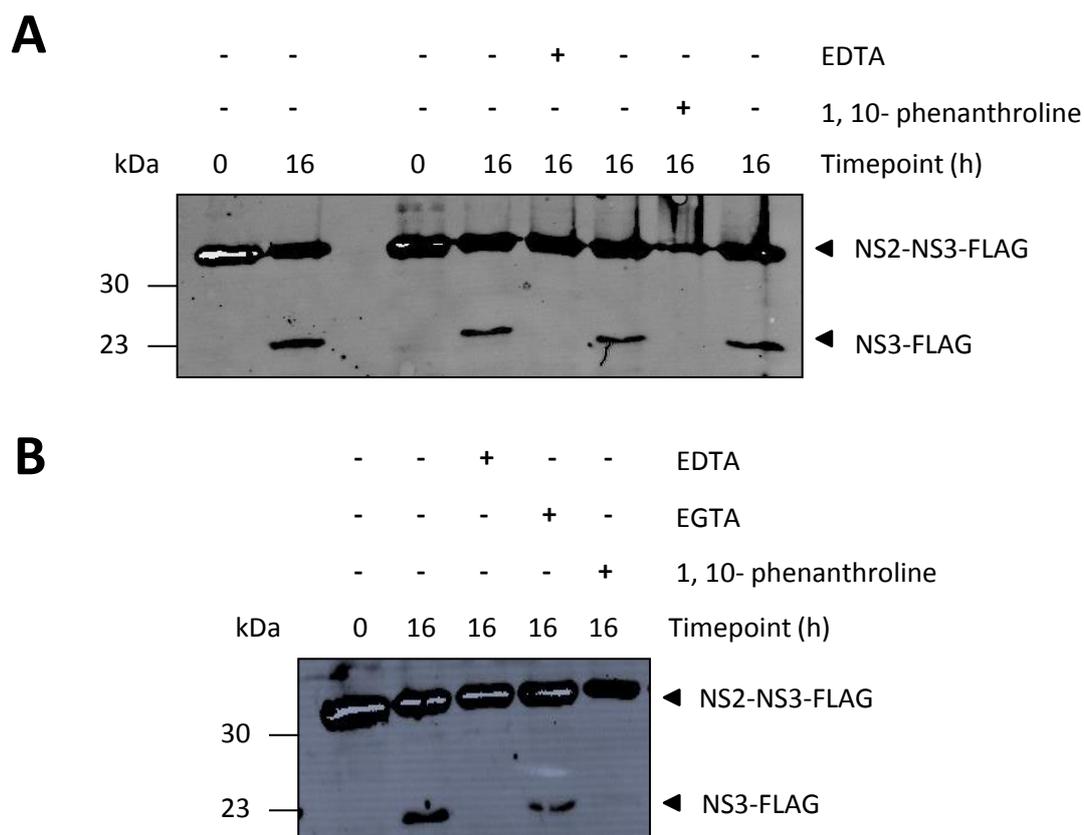
Figure 3.4: Abrogation of autoprotease activity by zinc chelation.

Refolding reactions were allowed to proceed for either 0 hours or 16 hours in the absence (-) or presence (+) of EDTA (10 mM), 1,10-phenanthroline (10 mM) or EGTA (10 mM). Reactions were resolved by SDS-PAGE and analysed by western blot. A) Production of NS3-FLAG from J4 NS2-3 refolding reactions was prevented by addition of EDTA or 1,10-phenanthroline. B) NS3-FLAG autoproteolysis product was produced in the presence of EGTA.

A

Time (hours)

0 0.5 1 1.5 2 2.5 3 3.5 4 8 16 24 48 0

◄ NS2-NS3-FLAG

◄ NS3-FLAG

B

in vitro NS2-3 autoprocessing timecourse
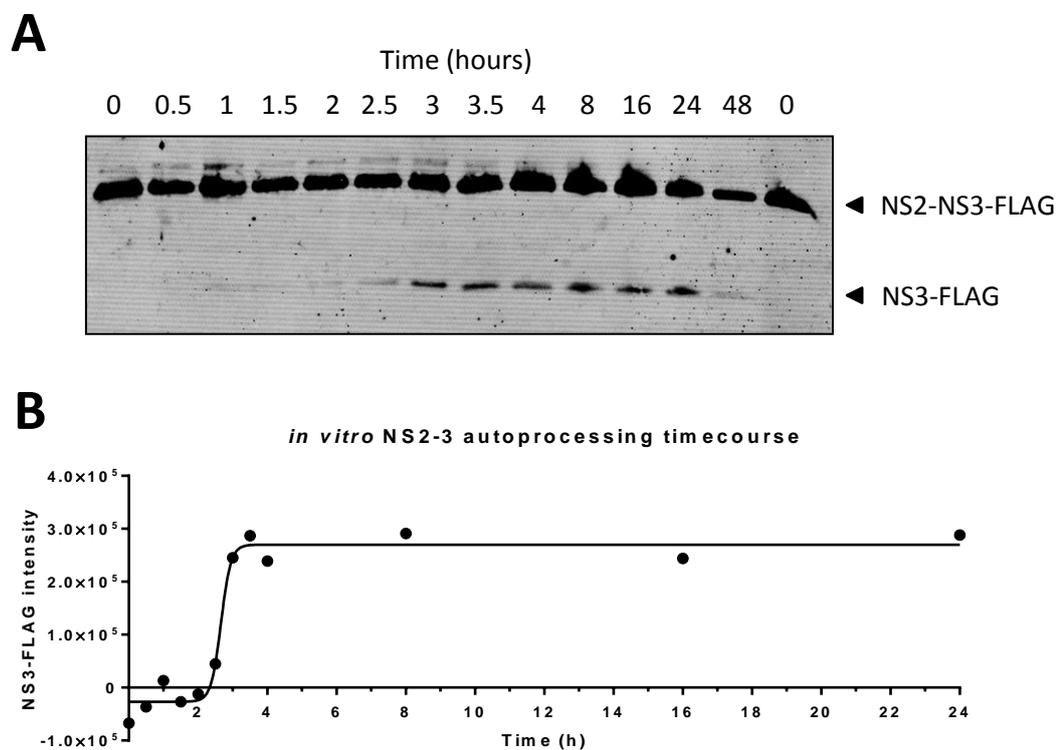
Figure 3.5: NS2-3 refolding and autoproteolysis timecourse.

A) JFH1 NS2-3 refolding reactions were allowed to proceed for indicated times before reactions were analysed by western blot. B) Levels of NS3-FLAG autoproteolysis product were quantified and plotted to fit non-linear regression. Autoproteolysis reactions were complete by 3.5 hours.

In order to test compounds against the NS2-3 refolding reaction, the assay was optimised for the effects of DMSO as a vector. A 16 hour refolding reaction was treated with concentrations of DMSO ranging from 0.25% to 50% (v/v) and compared to a 0 hour sample. NS3-FLAG was observed to similar levels when reactions were treated with 0.25% and 0.5% DMSO. NS3-FLAG was observed from 1% to 10% DMSO but appeared reduced (Figure 3.6A). To confirm these observations NS3-FLAG and NS2-NS3-FLAG were quantified (Figure 3.6B). DMSO levels below 1% produced similar NS3-FLAG intensities as an untreated, 0% DMSO, reaction. At 1% DMSO and above the levels of NS3-FLAG produced in the refolding reaction were reduced and by 4% DMSO were not above background. Additionally, higher quantities of DMSO reduced detection of total protein, with intensity of NS2-NS3-FLAG also reduced. The refolding reaction can therefore tolerate a final DMSO concentration < 1% (v/v). All subsequent *in vitro* NS2-3 refolding assays contain less than 1% final DMSO unless otherwise stated.

Previous reports have linked NS2-3 derived from the genotype 1b J4 isolate with decreased homogeneity after purification due to the presence of a contaminant, most likely a truncation product, along with reduced levels of reconstituted autoprotease activity (Foster et al., 2009). Throughout the current study both J4 and JFH1 isolates of NS2-3 were associated with some contaminants when eluted with high imidazole after fewer washes, most notable of which was small amounts of the NS3-FLAG cleavage product used to monitor the refolding reaction. Optimisation of the purification protocol allowed for selection of the purest fractions by gradually increasing imidazole levels (Figure 3.2B), a process required for both isolates of the NS2-3 protein. By this method fractions were selected whereby only unprocessed NS2-NS3-FLAG could be observed by Coomassie stain and anti-FLAG western blot analysis.

Despite comparable purity, direct comparison of 16 hour refolding reactions for both JFH1 NS2-3 and J4 NS2-3 confirmed that the JFH1 isolate was more efficient for reconstituting autoprotease activity *in vitro*. Refolded JFH1 NS2-3 produced significantly more NS3-FLAG than the J4 NS2-3 refolding reaction (Figure 3.7). Quantification of NS2-NS3-FLAG and NS3-FLAG from these reactions allowed calculation of the efficiency of the reconstituted autoproteolysis. The JFH1 NS2-3 autoprotease reaction was 37.6% efficient, in that 37.6% of the total protein quantified was present as NS3-FLAG. By contrast, the J4 NS2-3 refolding reaction was only 6.4% efficient (Figure 3.7). Previous timecourse analysis suggests this efficiency is unlikely to improve with extended time (Figure 3.5).
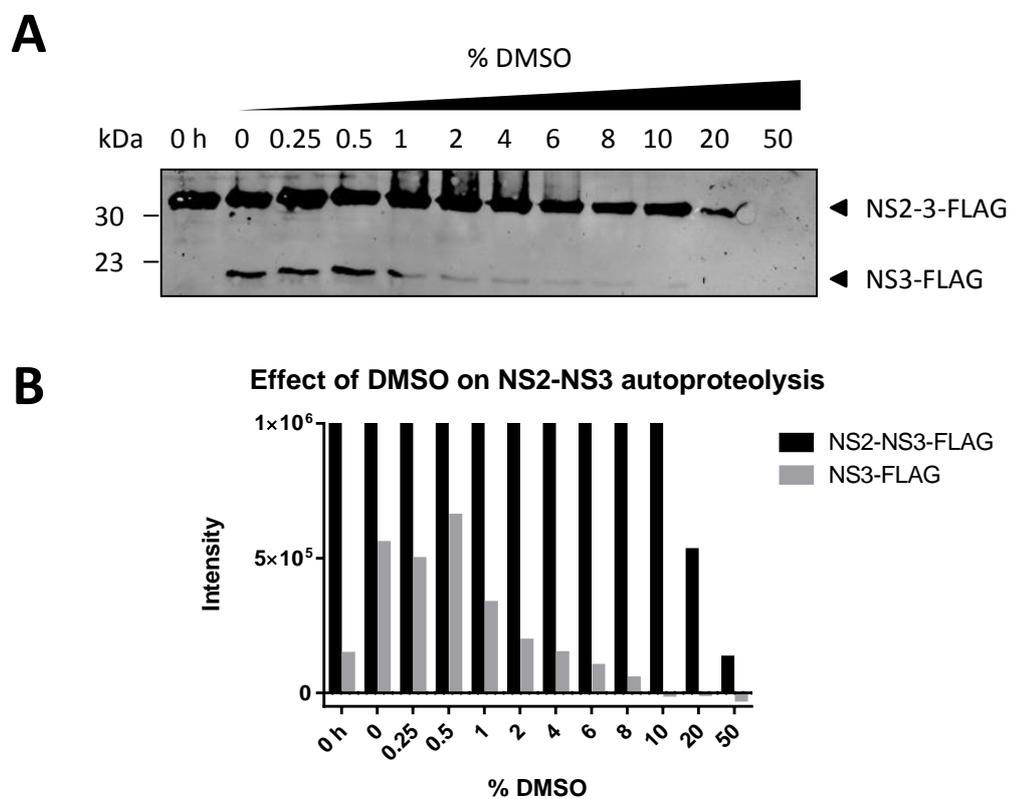
Figure 3.6: *In vitro* **NS2-3 autoproteolysis assay DMSO sensitivity.**

A) 16 hour NS2-3 refolding reactions were treated with indicated final DMSO concentration before analysis by western blot. NS3-FLAG autoproteolysis product was compared to a 0 hour control. B) Quantification of NS2-NS3-FLAG and NS3-FLAG bands from A.
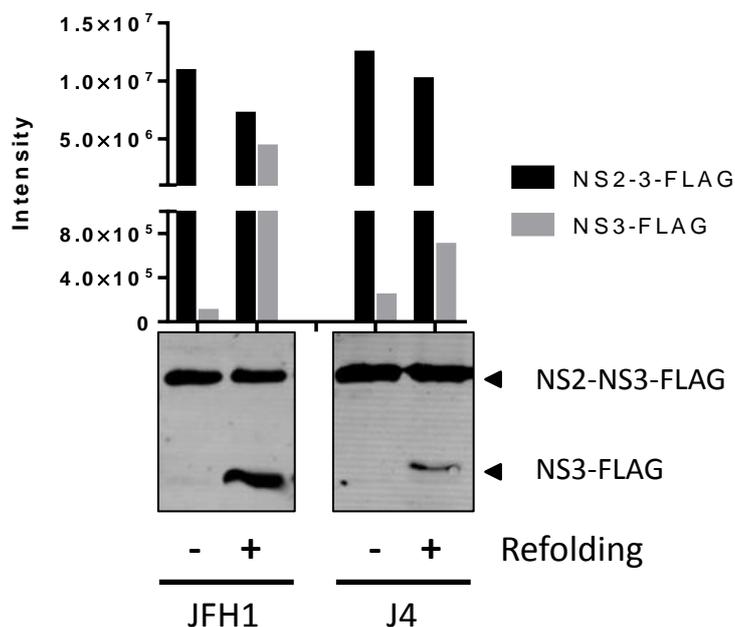
Figure 3.7: Comparison of *in vitro* autoproteolysis between different genotypes.

NS2-3 refolding reactions for the isolates JFH1 (gt 2a) and J4 (gt 1b) were compared at 0 hour (-) and 16 hour (+) timepoints by quantification of NS2-NS3-FLAG and NS3-FLAG following western blot analysis of reactions using αFLAG antibody.

As a negative control, NS2-3 refolding reactions were treated with an available DAA, the NS5A inhibitor Daclatasvir (BMS-790052) (Gao et al., 2010). This DAA is proposed to act upon the NS5A protein and as such should have no effect on NS2-NS3 processing. Firstly, to confirm that Daclatasvir does not act upon NS2, both SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] transcripts were electroporated into Huh7.5 cells and treated with a range of concentrations of Daclatasvir for 48 hours. Inhibition profiles were similar regardless of the presence or absence of NS2 in the SGR and calculated EC_{50} values were similar; 2.5 pM and 1.8 pM for [NS3-5B] and [NS2-5B] SGRs respectively (Figure 3.8A), similar to reported EC_{50} against JFH1 virus (28 ± 24 pM) (Gao et al., 2010).

After confirming that Daclatasvir does not exert any additional anti-viral effect in SGRs which are dependent on the NS2 autoprotease, a JFH1 NS2-3 refolding reaction was treated with the

same concentrations of Daclatasvir, ranging from 1 fM (10^{-3} pM) to 10 nM (10^4 pM) (Figure 3.8B). Quantification of NS3-FLAG from Daclatasvir treated samples confirmed no effect on levels of NS2-mediated autoproteolysis, even at active concentrations within the context of the SGR (Figure 3.8C). Thus active concentrations of the NS5A inhibitor Daclatasvir have no effect on the *in vitro* NS2-3 autoproteolysis assay and serve as a negative control.

The use of a positive control for screening assays is more challenging. As previously stated, several groups have attempted to identify inhibitors of the NS2 autoprotease without success, both in cell based screens and in similar *in vitro* refolding or *in vitro* transcription reactions using the minimal NS2-NS3 subunit. Initially, such *in vitro* reactions were used to characterise the nature of the NS2 autoprotease through the use of protease inhibitors, none of which showed any activity against NS2-NS3 processing (Pallaoro et al., 2001, Pieroni et al., 1997, Tedbury and Harris, 2007, Thibeault et al., 2001). The only agents reported as capable of blocking NS2 maturation *in vitro* are zinc chelators and some thiol alkylating agents. The inhibition of the NS2 autoprotease by thiol alkylating agents, along with mutagenic analysis, helped to define the NS2 protease as a cysteine protease.

As a positive control for the NS2-3 refolding reaction the halomethyl ketone thiol alkylating agent tosyl phenylalanine chloromethane ketone (TPCK) was used. NS2-3 refolding reactions were treated with TPCK at concentrations ranging from 1.7 µM to 300 µM alongside 0 hour and DMSO only controls. Western blot analysis confirmed the reduction or loss of NS3-FLAG in a dose-responsive manner above 9.5 µM TPCK (Figure 3.9A). NS3-FLAG was next quantified, normalised to DMSO controls and plotted to fit a non-linear regression, allowing the calculation of an EC_{50} of 15.3 µM (Figure 3.9B). These results confirm that the thiol alkylating agent TPCK blocks NS2-NS3 processing *in vitro* in a dose responsive manner. Furthermore, as a positive control TPCK demonstrates that quantification of NS3-FLAG from the NS2-3 refolding reaction allows the degree of NS2-mediated autoproteolysis to be accurately determined, to the extent that sensitivity of inhibition can be monitored through calculation of EC_{50}.

While useful as a positive control *in vitro*, toxicity analysis of TPCK in Huh7 cells revealed significant effects on cellular metabolism, with a CC_{50} of 3.3 µM (Figure 3.9C), not surprising considering the non-selective nature of this thiol alkylating agent. TPCK is therefore unlikely to be of any utility as a control in cell-based assays. Additional options for a positive control against NS2-dependent HCV replication in cell based assays are explored in Section 3.2.3.
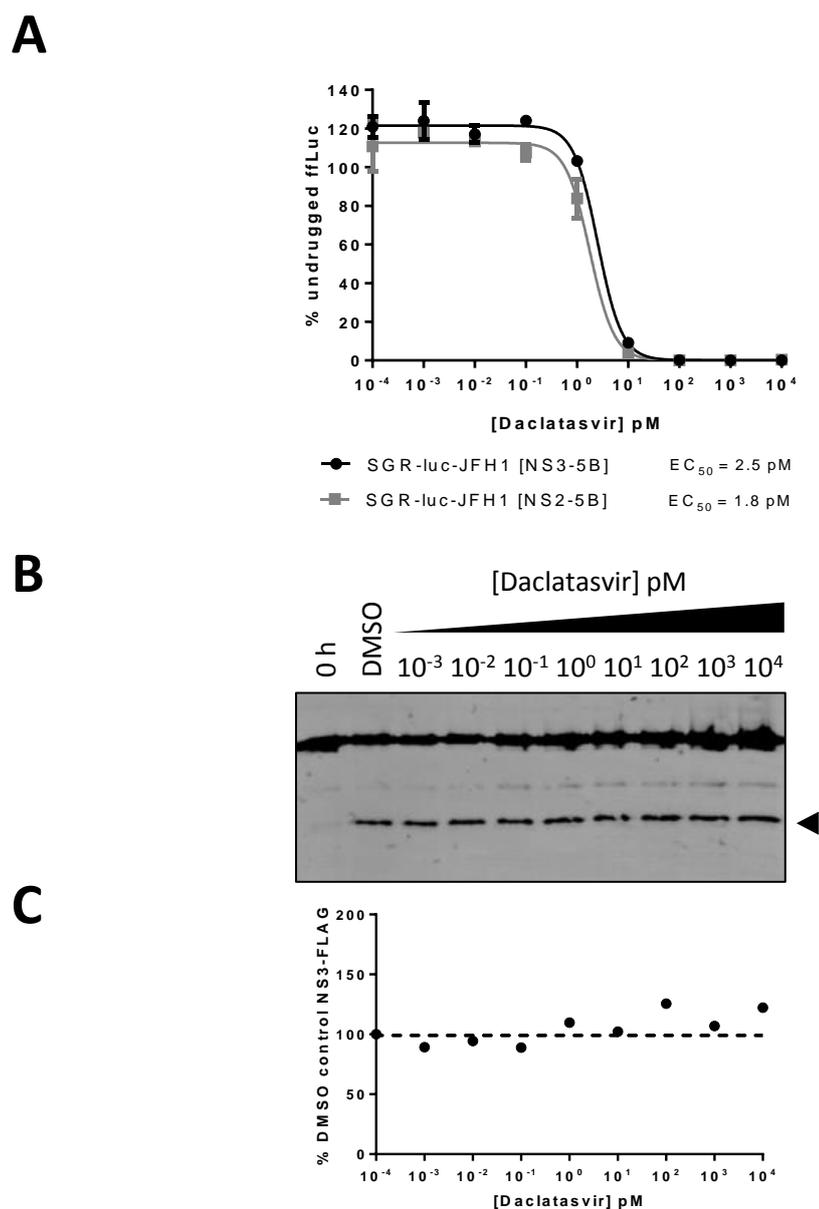
Figure 3.8: Daclatasvir has no effect on the NS2 autoprotease.

A) Luciferase reporter SGRs comprising NS3-5B or NS2-5B were electroporated into Huh7.5 cells and treated with a range of Daclatasvir concentrations from 1 fM to 10 nM for 48 hours. Luciferase activity was quantified and normalised to DMSO control (plotted as 10^{-4} pM). Non-linear regression was fitted to calculate EC_{50} B) JFH1 NS2-3 refolding reactions were treated with the same concentrations of Daclatasvir for 16 hours before assessing levels of NS3-FLAG (arrowhead) relative to 0 hour and DMSO only controls by western blot analysis. C) Quantification of NS3-FLAG from B confirmed no effect of Daclatasvir on in vitro NS2-3 autoproteolysis at active concentrations.
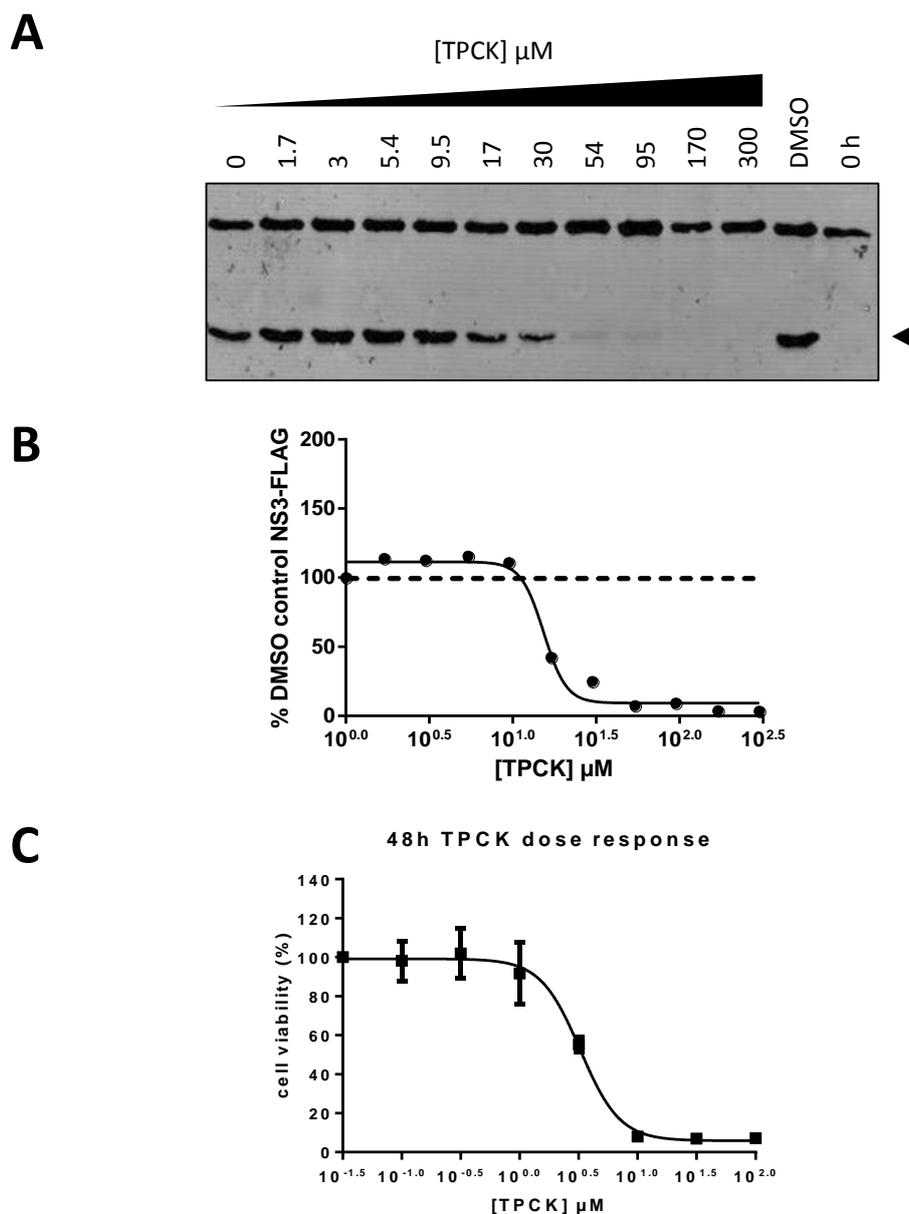
Figure 3.9: Thiol alkylating agents block *in vitro* NS2-3 autoproteolysis.

A) JFH1 NS2-3 refolding reactions were treated with a range of concentrations of tosyl phenylalanine chloromethane ketone (TPCK) from 1.7 μM to 300 μM for 16 hours alongside a DMSO only control and 0 hour control. Refolding reactions were analysed for levels of NS3-FLAG (arrowhead) by western blot using αFLAG antibody. B) Quantified NS3-FLAG from A was normalised to DMSO controls and plotted to fit a non-linear regression; EC_{50} = 15.3 μM. C) Huh7 cells were treated with a range of concentrations of TPCK from 0.1 μM to 100 μM for 48 hours before monitoring cellular metabolism by MTT assay. Data was normalised to DMSO controls and plotted to fit a non-linear regression and calculate CC_{50} = 3.3 μM.

3.2.2 NS2-5B SGR stable cell lines as a measure of NS2-dependent genome replication.

Previous work within this laboratory has utilised SGRs containing NS3-5B or NS2-5B with either a luciferase reporter in the first ORF (SGR-luc) for assays involving transient electroporation, or a NPT gene (SGR-neo) to allow selection of stable cell lines (Tedbury et al., 2011). More recently SGRs comprising NS3-5B derived from JFH1 and Con1 (BM4-5), a culture adapted genotype 1b SGR, were obtained as a kind gift from David Wyles, University of California. The first ORF of these SGRs encodes a ffLuc-NPT fusion protein. These constructs, termed SGR-feo, allow both analysis of SGR replication through luciferase activity and selection of stable cell lines (Wyles et al., 2009, Wyles et al., 2007). Due to elevated levels of luciferase achieved through constitutive SGR replication within stable cell lines and reduced variability usually introduced through the processes of preparing RNA transcripts and electroporation, these SGR-feo stable cell lines are more permissible to medium-high throughput screening.

To allow a similar system to be utilised for studying inhibitors of the NS2 autoprotease, the relevant NS2 coding sequence was introduced into both SGR-feo-JFH1 [NS3-5B] and SGR-feo-Con1 [NS3-5B] to produce SGR-feo-JFH1 [NS2-5B] and SGR-feo-Con1 [NS2-5B].

To produce SGR-feo-JFH1 [NS2-5B] a unique *EcoRI* site in the T7 promoter and a unique *PmeI* site at the start of the EMCV IRES were used to excise the HCV IRES and ffLuc-NPT fusion protein 'feo cassette' from SGR-feo-JFH1 [NS3-5B]. The same route was used to remove the NPT gene from SGR-neo-JFH1 [NS2-5B], allowing the NPT gene to be replaced by the feo cassette (Figure 3.10).

SGR-feo-Con1 [NS2-5B] was produced using the same *PmeI* site at the start of the EMCV IRES and a unique *XhoI* site within the NS5A gene to excise the 5' non-structural genes NS3-mid NS5A along with the EMCV IRES. The same fragment, containing NS2, was removed from SGR-neo-FK5.1 [NS2-5B] and ligated into the backbone of the SGR-feo-Con1 [NS3-5B] (Figure 3.11).

As this NS3-NS5A fragment contains all CAMs for both Con1 based SGRs (NS5A ΔS230 for BM4-5 or NS3 E176G, NS3 I254T, NS5A N137D and NS5A S225P for FK5.1), the produced SGR-feo-Con1 [NS2-5B] construct is essentially SGR-feo-FK5.1 [NS2-5B]. For simplicity SGR-feo-Con1(BM4-5) [NS3-5B] and SGR-feo-Con1(FK5.1) [NS2-5B] will both be referred to as Con1.
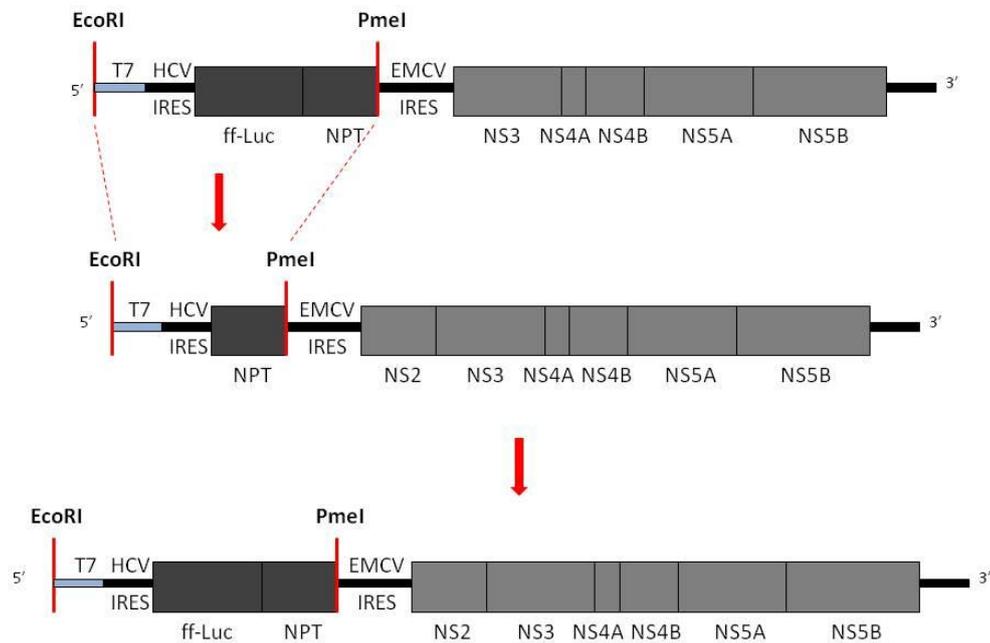
Figure 3.10: SGR-feo-JFH1 [NS2-5B] cloning strategy.

A unique EcoRI site at the start of the T7 promoter and a PmeI site unique to the start of the EMCV IRES were used to replace the neomycin phosphotransferase (NPT) gene of SGR-neo-JFH1 [NS2-5B] (middle) with the firefly luciferase-neomycin phosphotransferase fusion protein gene (ffLuc-NPT) from SGR-feo-JFH1 [NS3-5B] (top).
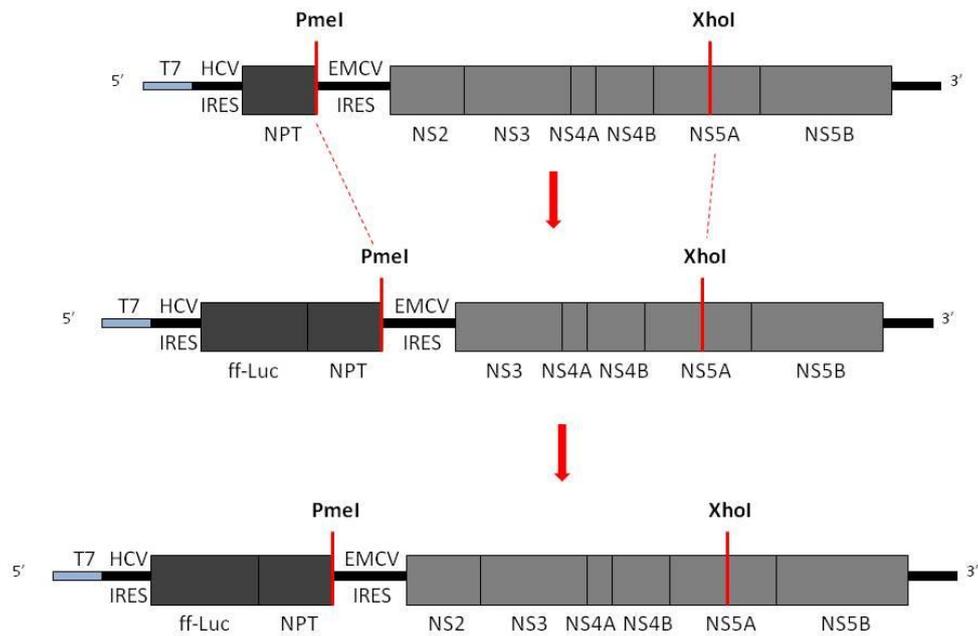
**Figure 3.11: SGR-feo-Con1 [NS2-5B] cloning strategy.**

A unique PmeI site at the start of the EMCV IRES and an XhoI site unique within the NS5A coding sequence were used to excise a fragment spanning the EMCV IRES – NS5A region and containing NS2 from SGR-neo-FK5.1 [NS2-5B] (top). This fragment was inserted in place of the corresponding EMCV IRES – NS5A region lacking NS2 of SGR-feo-Con1 [NS3-5B] (middle).

All SGR-feo [NS2-5B] constructs were confirmed by analytical digest and Sanger sequencing. Transcripts of SGR-feo-JFH1 [NS2-5B] and SGR-feo-Con1 [NS2-5B] were electroporated into Huh7.5 cells and subjected to selective pressure using G418. Separate plates were selected for either polyclonal or monoclonal stable cell lines, all of which produced detectable luciferase activity and could be cultured at up to 1 mg/ml G418. While luciferase activity was confirmed from monoclonal stable cell lines all future assays and the validation herein described for SGR-feo were performed with a polyclonal stable cell line population.

Once selection was complete and luciferase activity confirmed, stable cell lines were stored as frozen cell stocks as described in Section 2.2.3.3. Fresh stocks were thawed, cultured for at least two passages in the presence of G418 and seeded to the same cell density as after transient electroporation assays for analysis of luciferase signal and replication kinetics over various timepoints. In addition SGR-feo-JFH1 [NS3-5B] and SGR-feo-Con1 [NS3-5B] were analysed from frozen cell stocks in the same manner. SGR-feo [NS3-5B] stable cell lines were produced and selected by Yutaka Amako (University of Leeds) as polyclonal stable cell lines in Huh7.5 cells.

Previous reports have demonstrated a replication impaired phenotype for SGRs comprising NS2-5B (Lohmann et al., 1999). Such an effect has been observed in this laboratory using SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] in a transient luciferase reporter assay (Tedbury et al., 2011). Therefore if SGR-feo [NS2-5B] constructs rely on NS2-NS3 processing for genome replication, the same replication impaired phenotype should be apparent. In order to directly compare replication kinetics between SGR-luc and SGR-feo, SGR-feo-JFH1 [NS3-5B] or SGR-feo-JFH1 [NS2-5B] RNA was electroporated into Huh7 cells and analysed transiently i.e. at up to 72 hours post electroporation (h.p.e.) in the absence of G418. As a control for impaired replication when NS2 is present, transcripts of the previously reported SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] were also electroporated alongside a polymerase defective GND control (SGR-luc-JFH1 [NS3-5B] GND).

As shown in Figure 3.12, when SGR-feo-JFH1 [NS3-5B] or SGR-feo-JFH1 [NS2-5B] RNA was electroporated into cells and analysed transiently they showed comparable replication kinetics to previously reported constructs; SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B]. All constructs produced comparable 4 hour luciferase activity as a consequence of translation from the input RNA, irrespective of whether they contained luc or feo reporter genes. SGRs containing NS3-5B produced luciferase activity roughly 3 $\log_{10}$ higher than the GND SGR

negative control at 48-72 h.p.e. whether they contained luc or feo reporter genes. SGRs containing NS2-5B demonstrated a replication impaired phenotype, with luciferase activity over 1 $\log_{10}$ lower than the corresponding SGR spanning NS3-5B (Figure 3.12), in agreement with previously reported results (Lohmann et al., 1999, Tedbury et al., 2011). That the SGR-feo-JFH1 [NS2-5B] displays this phenotype as a transient experiment suggests a dependency on NS2-NS3 processing for replication.

However, following extensive selection with G418, cells stably-expressing the SGRs produced equivalent elevated levels of luciferase activity irrespective of the presence of NS2 in the polyprotein. Therefore, once selection was complete the phenotype of impaired replication due to NS2 appeared to be lost. As SGR replication is constitutive in these stable SGR cells, baseline luciferase activity is high at 4 hours post seeding (Figure 3.12). The subsequent 1 $\log_{10}$ increase in luciferase activity by 48 and 72 hours most likely represents an increase in cell numbers.
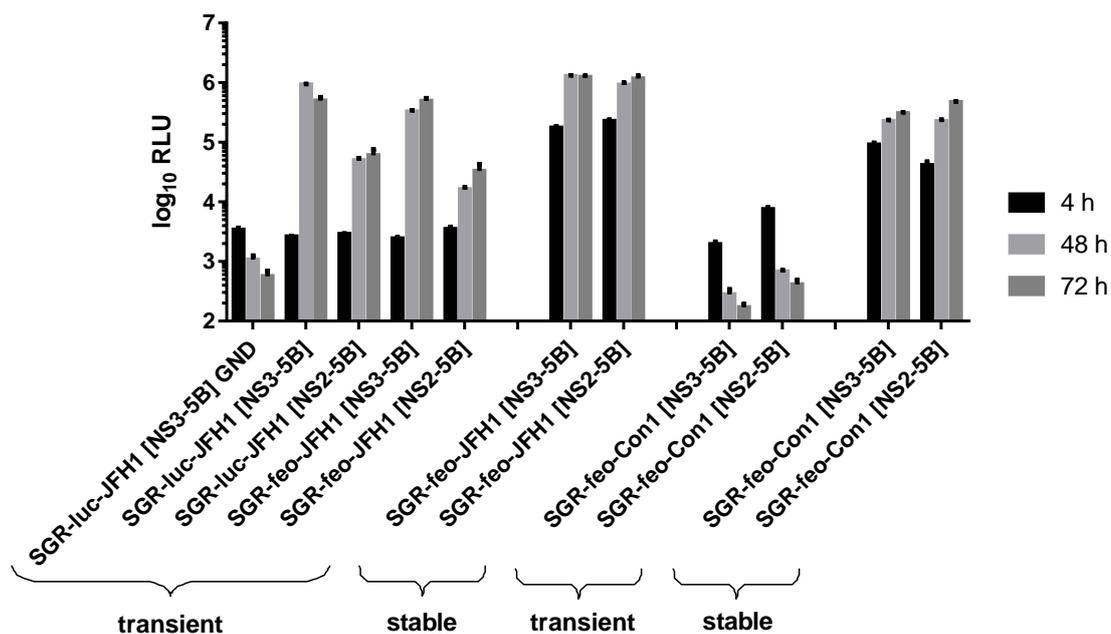
Figure 3.12: SGR replication kinetics assayed via luciferase activity.

Transcripts of indicated SGRs were either electroporated into Huh7 cells in the absence of selective pressure (transient) or electroporated into Huh7.5 cells and selected as polyclonal stable cell lines with G418. Stable cell lines (stable) were seeded alongside transient electroporations to directly compare replication kinetics. Luciferase activity at indicated timepoints post seeding is represented as relative luciferase units (RLU) on a logarithmic scale. Data represents the mean and standard deviation of one experiment seeded in triplicate.

Selection of SGR-feo [NS2-5B] as stable cell lines therefore appears to overcome the detrimental effect on replication imparted by the presence of NS2 in a 'transient' system. Providing that these stable cell lines are still dependent on NS2-NS3 processing for replication and have not acquired culture adapted mutations to improve replicative fitness (points that are addressed later in this chapter), the improved replication kinetics of a NS2-5B SGR as a stable cell line could be attributed to NS2-NS3 processing acting as a rate limiting step during the early establishment of HCV replication. As a transient experiment all non-structural proteins and RCs have to be produced *de novo* and as such replication may be more sensitive to slower polyprotein processing kinetics.

Stable cells harbouring SGR-feo constructs provide an ideal platform to screen for inhibitors due to the high luciferase activity compared to transient systems, resulting in a better signal-to-noise ratio. This effect is particularly pronounced with SGRs derived from genotype 1b. In a transient experiment both SGR-feo-Con1 [NS3-5B] and SGR-feo-Con1 [NS2-5B] produced no detectable luciferase activity over background, yet the selection of the same constructs as stable cell lines increased luciferase signal by over 3 $\log_{10}$ (Figure 3.12). The fact that the 'replication impaired' phenotype of an NS2-containing SGR is not observed in these stable cell lines is also beneficial for screening against NS2, but requires confirmation that genome replication within these cell lines is still dependent on NS2-NS3 processing. To validate these SGR containing stable cell lines further, the presence of the NS2 coding sequence in autonomously replicating RNA was confirmed and the potential for additional culture adapted mutations within these SGRs was examined. Additionally, the presence of mature viral proteins was explored to demonstrate normal processing of the polyprotein, particularly at the NS2-NS3 junction. Finally, the hypothesis that enhanced replication kinetics of SGR-feo [NS2-5B] was due to the selection of a sub-population of Huh7.5 cells more permissive to NS2-dependent HCV replication was tested.

Firstly, to confirm the presence of the NS2 coding region within the autonomously replicating RNA of SGR-feo [NS2-5B] stable cell lines, RNA was isolated by Trizol extraction following multiple passage of the cell population and used to perform an RT-PCR with NS2-specific primers. Available primer sets were used to amplify a 1280 base pair fragment for JFH1-based constructs and a 707 base pair fragment for Con1-based constructs. As a positive control for the fragments amplified with these primers a standard PCR was performed using the DNA template previously confirmed by Sanger sequencing. Amplified fragments of the correct size were observed for both positive control PCR reactions and RT-PCR reactions from total

extracted cellular RNA for SGR-feo-JFH1 [NS2-5B] and SGR-feo-Con1 [NS2-5B]. These products were not observed from identical reactions using SGR-feo-JFH1 [NS3-5B] or SGR-feo-Con1 [NS3-5B], demonstrating both JFH1 and Con1 SGR-feo [NS2-5B] stable cell lines retained the NS2 coding region within the autonomously replicating SGR RNA after the process of selection and further passages (Figure 3.13).

Improved replication kinetics of NS2-containing SGRs as stable cell lines could be attributed to culture-adapted mutations. To explore this possibility total cellular RNA from SGR-feo-JFH1 [NS2-5B] was subjected to RT-PCR with primers designed to amplify the entire coding region in multiple overlapping fragments. Amplified fragments were confirmed by agarose gel electrophoresis and sequenced to confirm the presence of the entire NS2 coding region and identify any potential culture adapted mutations. Analysis revealed 100% nucleotide identity to the sequence of the DNA construct throughout the entire SGR, demonstrating that the improved replication kinetics of SGR-feo-JFH1 [NS2-5B] in the context of a stable cell line could not be attributed to culture adapted mutations. Similar sequencing analysis was performed for SGR-feo-Con1 [NS2-5B], revealing no amino acid changes within NS2 or the NS3 N-terminus.

The cell lines were analysed for efficient processing at the NS2-NS3 boundary by probing cellular lysates for mature NS2. Replication kinetics of all four SGR-feo cell lines were directly compared in three independent experiments by seeding several timepoints in parallel before analysing luciferase activity, again confirming comparable kinetics. Lysates were analysed for the presence of post-cleaved NS2 and the endogenous marker GAPDH by western blot. As expected anti-NS2 antibody detected mature NS2 (23 kDa) from lysates of SGR-feo-JFH1 [NS2-5B] and infectious JFH1 virus, with none detected from lysates of SGR-feo-JFH1 [NS3-5B] (Figure 3.14A). The antibody used was raised against JFH1 NS2 (Jirasko et al., 2008) and hence showed no reactivity to either SGR-feo Con1 [NS2-5B] or SGR-feo Con1 [NS3-5B] lysates. The detection of a previously reported 17 kDa NS2 truncation product tNS2 (Boson et al., 2011, Jirasko et al., 2008, Jirasko et al., 2010) was also noted in both SGR-feo-JFH1 [NS2-5B] lysates and from infectious JFH1 virus lysates. In addition lysates were analysed for the presence of other key proteins; feo (anti-NPT), NS3 and NS5A (Figure 3.14B).
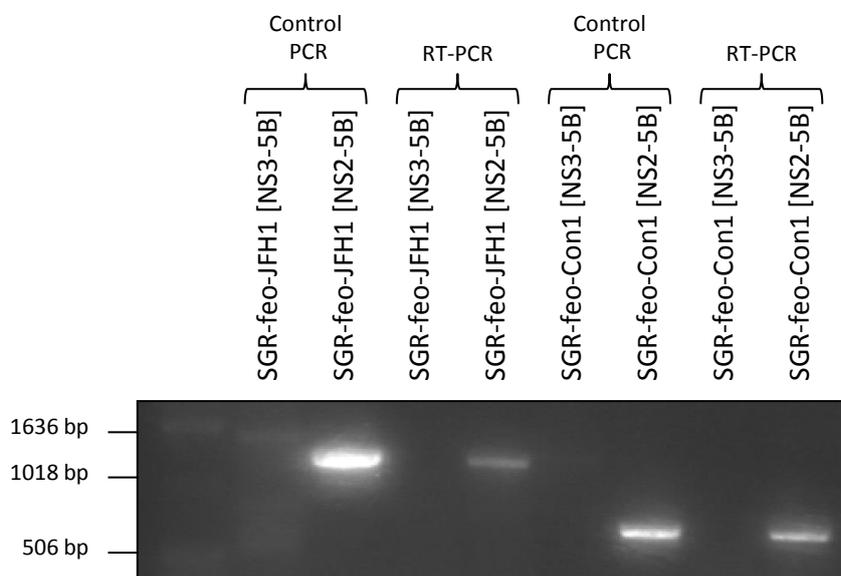
Figure 3.13: NS2-dependent RT-PCR from SGR-feo cell lines.

SGR-feo cell lines were passaged at least eight times before extraction of total cellular RNA followed by reverse transcription to cDNA using forward primers within the NS2 coding region (RT-PCR). Conventional PCR amplification of initial DNA constructs with the same primers was used as a control (Control PCR). PCR reactions were analysed by gel electrophoresis.
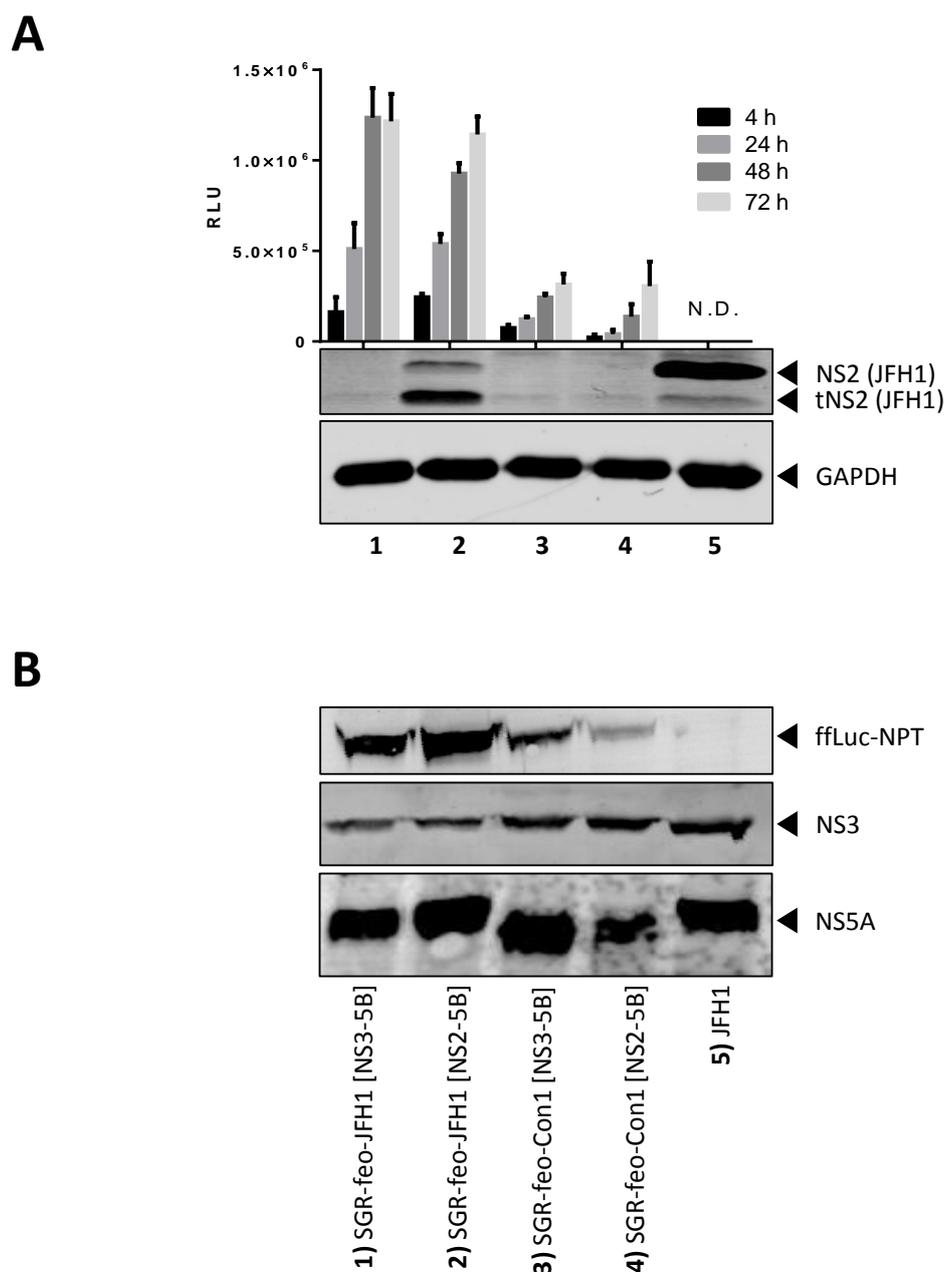
Figure 3.14: Comparative analysis of SGR-feo cell lines.

A) Luciferase activity from SGR-feo cell lines was analysed after indicated timepoints (N.D. – not determined). Data is the mean and standard deviation of three independent experiments. Below, lysates were analysed by western blot and probed with anti-JFH1 NS2 and anti-GAPDH antibody. 1) SGR-feo-JFH1 [NS3-5B], 2) SGR-feo-JFH1 [NS2-5B], 3) SGR-feo-Con1 [NS3-5B], 4) SGR-feo-Con1 [NS2-5B], 5) JFH1 HCVcc. B) Lysates probed with anti-NPT, anti-NS3 and anti-NS5A antibody.

One final consideration to explain the enhanced replication kinetics of stable NS2-5B SGRs compared to transient NS2-5B SGRs could be the selection of a cell population more permissive to the replication of NS2-containing SGRs. To exclude this possibility the replicon was cleared from SGR-feo-Con1 [NS2-5B] cell lines by treatment with 1 µM Daclatasvir (approximately 140,000 x EC_{50}) for 48 hours. Cells were maintained for a further 72 hours in the absence of Daclatasvir or G418 to remove the drug. Transcripts of SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] were then electroporated into the replicon-cleared cells and into both Huh7 and Huh7.5 cells for direct comparison. Similar to the data reported in Figure 3.12, SGRs containing NS2-5B had a severely impaired phenotype relative to SGRs containing NS3-5B. However this phenotype was much less pronounced in Huh7.5 cells (Figure 3.15). Cells derived from cleared SGR-feo-Con1 [NS2-5B] showed no further enhancement in the replication kinetics of a SGR spanning NS2-5B, with a slight replication impaired phenotype still observed (Figure 3.15). Luciferase from mock electroporated cleared-SGR cells was within background variation confirming the initial SGR-feo-Con1 [NS2-5B] had been cleared from these cells. Taken together these results demonstrate that while Huh7.5 cells allow for a significant improvement in the replication kinetics of NS2-containing SGRs, no further enhancement is observed from the cell population present at the end of SGR-feo-Con1 [NS2-5B] selection. Therefore a major influence on the replication kinetics of SGR-feo [NS2-5B] cell lines may be the use of Huh7.5 cells, though the process of constitutive autonomous replication may further elevate luciferase signals relative to transient electroporations.

SGR-feo [NS2-5B] stable cell lines therefore offer a robust cell-based assay with which to test inhibitors of NS2-dependent genome replication. These cell lines undergo polyprotein processing as a requirement for genome replication in the absence of additional CAMs. Genome replication can be monitored through luciferase activity with a larger signal-to-noise ratio than obtained in 'transient' systems and in the absence of impaired replication kinetics due to the presence of NS2.

Additional evidence that SGR-feo [NS2-5B] cell lines still require an active NS2 autoprotease for genome replication could be provided through the use of a positive control for selective inhibition of the NS2 autoprotease. As no specific small molecule inhibitor of the NS2 autoprotease has been reported, the next Section (3.2.3) will explore the potential of antagonists of cellular chaperones implicated in the NS2 autoprotease activity to serve as such a control.

Figure 3.15: SGR replication kinetics in cleared SGR-feo cell populations.

SGR-feo-Con1 [NS2-5B] was cleared with Daclatasvir and the resultant Huh7.5 derived cell population (Cleared SGR-feo (Huh7.5)) were subject to electroporation with SGR-luc-JFH1 [NS3-5B] or SGR-luc-JFH1 [NS2-5B] or mock treated. The same constructs were electroporated into Huh7 cells and Huh7.5 cells for comparison. Luciferase activity was measured after indicated timepoints and represents the mean and standard deviation of one experiment seeded in triplicate.

3.2.3 Chaperone antagonists as a positive control against NS2-dependent replication

Inhibitors of the cellular peptidyl-prolyl isomerase (PPI) cyclophilin A (CypA) are of interest as anti-viral agents targeting HCV replication, as described in Section 1.6.

In 2009, it was reported that both infectious virus and SGRs containing NS2-5B displayed enhanced sensitivity to CypA antagonists compared to SGRs containing NS3-5B (Ciesek et al., 2009). It was concluded that the isomerase activity of CypA played a role in the folding or activity of the NS2-NS3 auto-proteolysis event. If so, the differential sensitivity to CypA antagonsists observed between NS2-5B and NS3-5B replicons could serve as a useful positive control when screening compounds for specific inhibition of NS2-NS3 autoproteolysis. A control of this nature would benefit a screening programme aiming to identify a small molecule inhibitor of the NS2 autoprotease.

3.2.3.1 Effects of NS2 on SGR sensitivity to cyclosporin A

To explore previously reported effects of the CypA antagonist CsA on SGRs comprising NS3-5B and NS2-5B, SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] transcripts were electroporated into Huh7.5 cells. Electroporated cells were treated with CsA at a range of concentrations at 4 h.p.e. until 48 h.p.e. before measuring luciferase activity. EC_{50} values were calculated and the obtained data agreed with that reported (Ciesek et al., 2009); a shift in sensitivity was observed when comparing the two SGRs. SGR-luc-JFH1 [NS3-5B] exhibited a CsA EC_{50} of 481.4 nM, while the NS2-containing SGR gave an EC_{50} of 121.6 nM, a 4-fold increase in sensitivity (Figure 3.16A).

CsA EC_{50} values were next determined for the SGR-feo cell lines described in Section 3.2.2. Cells were seeded and again treated with a range of concentrations of CsA for 48 hours. Toxicity analysis confirmed that any effects observed were a direct effect on genome replication, but the same shift in sensitivity was not observed. While SGR-feo-Con1 [NS2-5B] (CsA EC_{50} = 499.2 nM) did appear slightly more sensitive than SGR-feo-Con1 [NS3-5B] (CsA EC_{50} = 686.3 nM) (Figure 3.16B), SGR-feo-JFH1 [NS2-5B] (CsA EC_{50} = 702.9 nM) was less sensitive to CsA inhibition than SGR-feo-JFH1 [NS3-5B] (CsA EC_{50} = 480.5 nM) (Figure 3.16C). These observations were confirmed by calculation of CsA EC_{50} values from three technical repeats (Table 3.1). SGR-feo-Con1 [NS2-5B] was 1.6 ± 0.3 fold more sensitive to CsA inhibition than SGR-feo-Con1 [NS3-5B], while SGR-feo-JFH1 [NS3-5B] was actually 1.3 ± 0.2 fold more sensitive than SGR-feo-JFH1 [NS2-5B] to CsA inhibition.

A

Huh7.5

SGR-luc-JFH1 [NS3-5B] EC_{50} = 481.4 nM
SGR-luc-JFH1 [NS2-5B] EC_{50} = 121.6 nM

B

Huh7.5 SGR-feo-JFH1

SGR-feo-JFH1 [NS3-5B] EC_{50} = 480.5 nM
SGR-feo-JFH1 [NS2-5B] EC_{50} = 702.9 nM
SGR-feo MTT CC_{50} = > 5,600 nM

C

Huh7.5 SGR-feo-Con1

SGR-feo-Con1 [NS3-5B] EC_{50} = 686.3 nM
SGR-feo-Con1 [NS2-5B] EC_{50} = 499.2 nM
SGR-feo MTT CC_{50} = > 5,600 nM
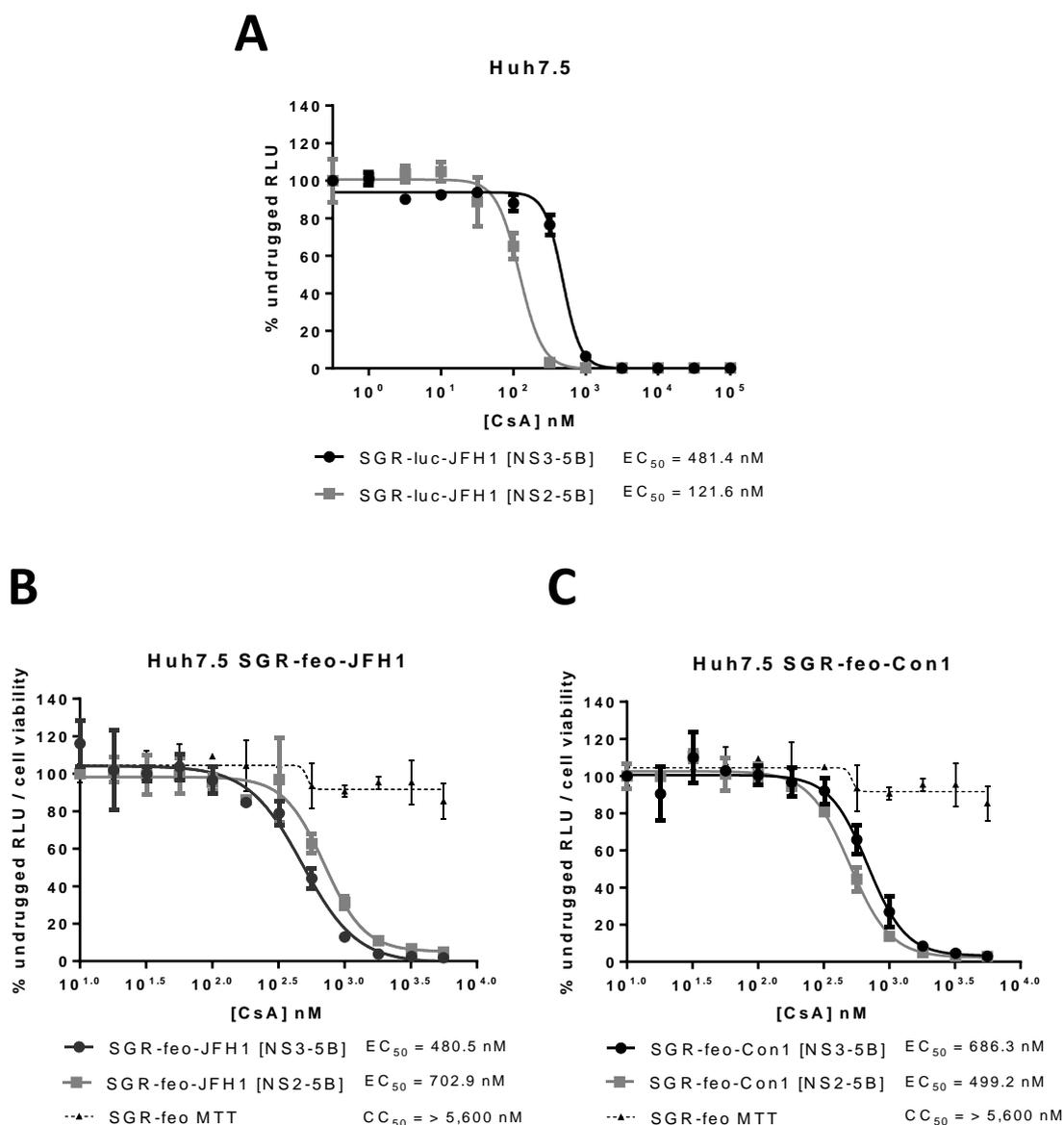
Figure 3.16: Cyclosporin A sensitivity in SGRs comprising NS3-5B or NS2-5B.

A) SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] were electroporated into Huh7.5 cells and treated with a range of cyclosporine A (CsA) concentrations from 1 nM to 100 µM for 48 hours to calculate EC_{50}. B) SGR-feo-JFH1 [NS3-5B] and SGR-feo-JFH1 [NS2-5B] stable cell lines treated with a range of cyclosporine A (CsA) concentrations from 10 nM to 5.6 µM for 48 hours to calculate EC_{50}. Cell viability was analysed by MTT assay. C) SGR-feo-Con1 [NS3-5B] and SGR-feo-Con1 [NS2-5B] stable cell lines treated with a range of cyclosporine A (CsA) concentrations from 10 nM to 5.6 µM for 48 hours to calculate EC_{50}. Cell viability was analysed by MTT assay.

Table 3.1: Sensitivity of SGR-feo cell lines to CsA.

EC_{50} calculated from treatment with a range of CsA concentrations over 48 hours in indicated SGR-feo cell lines. Three technical repeats for each assay are reported. Example dose response curves used for the calculation of EC_{50} are shown for experiment #1 in Figure 3.16B and C.

CsA sensitivity				
Experiment	[NS3-5B] EC_{50}	[NS2-5B] EC_{50}	Fold change in NS2-5B sensitivity	Average change in sensitivity
SGR-feo-Con1				
#1	686.3 nM	499.2 nM	x 1.4	X 1.6
#2	542.8 nM	355 nM	x 1.5	
#3	412 nM	211.9 nM	x 1.9	
SGR-feo-JFH1				
#1	480.5 nM	702.9 nM	x 0.7	X 0.8
#2	515.2 nM	587.3 nM	x 0.9	
#3	184.9 nM	234.1 nM	x 0.7	

Therefore neither SGR-feo [NS2-5B] stable cell line displays increased CsA sensitivity such as that observed with transient SGR-luc-JFH1 [NS2-5B], despite the presence of the NS2 coding region (Figure 3.13) and the observation of NS2 in lysates from SGR-feo-JFH1 [NS2-5B] (Figure 3.14). The shift in CsA sensitivity observed with transient replicons containing NS2-5B was therefore explored further to confirm a specific activity against NS2.

One key difference between the transient SGR-luc and the SGR-feo stable cell lines is that selection of stable cell lines may have improved the replication kinetics of NS2-containing SGRs

(Figure 3.15). It was therefore considered that CsA sensitivity may be linked to the replicative ability of NS2-5B SGRs. CsA sensitivity of transient NS2-5B and NS3-5B SGRs was therefore examined in the context of different cell lines. Previous electroporation of transient SGR-luc for CsA treatment was performed in Huh7.5 cells, as was the selection of SGR-feo stable cell lines. As described in Section 1.7, this cell line is more permissive to HCV replication. To test any potential effect of replication kinetics on sensitivity to CsA-mediated inhibition, SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] were transiently electroporated into the Huh7 cell line, where SGR replicative ability is reduced, an effect particularly pronounced in SGRs containing NS2-5B (Figure 3.12).

To confirm any observations made in NS3-5B and NS2-5B SGR-luc, identical experiments were performed with the same two SGRs containing a CsA resistance mutation within NS5A; D316E (Yang et al., 2010). SGR-luc-JFH1 [NS3-5B] (D316E) was previously cloned by Douglas Ross-Thriepland (University of Leeds). To produce SGR-luc-JFH1 [NS2-5B] (D316E), a fragment spanning NS3-NS5A was excised from SGR-luc-JFH1 [NS3-5B] (D316E) and ligated into *SpeI* and *RsrII* digested SGR-luc-JFH1 [NS2-5B]. The construct was confirmed by analytical digest and Sanger sequencing and *in vitro* transcribed and electroporated as before. This allowed analysis of the effect of NS2 on CsA sensitivity both within wt SGR-luc and in the context of a partially CsA resistant SGR-luc.

To explore the replication kinetics of wt and D316E SGRs spanning NS3-5B and NS2-5B, transcripts were electroporated in parallel into Huh7 and Huh7.5 cells. To account for potential differences in input RNA, luciferase activity is normalised as fold increase from 4 h.p.e. (at 4 h.p.e. no SGR replication is likely to have occurred and hence all luciferase present is a consequence of translation from the input RNA).

As shown in Figure 3.17, all SGRs produced detectable luciferase 48 h.p.e. significantly above GND negative control. SGR-luc-JFH1 [NS2-5B] (D316E) is replication competent with comparable replication kinetics in both cell lines to wt SGR-luc-JFH1 [NS2-5B].

The NS2 containing SGRs, both wt and D316E, had significantly impaired levels of replication 48 h.p.e. in Huh7 cells relative to identical SGRs containing NS3-5B. All SGRs showed comparable increases in luciferase signal within the context of Huh7.5 cells (Figure 3.17). A similar profile for wt NS2-5B and NS3-5B SGRs was described in Figure 3.15.

After confirming that all constructs were replication competent, CsA EC_{50} values were calculated for all constructs in parallel in both Huh7 and Huh7.5 cells (Figure 3.18).
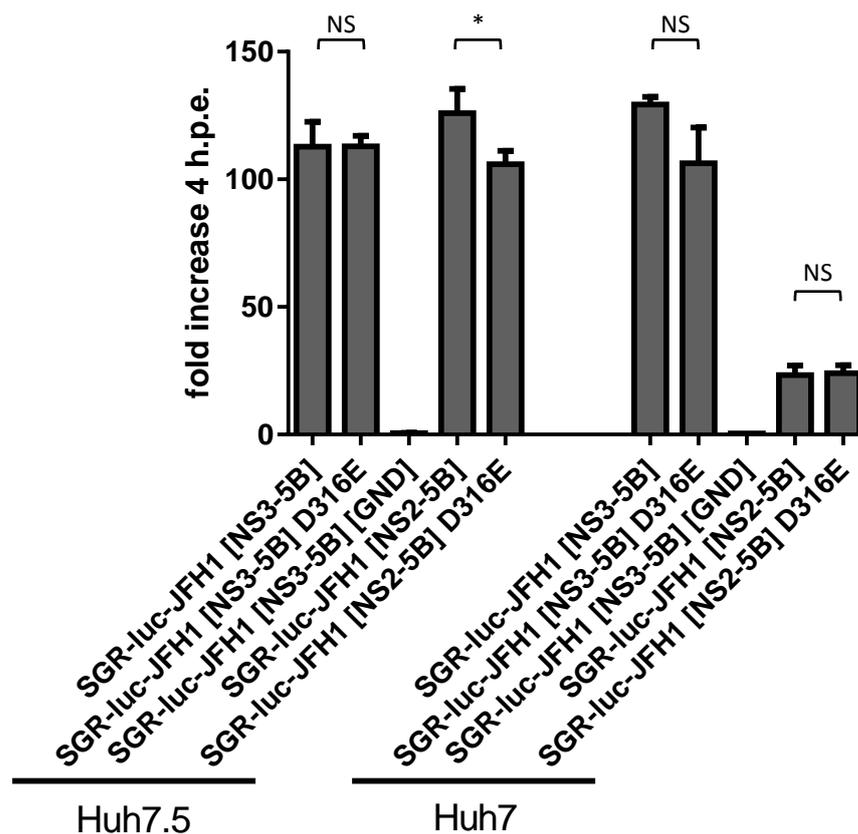
Figure 3.17: Effect of cell context on SGR replication kinetics.

Indicated SGRs were electroporated into Huh7.5 cells (left panel) or Huh7 cells (right panel) and luciferase activity monitored after 48 hours. Luciferase activity as fold increase from signal 4 hours post electroporation (h.p.e.) is plotted as the mean and standard deviation of each electroporation seeded in triplicate. Statistical significance was determined by unpaired Student's t test (* p < 0.05).
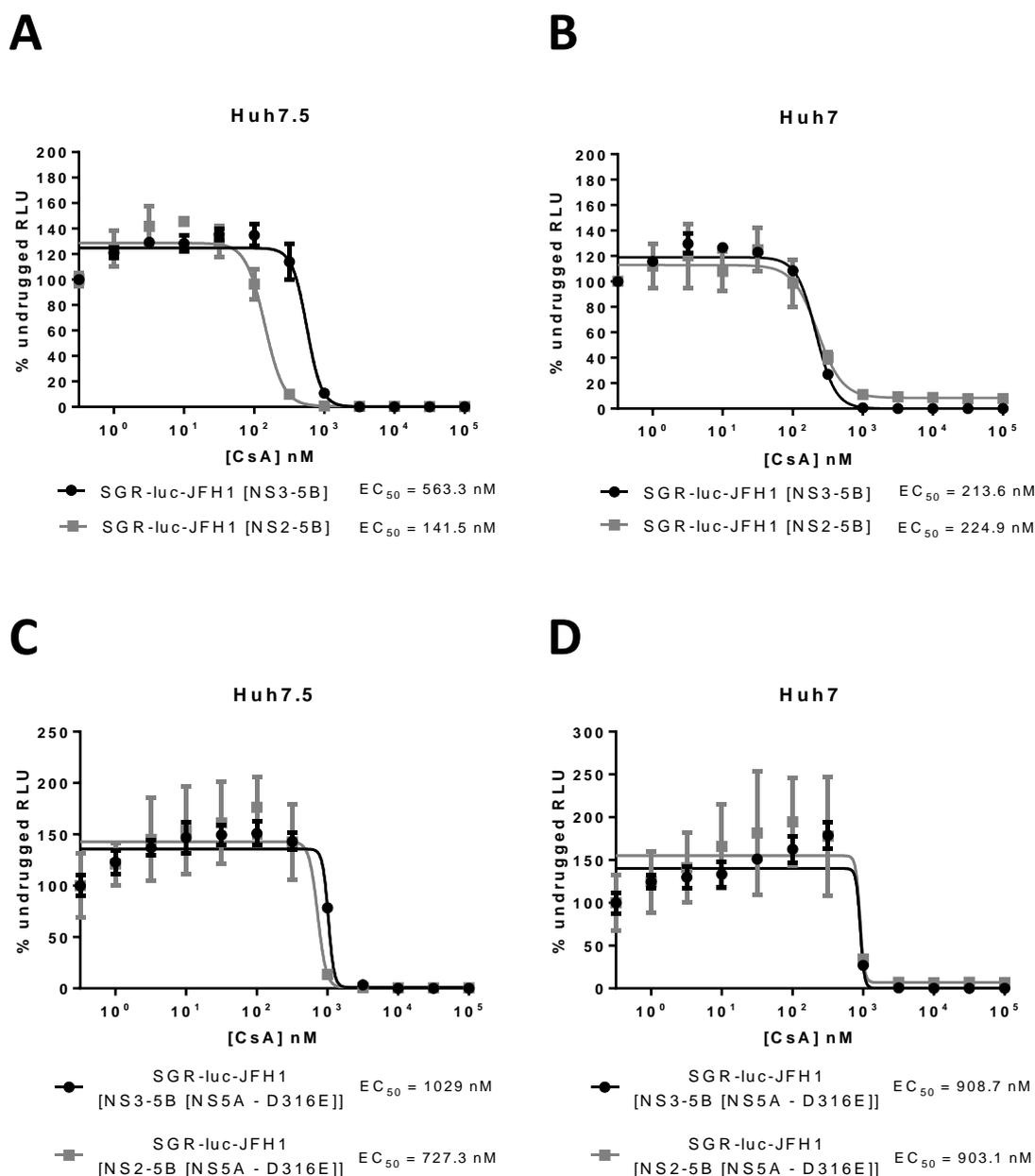
Figure 3.18: Effect of cell context on the CsA sensitivity of NS3-5B and NS2-5B SGRs.

A) wt SGR-luc-JFH1 [NS3-5B] and wt SGR-luc-JFH1 [NS2-5B] electroporated into Huh7.5 cells and treated with a range of cyclosporine A (CsA) concentrations from 1 nM to 100 µM for 48 hours to calculate EC_{50}. B) As in A with SGRs electroporated into Huh7 cells. C) As in A using SGRs containing the D316E CsA resistance mutation in NS5A. D) As in C with mutant SGRs electroporated into Huh7 cells.

As previously observed in Huh7.5 cells (Figure 3.16A), CsA treatment of transient SGR-luc-JFH1 [NS2-5B] produced a shift in sensitivity compared to the SGR-luc-JFH1 [NS3-5B] (Figure 3.18A). By contrast, when the same constructs were electroporated into Huh7 cells the sensitivity of SGR-luc-JFH1 [NS3-5B] and SGR-luc-JFH1 [NS2-5B] was much more similar (Figure 3.18B). Furthermore, a similar pattern was observed with SGRs containing the D316E CsA resistance mutation in NS5A. SGR-luc-JFH1 [NS2-5B] (D316E) was more sensitive to CsA inhibition than SGR-luc-JFH1 [NS3-5B] (D316E) when the experiment was performed in Huh7.5 cells (Figure 3.18C), but both SGRs showed identical sensitivity to CsA when the experiment was performed in Huh7 cells (Figure 3.18D). It was also noted that the presence of the D316E mutation was sufficient to impart resistance to CsA in both NS3-5B and NS2-5B SGRs in all cell lines.

As a mutation in NS5A is able to confer resistance to CsA regardless of the absence or presence of NS2, NS5A is likely to be directly involved in CsA-mediated inhibition. Furthermore, a shift in sensitivity to CsA-mediated inhibition between NS2-5B and NS3-5B SGRs is only observed in Huh7.5 cells and not Huh7 cells. This observation was made for both wildtype SGRs and partially CsA-resistant SGRs and implicates either an effect influenced by differential host factors between these two cell lines or an effect due to differences in SGR replication kinetics (Figure 3.17).

To explore these possibilities, the same experiments were performed using genotype 1b SGRs. These constructs have significantly reduced replication kinetics in cell culture, with replicative ability remaining poor even in Huh7.5 cells. SGR-luc-FK5.1 [NS3-5B] and SGR-luc-FK5.1 [NS2-5B] were electroporated into Huh7.5 cells and treated with a range of concentrations of CsA as before. Due to the reduced replicative ability of FK5.1 SGRs and the corresponding low luciferase signal, the experiments were performed over 72 hours for accurate calculation of EC_{50} values. Nonetheless as expected the sensitivity of SGR-luc-FK5.1 [NS2-5B] and SGR-luc-FK5.1 [NS3-5B] to CsA were identical, despite being performed in Huh7.5 cells (Figure 3.19A).

The shift in CsA sensitivity observed with NS2-containing SGRs (Ciesek et al., 2009) therefore appears to be an artefact of differences in replication efficiency specific to the strong levels of replication observed in JFH1-derived SGRs, and not a difference in host factor expression or activity. In agreement with this western blot analysis of lysates from Huh7 cells, Huh7.5 cells and from a range of SGR-feo constructs selected as stable cell lines did not show any significant differences in levels of endogenous CypA (Figure 3.19B).
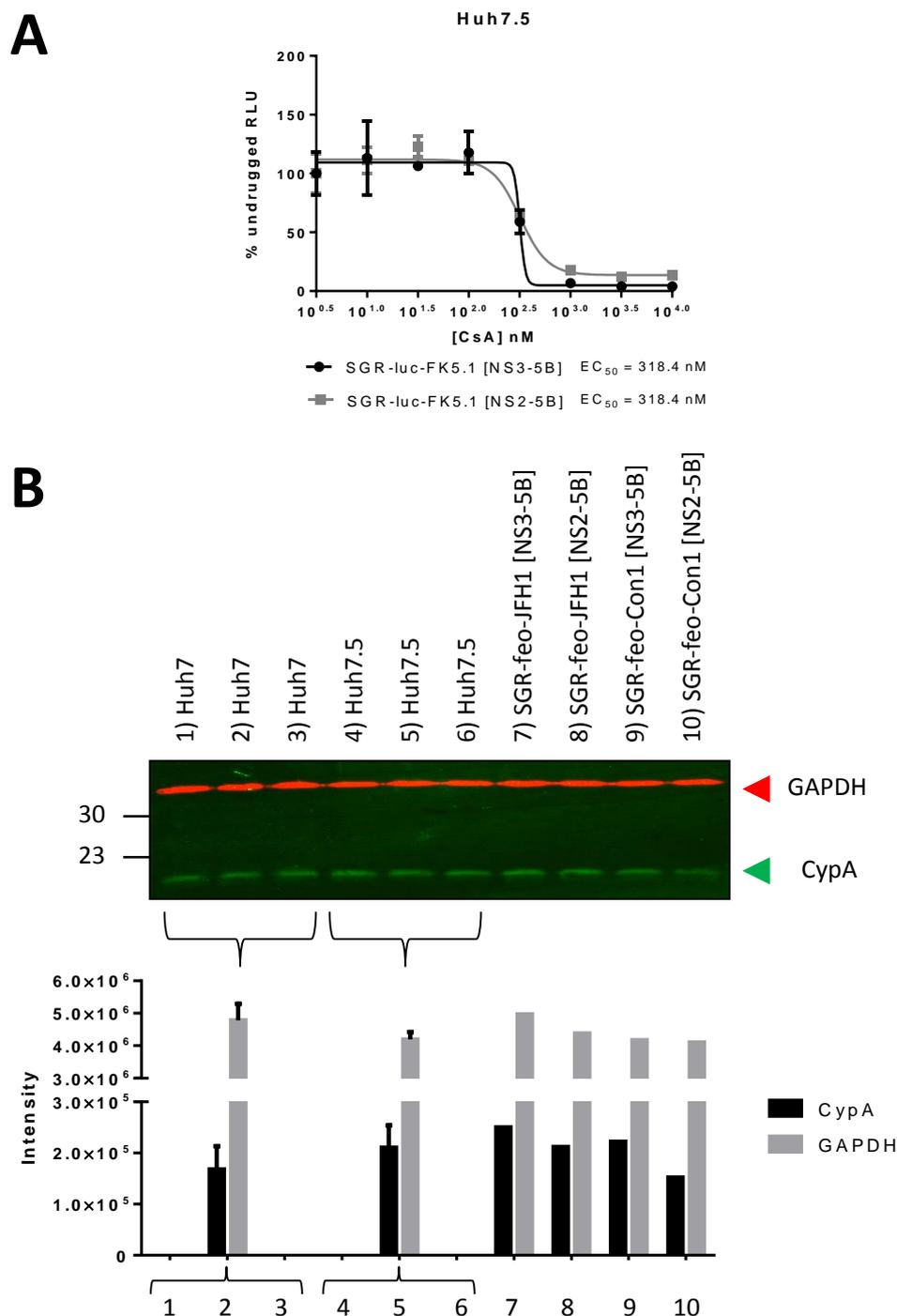
Figure 3.19: CsA sensitivity is not directly linked to cell context.

A) SGR-luc-FK5.1 [NS3-5B] and SGR-luc-FK5.1 [NS2-5B] electroporated into Huh7.5 cells and treated with a range of CsA concentrations from 3.16 nM to 10 µM for 72 hours to calculate EC_{50}. B) Cellular abundance of CypA and GAPDH in cell lysates analysed by western blot. Quantification is plotted with the mean and standard deviation from Huh7 and Huh7.5 shown.

3.2.3.2 Effects of CypA on NS2 refolding and autoproteolysis *in vitro*

Because antagonism with CsA is not directly affected by the NS2 protein, but appears to correlate with the replication kinetics of the SGR, this suggests that CypA plays no direct role in either the folding or autoprotease activity of NS2. To confirm this observation the effects of CypA were explored in the context of the *in vitro* NS2-3 refolding and autoprotease assay described in Section 3.2.1.

Ciesek *et al.* briefly described that CsA or other CypA antagonists had no effect on NS2-NS3 autoproteolysis in an assay based on *in vitro* translation of the minimal subunit required for NS2-NS3 processing in the presence of microsomal membranes (Ciesek et al., 2009). The design of this assay is somewhat similar to the *in vitro* NS2-NS3 autoproteolysis assay described in Section 3.2.1. However, the assay in the current study uses purified bacterially expressed NS2-NS3 and hence lacks any eukaryotic cellular proteins such as CypA. The successful autoproteolysis of NS2-3 in this setting demonstrates that NS2 does not have an absolute requirement for CypA for correct folding, though the inefficiency of this *in vitro* reaction could be explained by a lack of important cellular factors such as chaperones.

To examine whether a potential role of CypA in binding NS2-3 and aiding correct folding could be reconstituted *in vitro*, using autoprotease activity as a readout for the degree of correct NS2-3 folding, CypA was bacterially expressed as a GST fusion protein and purified by batch purification on Glutathione Sepharose 4B beads before being added to the *in vitro* NS2-3 autoprotease assay. GST-CypA has been confirmed to enhance the ability of NS5A domain 2 to bind RNA *in vitro*, suggesting functional isomerase activity (Foster et al., 2011).

Overexpression and purification of GST-CypA was performed both for wt CypA and a H126Q CypA mutant. This CypA mutation has been demonstrated to show less than 1% of wt PPIase activity (Chatterji et al., 2010) and thus served as a negative control to show whether observed effects were dependent on PPIase activity. Over-expression and purification by GST resin produced good yields of soluble protein; roughly 24 mg of protein from a 500 ml culture (Figure 3.20A). A series of dilutions for wt and H126Q mutants were analysed by western blot to confirm the purified protein was CypA (Figure 3.20B).

Purified CypA (wt or H126Q) was added to a JFH1 NS2-3 refolding reaction at 5 x the NS2-3 concentration (1 μM). After 16 hours the total amount of autoproteolysis was measured by quantifying the NS3-FLAG proteolysis product. Comparable levels of NS3-FLAG product were produced irrespective of the presence of GST-CypA (Figure 3.21A) suggesting that the addition

of 5x concentration of GST-CypA had no effect on the final amount of NS2-3 autoproteolysis within the context of the minimal NS2-3 subunit used in the assay. In addition, a timecourse was performed of the JFH1 NS2-3 refolding reaction in the absence or presence of 1 μM wt GST-CypA. Refolding reactions were terminated every 30 minutes by the addition of Laemmli SDS-PAGE running buffer before analysing by western blot for the levels of NS3-FLAG. Both refolding profiles appear similar with the control refolding reaction complete by 3 h and a reaction in the presence of 1 μM GST-CypA complete by 3 h (Figure 3.21B).

These results demonstrate that *in vitro* and within the context of the JFH1 NS2 C-terminal domain and the NS3 N-terminal domain, bacterially expressed GST-CypA does not increase the total amount of refolding and auto-proteolysis or the rate of NS2-3 refolding. These results support the previous data using NS2-containing SGRs which suggest that NS2 is not directly a substrate for the cellular PPIase CypA and as such does not directly influence the anti-HCV activity of CsA. Therefore, antagonists of CypA do not represent a viable positive control for specific inhibition of NS2-dependent HCV replication.
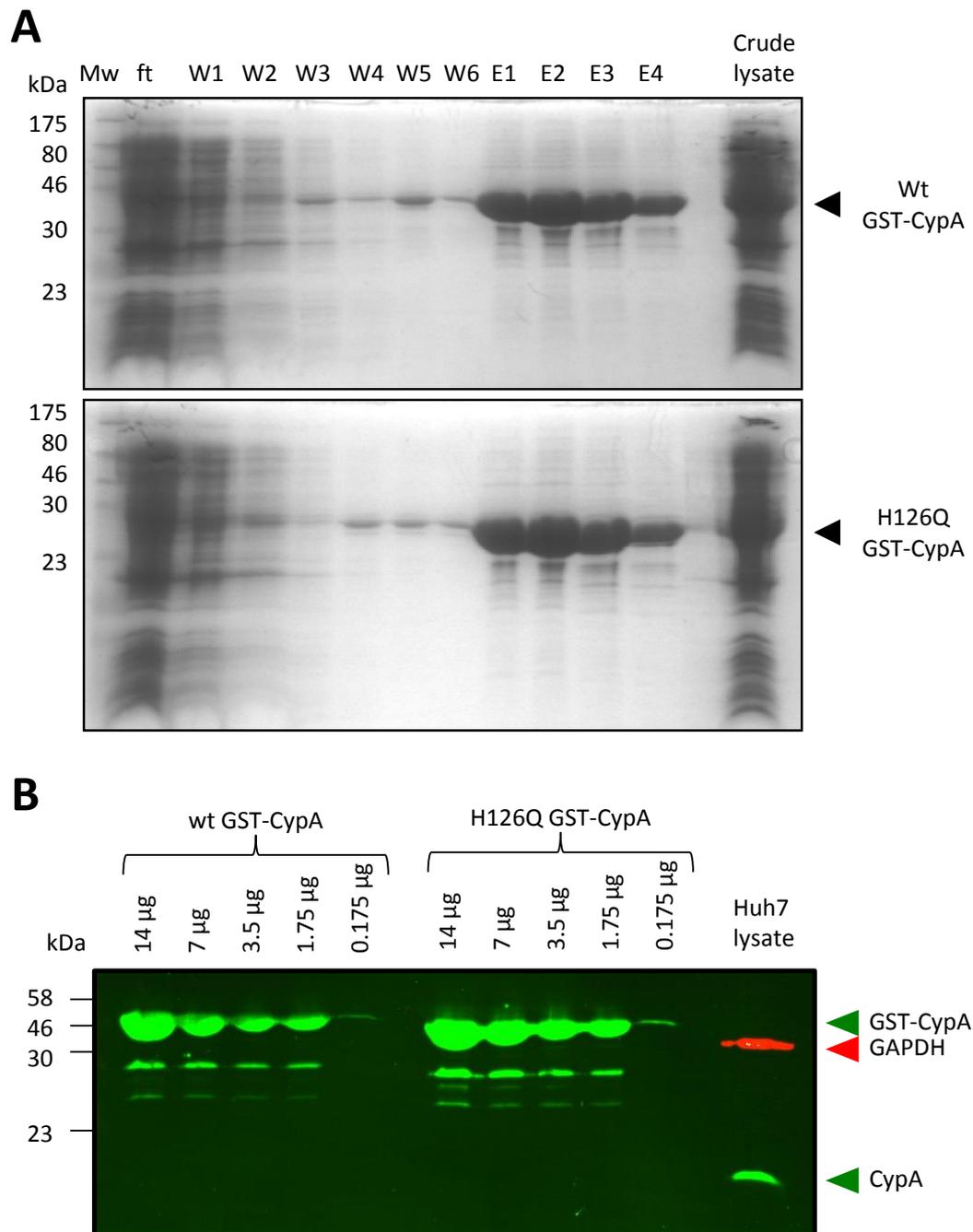
Figure 3.20: Overexpression and purification of GST-CypA.

A) Coomassie Brilliant Blue stained SDS-PAGE of fractions from wt and H126Q GST-CypA overexpression and purification. B) Western blot analysis of indicated amounts of GST-CypA from elution fractions. A Huh7 lysate control is confirmed by GAPDH and demonstrates reactivity of the antibody to endogenous CypA.
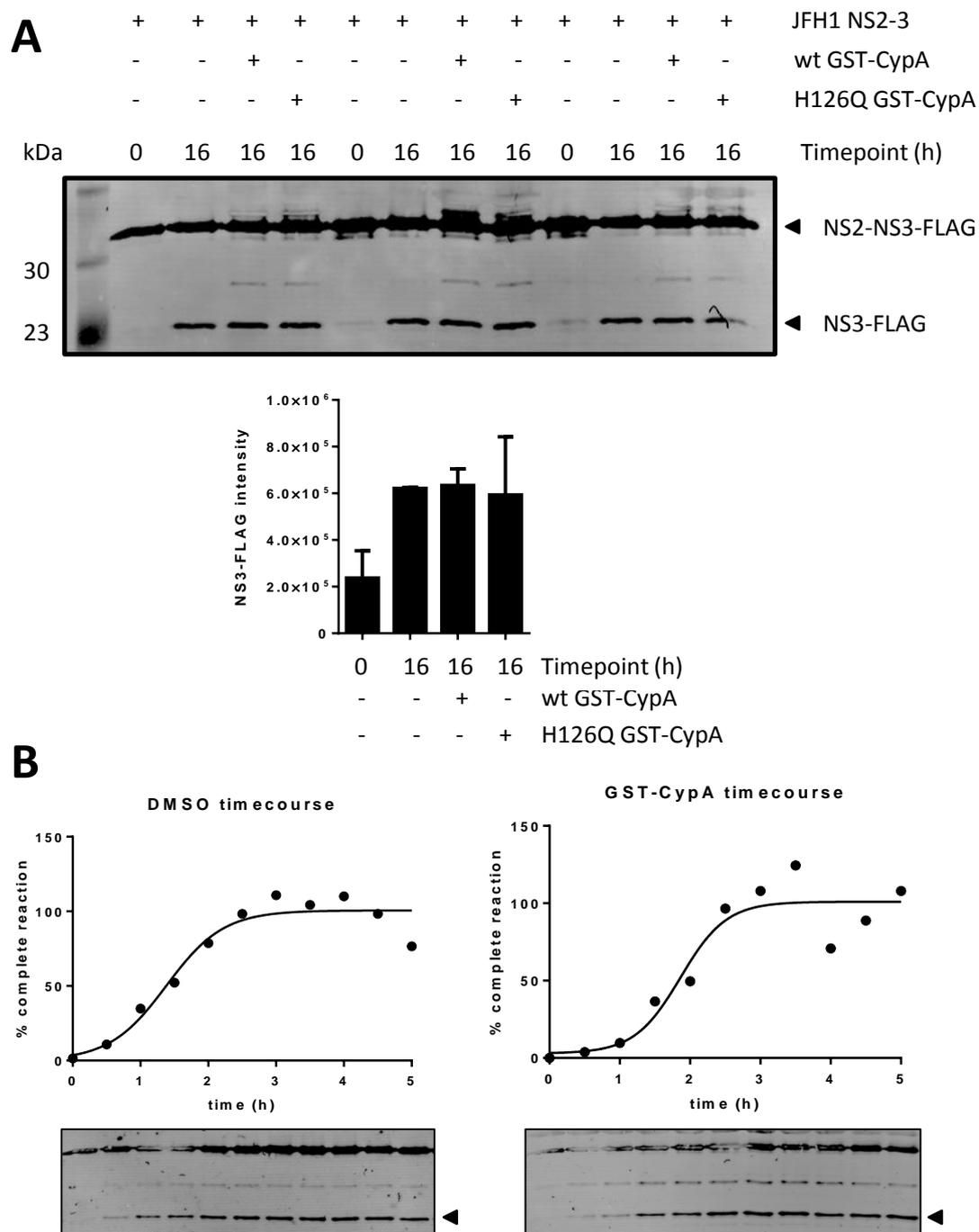
Figure 3.21: Effect of GST-CypA on NS2-3 refolding and auto-proteolysis *in vitro*.

A) Addition of wt GST-CypA (1 µM) or H126Q GST-CypA (1 µM) to JFH1 NS2-3 (200 nM) had no effect on the levels of NS3-FLAG produced by auto-proteolysis following western blot analysis. B) Addition of wt GST-CypA (1 µM) did not increase the rate at which refolding and auto-proteolysis occurs. Control (0.75% DMSO) refolding reaction complete by 3 h, GST-CypA treated refolding reaction complete by 3 h. Quantified NS3-FLAG is indicated by the arrowhead.

119

3.2.3.3 Analysis of NS2-dependent anti-HCV activity of Hsp90 inhibitors

In addition to observations linking CypA to the folding of NS2, the cellular chaperone Heat-shock protein 90 (Hsp90) has been implicated. An interaction of Hsp90 with NS2-3 *in vitro* has been demonstrated and the same system showed activity for the Hsp90 inhibitors geldanamycin and radicicol against NS2-NS3 processing *in vitro*. Subsequently these Hsp90 inhibitors were reported to block NS2-mediated proteolysis in a cell-based system with EC_{50} values in the nM range (Waxman et al., 2001). However, this system used NS2 proteolysis to stabilise a reporter in Jurkat cells, hence this system does not reveal whether Hsp90 is required for NS2 function in hepatoma cells or whether Hsp90 inhibitors display activity against NS2-dependent HCV replication.

To explore the potential for Hsp90 inhibitors to serve as a positive control for inhibition of NS2-containing SGRs two Hsp90 inhibitors; radicicol and geldanamycin, were tested for anti-viral activity against SGR-feo-JFH1 [NS3-5B] and SGR-feo-JFH1 [NS2-5B] stable cell lines alongside toxicity analysis. As shown in Figure 3.22, both compounds showed significant cytotoxicity which matched luciferase activity. This toxicity precluded the observation of any anti-viral effect. A dose response treatment with radicicol yielded a CC_{50} of 695.4 nM, with similar EC_{50} values calculated from luciferase activity demonstrating that the reduction of genome replication correlates with toxic effects on the cells (Figure 3.22A). A geldanamycin dose response curve produced a CC_{50} of roughly 74 nM, with the high degree of toxicity observed preventing accurate calculation of EC_{50} values from luciferase activity within this concentration range. Despite this the luciferase activity again appears to match the levels of cellular metabolism for all concentration points indicating effects are not anti-viral.

The high degree of toxicity observed with Hsp90 inhibitors in Huh7.5 cells therefore does not allow any anti-viral effect to be examined and prevents their use as a positive control in SGR-based assays.
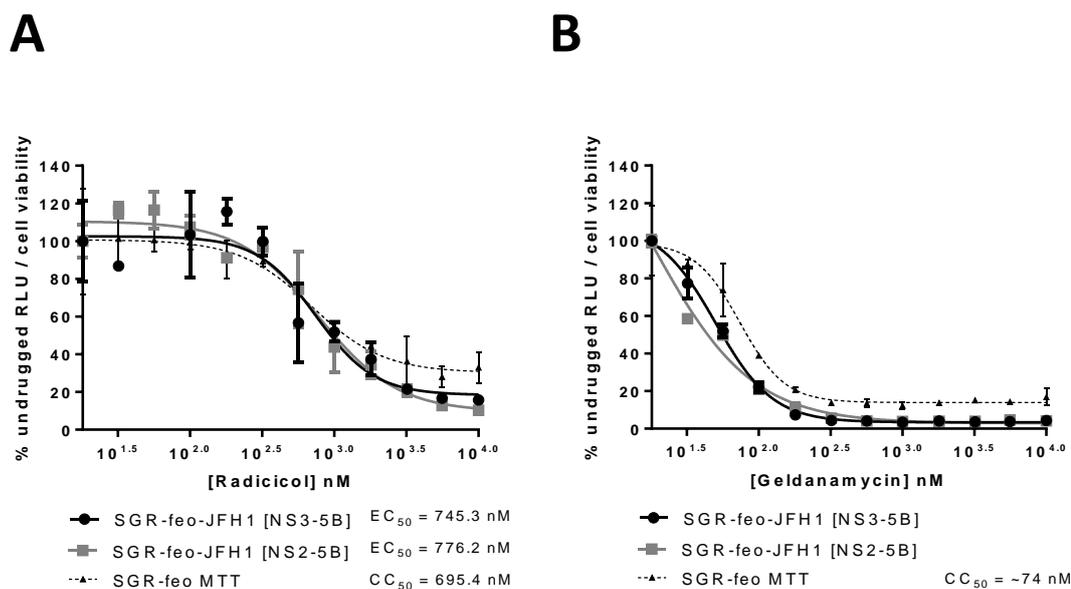
Figure 3.22: Anti-viral effect of Hsp90 inhibitors.

A) SGR-feo-JFH1 [NS3-5B] and SGR-feo-JFH1 [NS2-5B] were treated with a range of Radicicol concentrations before quantifying luciferase activity alongside toxicity analysis. Significant cytotoxicity was observed, with a CC_{50} of 695.4 nM, preventing the analysis of anti-viral effects specific for NS2. B) Identical experiment as in A) with Geldanamycin. Geldanamycin also shows significant toxicity to the extent that EC_{50} curves could not be accurately fitted within the tested concentration range.

3.3 Discussion

3.3.1 NS2-3 *in vitro* refolding as an assay for NS2-mediated proteolysis

Results in Section 3.2.1 describe the optimisation of a method of refolding a bacterially expressed, purified NS2-3 precursor for efficient processing at the NS2-NS3 boundary by the NS2 autoprotease. NS2-mediated autoproteolysis could be monitored through quantitative western blot analysis (Schutz-Geshwender, 2004) by detection of the NS3 proteolysis product via the FLAG tag. This refolding reaction as a measure of autoprotease activity can tolerate suitable levels of DMSO as a vector for the screening of compounds and can be performed over 4-16 hours. Positive controls for inhibition of autoproteolysis, as measured through prevention of NS3-FLAG production, are available by disruption of the NS2-NS3 protein fold with zinc chelators, through alkylation of thiol groups with TPCK and by comparison to refolding reactions terminated before refolding is allowed to occur (0h). HCV DAAs targeting the alternative non-structural protein NS5A serve as a negative control, with no effect on NS2 autoproteolysis at concentrations active in cell culture models. An optimal signal to noise ratio from this assay is achieved with the JFH1 isolate of NS2-3, with sufficient signal to accurately fit a non-linear regression for the calculation of EC_{50} values demonstrated by a dose-response treatment with TPCK (Figure 3.9). The ability to fit EC_{50} models allows direct comparison of activity against the NS2 autoprotease for ranking of compounds during screening and for structure activity relationship (SAR) analysis of hit compounds.

Due to the western blot readout of the assay, the NS2-3 refolding assay is not amenable to HTS and is technically demanding. It does however offer a sensitive assay for assessing activity directly against the NS2 autoprotease in an *in vitro* setting in the absence of potential off-target effects. Additionally, the reasonably high tolerability for DMSO should allow screening of compounds at high concentrations.

3.3.2 SGR-feo stable cell lines as an assay for NS2-dependent HCV replication amenable to HTS.

The selection of SGR-feo [NS2-5B] produces stable cell lines with luciferase activity levels comparable to SGR-feo [NS3-5B] cell lines. The elevated luciferase from these stable cell lines can be at least partially attributed to the use of the more permissive Huh7.5 cell line, but may also be influenced by constitutive autonomous replication overcoming a more pronounced

rate-limiting effect of NS2-NS3 processing at early stages in the establishment of replication. SGR-feo [NS2-5B] cell lines produce mature NS2 and contain no culture adapted mutations following selection indicating a normal requirement for processing at the NS2-NS3 junction to allow replication.

SGR-feo [NS2-5B] stable cell lines offer a robust, sensitive cell-based assay to screen for inhibitors of NS2-dependent replication. Moreover, the nature of the NS2 autoprotease offers a unique opportunity in terms of a selective cell-based screen. As replication only requires the NS2 autoprotease in the presence of an NS2-NS3 precursor in the polyprotein, SGRs spanning NS2-5B and NS3-5B are both replication-competent. Therefore SGR constructs that differ only by the presence or absence of NS2 provide one construct (SGR [NS2-5B]) dependent on the NS2 autoprotease for replication, and another (SGR [NS3-5B]) which can replicate independently of NS2 autoprotease activity. As such if a compound were to block the NS2 autoprotease and exert an inhibitory effect, this effect should only be seen in the NS2-containing SGR, and not in the corresponding SGR comprising NS3-5B. The NS3-5B SGR therefore provides a useful counter-screen with which to rule out false positives that are cytotoxic, that inhibit alternate viral functions required for replication and even that act against the luciferase reporter. Any compound reducing luciferase from the SGR [NS2-5B] but not the SGR [NS3-5B] would be expected to exert its action through the NS2 protein. The fact that as stable cell lines these SGRs have comparable luciferase signal should allow for reliable interpretation of results from screening against both in parallel.

3.3.3 Effects of chaperone inhibitors on the NS2 autoprotease

Unfortunately no suitable positive control could be identified to selectively inhibit a SGR comprising NS2-5B. Despite previous reports suggesting Hsp90 inhibitors could potently block NS2-NS3 processing in a cell-culture based system, in Huh7.5 cells the inhibitors geldanamycin and radicicol show potent cytotoxicity with no additional anti-viral effect. While the cell-based assay in Jurkat cells described by Waxman *et al.* did not describe any counter-measures for false positives due to cytotoxicity, it is possible that these compounds have considerably different toxicity profiles within Huh7.5 cells and Jurkat cells.

An alternative option was to use the increased CsA sensitivity previously reported with NS2-containing SGRs (Ciesek et al., 2009). However, while this effect could be reproduced in a transient system in Huh7.5 cells, the same shift in sensitivity was not observed between SGR-feo [NS2-5B] and SGR-feo [NS3-5B] stable cell lines. Upon further analysis, the shift in CsA

sensitivity when the SGR includes NS2 appeared specific to transient electroporations in Huh7.5 cells and was not observed in Huh7 cells. In line with this all previous experiments reporting this effect have been performed in cleared SGR Huh7 cell lines similar to Huh7.5 cells (Ciesek et al., 2009, Kaul et al., 2009) (Lohmann, pers comm).

Analysis of calculated EC_{50} values offers further insight into the cause of this effect. If the comparable CsA EC_{50} values in Huh7 cells for SGRs containing NS3-5B and NS2-5B is taken as the 'normal' CsA sensitivity, then in a Huh7.5 system, the EC_{50} of the NS3-5B SGR appears to have the greatest shift in sensitivity. While the CsA EC_{50} in Huh7 cells is 219.3 ± 8 nM (n=2; Figure 3.18B), the CsA EC_{50} of a SGR containing NS3-5B in Huh7.5 cells is 522.4 ± 58 nM (n=2; Figure 3.16A and Figure 3.18A), a 2.4-fold shift in sensitivity.

This effect could be due to the rapid replication kinetics of a SGR derived from JFH1 that comprises NS3-5B. Such SGRs replicate very well in Huh7 cells. Replication kinetics by luciferase activity in Huh7.5 cells do not appear to be enhanced any further for these constructs (Figure 3.17), though it is possible these systems are saturated. SGR-JFH1 [NS3-5B] may be unique in that its replication cycle is limited only by the availability of essential cellular co-factors. As such when a proportion of HCV replication is inhibited by an antagonist, it may free essential host factors which can facilitate the formation of new RCs using additional NS proteins which would otherwise be redundant. A greater concentration of antagonist would therefore be required to exert an anti-viral effect. It is important to note that in the context of differential CsA sensitivity this 'rate-limiting' host factor is not proposed to be CypA but could be one of a number of other host factors known to be important for formation of the RC, such as the lipid kinase PI4K-α, endosomal Rab proteins or a member of the VAP family of proteins involved in vesicular trafficking, amongst others (Berger and Randall, 2009).

In effect, NS3-5B SGRs derived from JFH1 may replicate better in Huh7.5 cells than Huh7 cells, but the saturation of a cellular host factor(s) may mask this effect in a 48 hour luciferase timepoint, as the SGR essentially already replicates to the maximal level permitted by the essential host co-factor(s). This effect is therefore only observed when replication is inhibited pharmacologically, wherein inhibiting replication only serves to free the co-factor(s) which can contribute to new RCs. As such a greater concentration of antagonist is required to block replication, and hence the SGR is less sensitive to inhibition. This effect is linked to replication kinetics, but only to the remarkable replication kinetics of a SGR comprising NS3-5B derived from JFH1. Replication kinetics of JFH1 replicons containing NS3-5B may be reduced in Huh7

cells relative to Huh7.5 cells (though the saturation of host factors limits replication in Huh7.5 cells and masks this effect). As SGRs derived from genotype 1b isolates have reduced replication kinetics in Huh7.5 cells the rate-limiting factor for replication is virally-encoded instead of host-factor derived, hence no shift in CsA sensitivity is observed.

Evidence for such an effect has recently been reported. Madan *et al.* recently demonstrated that an NS2-mediated shift in CsA sensitivity is not observed for SGRs derived from a genotype 1b isolate, as observed in this study (Madan et al., 2014). They also reported observing no shift in CsA sensitivity between SGRs comprising NS3-5B and NS2-5B which contained mutations in NS4B impairing the fitness of the SGR, despite all these experiments being performed in cleared SGR-cell lines. This evidence led to the conclusion that SGRs dependent on NS2-NS3 processing are more sensitive to CsA inhibition as the NS2 autoprotease reaction provides a rate-limiting step for replication of JFH1-derived SGRs (Madan et al., 2014). The data presented herein suggest a modified hypothesis, wherein a JFH1 SGR dependent on NS2-NS3 processing represents the 'normal' sensitivity to CsA. The effect observed is in fact a reduced sensitivity to CsA for SGRs lacking a virally encoded rate-limiting step (such as NS2-mediated processing or an alternative rate-limiting mutation) as in these cases replication is subject to a host-factor derived rate-limiting step. Assuming replication is only limited by the availability of host factors, in order to observe an actual reduction in replicative capacity an antagonist must first reduce the number of RCs to a point where host factor saturation does not occur. Such an effect would be observed through any pharmalogical inhibition of genome replication in this 'replication-saturated' state. In agreement, Madan *et al.* reported a correlation between replication kinetics of SGRs and sensitivity to NS5A inhibitors and NI NS5B inhibitors. This effect may be easier to further study through the comparative use of Huh7 and Huh7.5 cells as in this study, though previous analysis of sensitivity to an NS5A inhibitor using SGR-luc-JFH1 [NS2-5B] and SGR-luc-JFH1 [NS3-5B] in Huh7.5 cells suggested only a slight shift in sensitivity (Figure 3.8A). While Madan *et al.* suggested replication kinetics may be a useful determinant for the success of therapies with these compounds, the fact that the effect is specific to the JFH1 isolate, derived from a case of fulminant hepatitis, means that these findings may only be relevant in the rare cases of HCV causing fulminant hepatitis.

An alternative theory involves an effect on the cellular immune response. While overwhelming evidence confirms that CypA antagonists do not exert their anti-viral effect through the immunosuppressive effect of mediating a CypA-CsA-calcineurin interaction (Paeshuyse et al., 2006) (such an interaction blocks the phosphatase activity of calcineurin which would

otherwise mediate transcription of ISGs), recent evidence has linked CsA effects to immune responses in a different manner. CypA is now thought to form a double-stranded RNA (dsRNA) dependent interaction with protein kinase R (PKR) (Daito et al., 2014). Such a binding event leads to PKR phosphorylation, ultimately resulting in a block in translation of ISGs. CypA antagonists have been shown to prevent CypA-PKR interaction in the presence of dsRNA and as a consequence of alleviating a block on translation, upregulate the levels of ISG-encoded proteins (Daito et al., 2014). If CsA and other CypA antagonists can influence levels of ISG product translation, then altered levels of ISG transcription due to defective RIG-I in Huh7.5 cells could further influence how CsA affects a host immune response. How this effect might be mediated through the NS2 protein or the presence of the NS2 gene in the SGR RNA and why it is specific to the JFH1 isolate remains a mystery, and gaining a complete understanding of any potential effect would prove difficult due to the extensive feedback systems inherent in immune response pathways.

In either case significant evidence now demonstrates that the mechanism of action of CsA is not directly linked to the NS2 autoprotease activity. In agreement with this, data presented in Section 3.2.3.2 demonstrated that bacterially purified GST-CypA has no effect on the folding of NS2-3 *in vitro*. Other groups have reported similar findings, with no interaction of NS2 and CypA observed in cell culture (Madan et al., 2014). The same group failed to identify resistance to CsA within the NS2 coding sequence and noted that the presence of NS2 did not affect the ability of an NS5A encoded resistance mutation to impart CsA resistance. Both findings were replicated within the course of this study.

It therefore seems unlikely that CypA is required for correct folding of the NS2 autoprotease and hence CypA antagonists exert no specific anti-NS2 effect.

Despite failing to identify a positive control for specific inhibition of NS2-dependent replication, the use of SGR-feo-[NS2-5B] and SGR-feo-[NS3-5B] stable cell lines in parallel should aid identification of a small molecule inhibitor of the NS2 autoprotease. This cell-based screening assay can be complemented with the *in vitro* NS2-3 refolding assay to directly assess effects of a compound against NS2-mediated proteolysis.

Chapter 4 - Identification of small molecule inhibitors of the NS2 autoprotease

4.1 Introduction

A common route to developing a protease inhibitor is the use of a mechanism-based inhibitor which produces a covalent interaction with the catalytic residues through a reactive electrophilic unit often referred to as a 'chemical warhead' (Powers et al., 2002). Such an approach was used in the development of NS3 protease inhibitors, with the inhibitory activity of proteolysis peptides enhanced by conjugating reactive α-ketoamides (Han et al., 2000). Similar methods have not proven as successful for the HCV NS2 autoprotease. While NS2 mediated proteolysis can be inhibited *in vitro* by reactive halomethyl ketones, such as TPCK, less reactive mechanism-based protease inhibitors have not shown any activity (Table 1.2). An example of one such protease inhibitor is E64, a covalent cysteine protease inhibitor based around an epoxide warhead conjugated to a peptide derived backbone (Figure 4.1C). Despite relatively broad activity against cysteine proteases, E64 shows no activity against the NS2 autoprotease (Figure 4.1) (Pieroni et al., 1997, Thibeault et al., 2001). Epoxysuccinyl peptides such as E64 are considered less reactive than more crude alkylating agents such as halomethyl ketones, as the rate limiting step of inhibition is the binding of the compound to the active site cavity in a non-covalent manner. A transient interaction is usually mediated by hydrogen bond interactions with S sites (each S site refers to a region of the active site occupied by the corresponding P residue of the substrate peptide sequence, as depicted in Figure 4.8). This transient interaction is required for correct orientation of the epoxide near the catalytic thiol before opening of the epoxide to form a covalent interaction, a process which, similar to peptide bond hydrolysis, is energetically favourable but kinetically slow (Bihovsky et al., 1993, Powers et al., 2002). This affords covalent inhibitors based around an epoxide warhead a higher degree of specificity, such that E64 is active against a wide range of cysteine proteases but inactive against all serine, aspartic- or metalloproteases.

The key determinant of the interaction is the peptide-like backbone binding to S sites before any covalent complex is formed. As such, this peptide sequence can be further tailored to make epoxysuccinate peptide derivatives with enhanced specificity for certain cysteine proteases (Powers et al., 2002, Schiefer et al., 2013). A common starting point for design of mechanism-based inhibitors is therefore a suitable substrate peptide. However, a range of NS2 substrate peptides exert little or no inhibition on the NS2 autoprotease (Table 1.2) (Pallaoro et al., 2001, Thibeault et al., 2001) and the conjugation of mechanism-based inhibitors to substrate peptides also failed to yield an inhibitor of the NS2 autoprotease (Pallaoro et al., 2001).

The first section of this chapter (4.2.1) describes an examination of epoxide based warheads not conjugated to NS2 substrate peptides. Substrate peptide sequences seem unable to bind with high affinity into the NS2 active site, as evidenced by the lack of any potent inhibitory effect, so alternative backbone scaffolds were tested for their ability to allow an epoxide warhead to occupy the active site of NS2 and exert an inhibitory effect. A range of bis-amido epoxide compounds with functionalised aromatic groups were tested for the ability to block NS2-NS3 autoproteolysis *in vitro*. These compounds were designed and synthesised at the University of Leeds (Craig Avery, PhD Thesis) and show nM activity *in vitro* against an alternative cysteine containing system, factor XIIIa. Additionally, compounds from this series do not interact with free thiol agents in the form of up to 10 mM glutathione and exerted no activity up to 100 μM against an alternative cysteine protease Cathepsin S (Craig Avery, PhD Thesis). This selectivity against factox XIIIa suggests that selectivity to the NS2 autoprotease could be achieved.

While the use of an epoxide as a less reactive electrophilic warhead amenable to a greater degree of selectivity overcomes many of the pitfalls associated with covalent inhibitors, the majority of drug discovery programmes still aim to identify non-covalent, reversible inhibitors of their target. Primarily this is due to safety concerns surrounding covalent inhibitors due to off-target effects made more likely by the inherent reactivity of the molecules (Singh et al., 2011). Such a preference is evidenced by the progression of HCV NS3 protease inhibitors from reversible covalent inhibitors to non-covalent 2nd generation protease inhibitors. Identification of a more drug-like non-covalent inhibitor of the NS2 autoprotease requires screening of small molecule libraries.

As described in Section 1.9, hit rates for screening programmes can be enriched through CADD. As structural information is available for the post-cleaved form of the NS2 autoprotease, structure-guided CADD was applied to create an *in silico* enriched library for biological testing. Such an approach is particularly necessary for testing against the *in vitro* NS2-3 refolding assay described in Section 3.2.1 as this assay is not amenable to HTS, limiting the number of compounds that can be screened. Section 4.2.2 describes the application of structure-guided CADD and vHTS to a model of the NS2 protease using the aforementioned crystal structure so as to enrich a relatively small library with molecules predicted to be capable of forming favourable interactions with the NS2 active site.

The use of a more manageable screening library allows compounds to be tested against both the *in vitro* NS2-3 refolding assay, to determine ability to block NS2-mediated proteolysis, and

the cell-based assay using stable SGR cell lines, to test for specific effects against HCV genome replication which is dependent on the activity of the NS2 autoprotease. Section 4.2.3 concerns the biological testing of the *in silico* enriched library described in 4.2.2 and the use of these two independent screening assays to eliminate false positives and identify a more lead-like small molecule inhibitor of the NS2 autoprotease.

4.2 Results

4.2.1 Epoxide based inhibitors of the NS2 autoprotease

As demonstrated in Section 3.2.1, the NS2 autoprotease can be inhibited *in vitro* by halomethyl ketone alkylating agents such as TPCK (Figure 3.9). However, such reactive alkylating agents offer little scope for specificity and consequently TPCK shows off-target effects in cell culture in the form of cytotoxicity (Figure 3.9C). By contrast, less reactive alkylating agents such as the epoxysuccinate cysteine protease inhibitor E64 exert no activity against the NS2 autoprotease (Figure 4.1) (Pieroni et al., 1997, Thibeault et al., 2001). However, E64 is based around a peptide-mimetic backbone. Previous observations have shown only weak or undetectable inhibitory activity of substrate peptides against NS2-mediated proteolysis, indicative that the NS2 autoprotease may not be sensitive to the binding of peptides in *trans*.
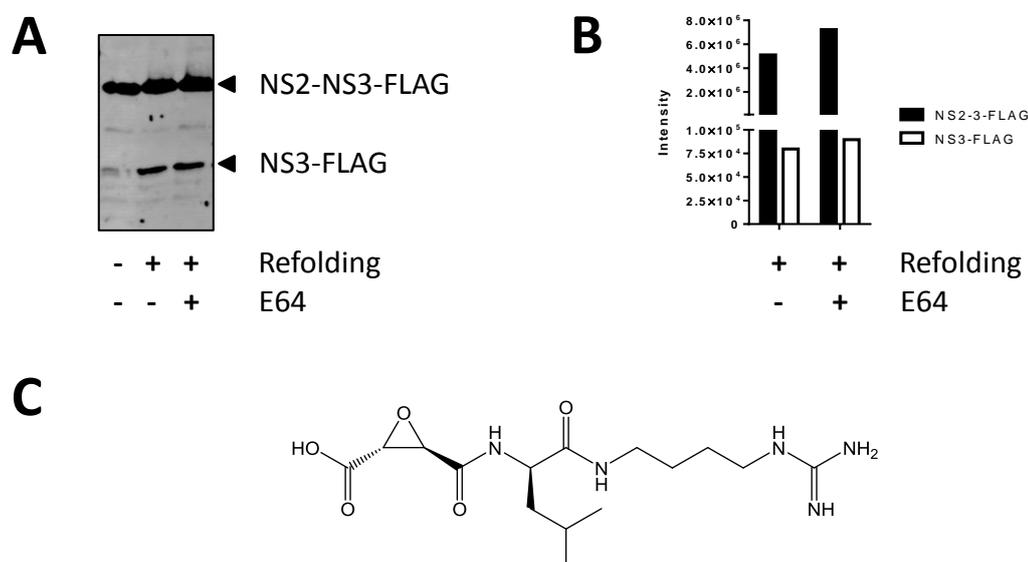

Figure 4.1: The NS2 autoprotease is not sensitive to E64.

A) JFH1 NS2-3 refolding reactions halted immediately or allowed to proceed in the absence or presence of E64 (300 μM). Reactions were analysed by western blot using αFLAG antibody. B) Quantification of NS2-3-FLAG and NS3-FLAG bands from refolding reactions allowed to proceed in A. C) Structure of E64.

To explore the use of an epoxide warhead conjugated to alternative scaffolds as a mechanism-based inhibitor of the NS2 autoprotease, NS2-3 refolding reactions were treated with a range of concentrations of bis-amido epoxides conjugated to simple or functionalised aromatic rings. Reactions were treated with a range of concentrations of each epoxide based compound numbered **1-12** (Table 4.1) and analysed by western blot for levels of NS3-FLAG (Figure 4.2). Dose response curves were plotted to allow calculation of EC_{50} values (Table 4.1). Several of the epoxide based compounds appeared to show activity against NS2-NS3 autoproteolysis *in vitro*. Initial testing identified compound **1**, which blocks production of the NS3-FLAG proteolysis product in a dose responsive manner (Figure 4.2, Figure 4.4A). Quantification of NS3-FLAG from two separate experiments yielded an EC_{50} of 91.6 ± 1.6 µM. To explore whether compound **1** exerted a pan-genotype effect against the NS2 autoprotease, identical concentrations were tested against the J4 NS2-3 refolding reaction. Again, compound **1** showed a dose responsive effect with an EC_{50} of 79 µM (Figure 4.4B). Of interest, in the absence of the 3-chloro group of the benzene ring (**2**) activity was abolished up to 300 µM (Table 4.1, Figure 4.2)

Along with removal of the 3-chloro substituent from the benzene (**2**), introduction of 3, 4-dimethyl groups (**5**) reduced activity to below tested levels. The presence of a 4-bromo substituent reduced activity (**3**) as did a 3-chloro, 4-bromo substituted group (**4**).

Compounds **7** and **8**, with a 4-amino-phenyloxy or 4-nitro-phenyloxy substituent respectively, showed some activity but to a lesser extent to that shown by **1**, though a 3-phenyloxy substituent did enhance activity (**6**). A larger substituent containing a N-Boc protected piperidinileoxy at the 4-position and a 3-chloro substituent (**10**) showed similar activity to that shown by **1**.

The most potent inhibitor was a 4-benzyloxy, 3-chloro substituted system (**9**) yielding an EC_{50} of 55 µM. A bulky sulphonamide-based side chain at the *para* position (**11**) had reduced activity relative to **1**, though in the absence of a 3-chloro substituent further SAR is required to determine the precise effect of this bulky sidechain. Integrity of all compounds in Table 4.1 was confirmed by mass spectrometry (Appendix I).

The observation that activity varies between the different epoxide based compounds suggested that features other than the epoxide warhead alone are required for activity against the NS2 autoprotease. Additional contacts formed between the R_1 group of this range of compounds could facilitate binding to NS2-3 in a favourable conformation to allow opening of

the epoxide by the catalytic thiol to form a covalent interaction. In agreement with this the epoxide warhead alone (**12**) showed no activity against autoproteolysis when R_1 was a primary alcohol (Table 4.1,Figure 4.2). It was also noted that the natural product cerulenin, based around a similar epoxide warhead with an extended hydrocarbon chain, showed no activity against NS2-3 autoproteolysis (Figure 4.5A).

In addition to greater specificity than other alkylating agents such as TPCK, epoxide warhead based inhibitors have been shown to exhibit stability under physiological conditions and show little reactivity to free thiols, an observation previously made for the current range of compounds (Craig Avery, PhD Thesis). As such epoxide based protease inhibitors have more pharmaceutical applications than other mechanism-based inhibitors (Powers et al., 2002).

In order to test whether epoxide based inhibitors of the NS2 autoprotease could mediate an effect against NS2-dependent HCV genome replication, it was first necessary to confirm they were not cytotoxic within the concentration ranges active *in vitro*. The most potent compounds had an $EC_{50} < 100$ µM against NS2-3 (**1**, **6**, **7**, **9** and **10**). Compounds **1**-**12** were therefore analysed for toxicity up to 100 µM. Huh7 cells were treated with a range of concentrations from 1 nM to 100 µM and after 72 hours cellular metabolism was quantified by MTT assay (Figure 4.3) and data plotted to calculate CC_{50} (Table 4.1). Compounds **1**-**12** showed a range of toxicity profiles, with no correlation to *in vitro* activity against NS2-3.

Table 4.1: Activity and toxicity analysis of epoxide based inhibitors of the NS2 autoprotease.

EC$_{50}$ against JFH1 NS2-3 refolding reaction is calculated from data in Figure 4.2.
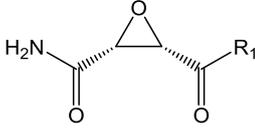
Compound #	R$_1$	NS2-3 refolding EC$_{50}$ (µM)	Huh7 CC$_{50}$ (µM)
1		93	>100
2		>300	>100
3		>200	17
4		>200	2
5		>300	>100
6		70	42
7		96	>100
8		119	62
9		55	12.6
10		98.5	>100
11		>200	>100
12		>300	>100
cerulenin		>300	31.9
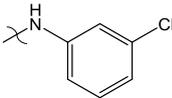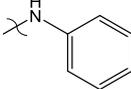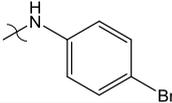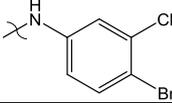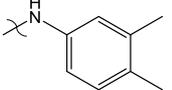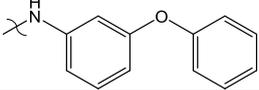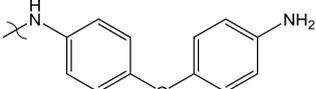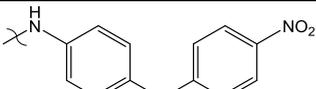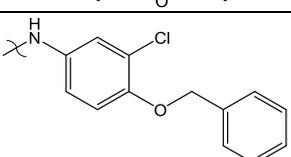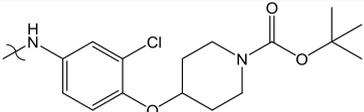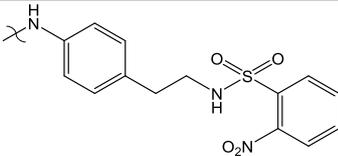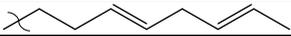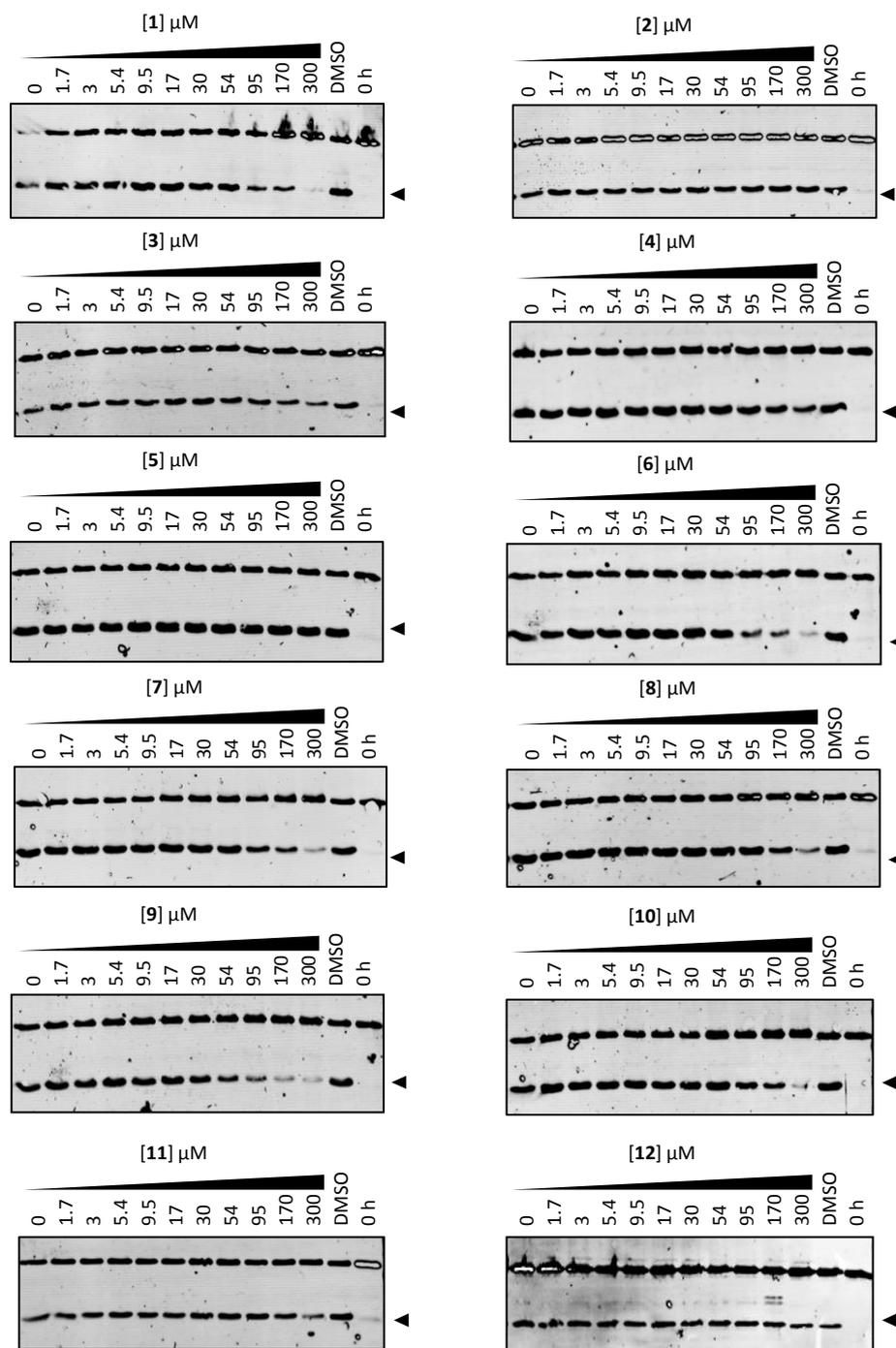
Figure 4.2: Activity of epoxide based compounds on *in vitro* NS2 autoprotease activity

Refolding reactions treated with indicated concentrations of compounds 1-12 and analysed by western blot. Quantified NS3-FLAG is indicated by the arrowhead. Calculated EC50 values are reported in Table 4.1.
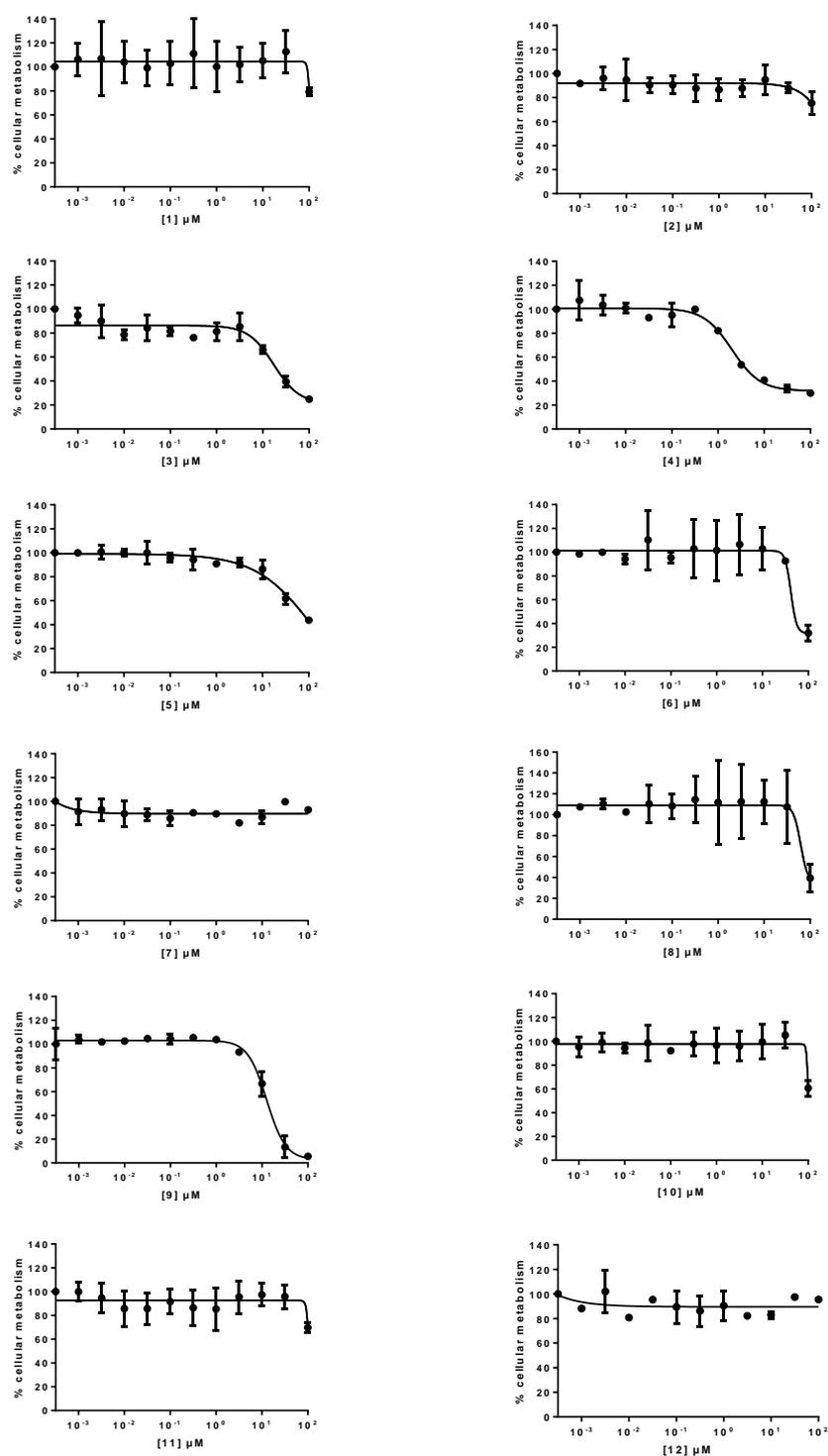
Figure 4.3: Cytotoxicity analysis of epoxide based compounds.

Huh7 cells were treated with indicated concentrations of compounds 1-12 for 72 hours before monitoring cellular metabolism by MTT assay. Data is normalised to DMSO control (plotted as $10^{-3.5}$ μM), n=2. Calculated CC_{50} values are reported in Table 4.1.
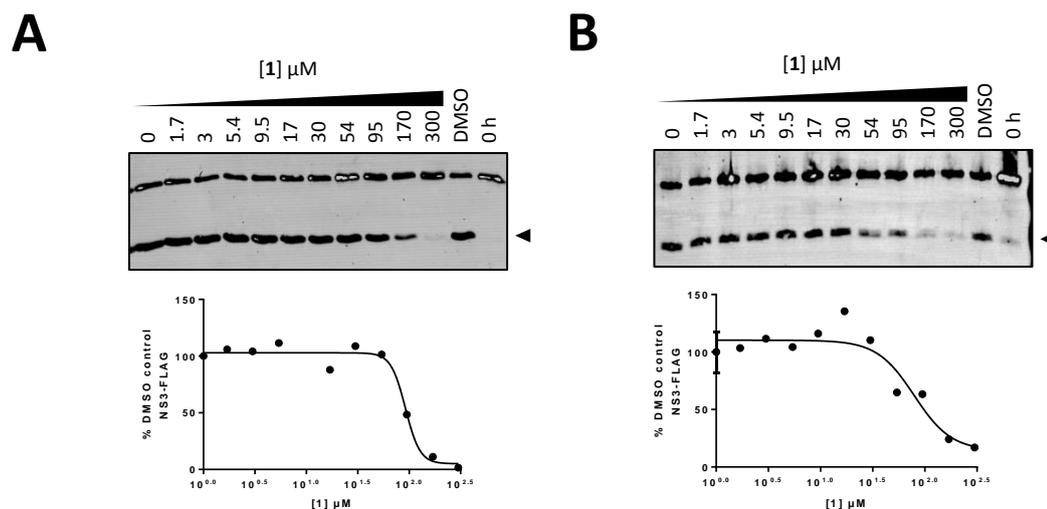
Figure 4.4: Pan-genotype activity of an epoxide warhead against the NS2 autoprotease.

*A) Top; JFH1 NS2-3 refolding reaction treated with indicated concentrations of an epoxide warhead (**1**) and analysed by western blot. Bottom; Quantification of NS3-FLAG (arrowhead) for calculation of $EC_{50} = 92$ μM. B) Identical experiment as in A) using J4 NS2-3, $EC_{50} = 79$ μM.*
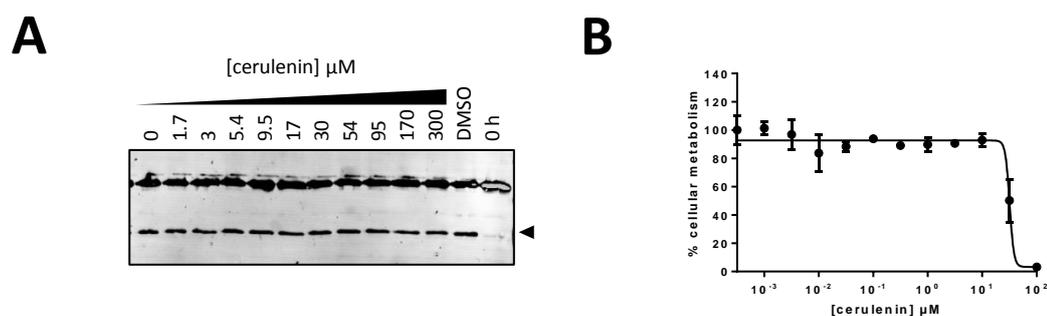
Figure 4.5: Effect of cerulenin on the NS2 autoprotease and cytotoxicity analysis.

A) Indicated concentrations of the natural product cerulenin tested in the JFH1 NS2-3 refolding assay and analysed by western blot. Quantified NS3-FLAG is indicated by the arrowhead. Calculated EC_{50} is reported in Table 4.1. B) Huh7 cells treated with indicated concentrations of cerulenin for 72 hours for toxicity analysis. Calculated CC_{50} is shown in Table 4.1.

137

Of the compounds showing activity *in vitro* only **1**, **7** and **10** lacked significant toxic effects on Huh7 cells up to 100 µM. Of these, compound **1** showed the greatest activity *in vitro*. Additionally, the lack of toxicity observed with compound **2**, which differs from **1** by a single atom, a 3-chloro group on the benzene ring, allowed for a comparative negative control as **2** showed no activity against the NS2-3 refolding reaction.

Compounds **1** and **2** were therefore tested against the SGR-feo-JFH1 [NS2-5B] cell line described in Section 3.2.2 at approximately the *in vitro* EC_{50}; 90 µM. To explore whether any potential effects were mediated against NS2, compounds were tested in parallel against SGR-feo-JFH1 [NS3-5B]. In all cases the luciferase signal after 48 hours was compared to analysis of cellular metabolism in the same cell lines. As a positive control, and to ensure similar sensitivity to DAA mediated inhibition between the two cell lines, both were treated with the NS3 protease inhibitor Telaprevir at an EC_{50} concentration; 150 nM.

Addition of compound **1** (90 µM) to SGR-feo-JFH1 [NS2-5B] gave a significant decrease in luciferase activity to 48.1 ± 10.7% of DMSO control (Figure 4.6A), while analysis by MTT assay demonstrated no significant effect on levels of cellular metabolism (Figure 4.6B). In contrast, addition of compound **2** (90 µM), which lacks activity *in vitro*, had no effect on either luciferase activity from SGR-feo-JFH1 [NS2-5B] or the degree of cellular metabolism (Figure 4.6). Furthermore, the same treatment of SGR-feo-JFH1 [NS3-5B] did not significantly impact luciferase activity after treatment with compound **1** or compound **2**. A slight decrease in the mean luciferase signal (which was not statistically significant) after treatment with compound **1** correlates with a small reduction in levels of cellular metabolism (Figure 4.6). Addition of Telaprevir (150 nM) produced a similar effect in both cell lines. SGR-feo-JFH1 [NS2-5B] luciferase activity was reduced to 53.1 ± 3.6% of DMSO control and SGR-feo-JFH1 [NS3-5B] luciferase activity was reduced to 54.9 ± 5.7% of DMSO control. Cellular metabolism of both cell lines upon Telaprevir treatment was unaffected (Figure 4.6).

Compound **1** therefore appears to block NS2-NS3 processing *in vitro* and inhibit NS2-dependent genome replication in cell culture at similar concentrations. The reduced luciferase activity upon treatment of SGR-feo-JFH1 [NS2-5B] with **1** is not related to any effects on cellular metabolism and appears specific to a SGR where genome replication is dependent on the NS2 autoprotease, as no significant decrease in luciferase activity is observed upon treatment of SGR-feo-JFH1 [NS3-5B] with **1.** Hence compound **1** appears inactive against a SGR that can replicate independently of NS2 autoprotease activity. Both SGR containing cell lines

show comparable sensitivity to Telaprevir suggesting this effect is not a consequence of different replication kinetics or response to DAA treatment.

Furthermore, compound **2**, which lacks only a 3-chloro group on the benzene ring, was inactive against NS2-mediated proteolysis *in vitro* up to 300 µM. A similar trend was observed in this cell based assay. Addition of compound **2** (90 µM) had no effect on the luciferase signal from either SGR-feo-JFH1 [NS2-5B] or SGR-feo-JFH1 [NS3-5B].
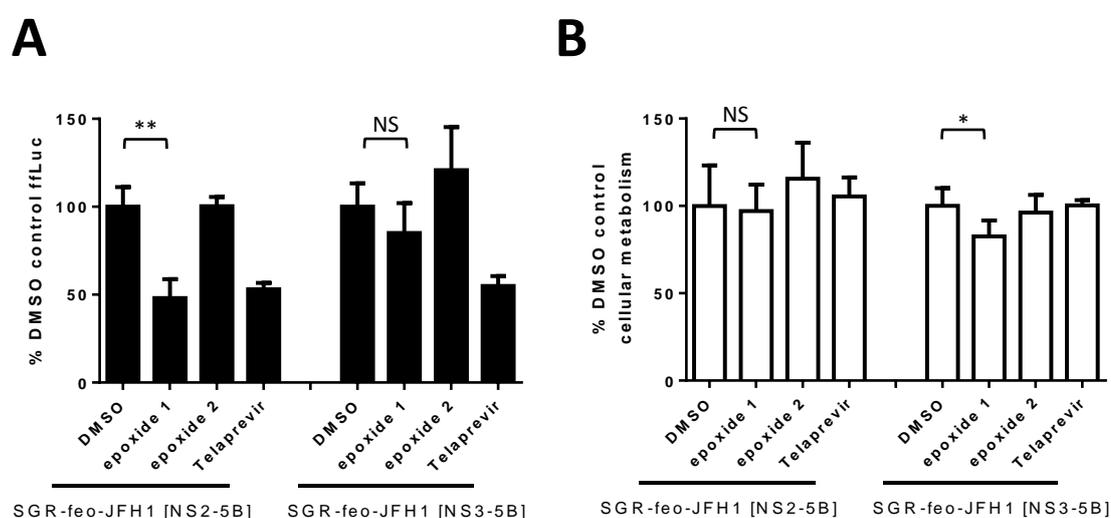
Figure 4.6: Effect of epoxide based compounds 1 and 2 on HCV genome replication.

*A) Treatment of SGR-feo-JFH1 [NS2-5B] and SGR-feo-JFH1 [NS3-5B] with compound **1** (90 µM) and compound **2** (90 µM) for 48 hours. Luciferase activity is normalised to DMSO control. Values and error bars represent the mean and standard deviation of two independent experiments performed in triplicate. Telaprevir (150 nM) was used as a control (one experiment, n=3). B) Toxicity analysis by MTT assay from the same experiments as in A. Statistical significance was determined by unpaired Student's t test (*p < 0.02, **p < 0.0001).*

In vitro analysis of the activity of compound **1** against the NS2 autoprotease indicated increased potency against genotype 1b (J4 NS2-3). To test for a similar trend against genome replication, compound **1** (90 µM) was added to both SGR-feo-JFH1 [NS2-5B] and a genotype 1b SGR cell line; SGR-feo-Con1 [NS2-5B], for 48 hours. The inactive compound **2** was included at the same concentration as a control. As previously observed (Figure 4.6), treatment with compound **1** reduced luciferase activity from SGR-feo-JFH1 [NS2-5B] to 52.2 ± 10.3% of DMSO control with a marginal effect on cellular metabolism. By contrast, addition of compound **1** to SGR-feo-Con1 [NS2-5B] reduced luciferase to 6.0 ± 1.0% of the signal from DMSO control. In both cases, the inactive compound **2** did not reduce luciferase signal. These findings suggest that, in line with *in vitro* observations, compound **1** appears more potent against NS2 mediated proteolysis for genotype 1b than genotype 2a. Taken together these results indicate that a small molecule inhibitor of NS2 mediated proteolysis *in vitro* is capable of specifically blocking NS2-dependent HCV genome replication.
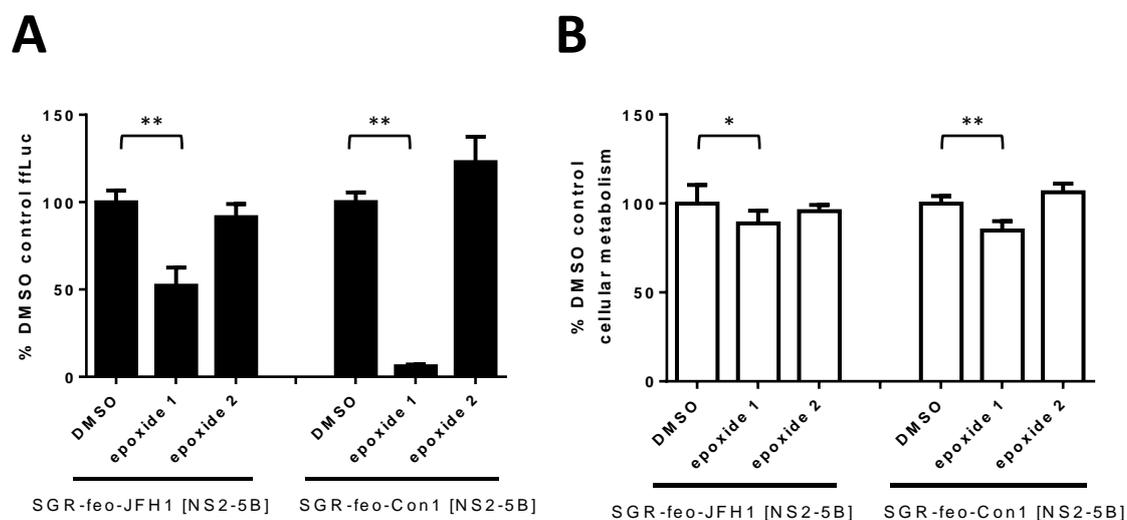
Figure 4.7: Sensitivity of different genotypes to epoxide compounds 1 and 2.

*A) Treatment of SGR-feo-JFH1 [NS2-5B] and SGR-feo-Con1 [NS2-5B] with compound **1** (90 µM) for 48 hours. Luciferase activity is normalised to DMSO control. Values and error bars represent the mean and standard deviation of three independent experiments performed in triplicate. Compound **2** (90 µM) was used as a control (one experiment, n=3). B) Toxicity analysis by MTT assay from the same experiments as in A. Statistical significance was determined by unpaired Student's t test (*p < 0.03, **p < 0.0001).*

4.2.2 CADD to produce an *in silico* enriched screening library

The observation that an epoxide-based compound which blocked NS2-mediated proteolysis *in vitro* inhibited NS2-dependent HCV genome replication suggested that the NS2 autoprotease represents a viable target for anti-virals. The screening assays described in Chapter 3 could therefore be applied to search for a more drug-like small molecule inhibitor of the NS2 autoprotease. To attempt to enrich hit rates, structure-guided CADD was applied to the structural information available for the post-cleaved form of the NS2 protease C-terminal domain (NS2^{pro}) (Lorenz et al., 2006).

Structural coordinates for NS2^{pro} were obtained from the protein data bank (PDB; 2HD0). Water molecules and buffer components such as detergents were removed manually using the Maestro interface (Schrodinger), as were additional subunits, leaving a single NS2^{pro} homodimer. As described by Lorenz *et al.*, the C-terminus of this structure represents the substrate sequence of the NS2 autoprotease, which is observed with P1 and P2 residues bound in the active site and postulated to exert an inhibitory effect. Therefore, to model an accessible active site, C-terminal residues were sequentially removed using the Maestro interface to yield the following NS2^{pro} models; NS2^{pro}ΔP1-P2, NS2^{pro}ΔP1-P5 and NS2^{pro}ΔP1-P10 (Figure 4.8A).

Structure guided vHTS was performed on the above NS2^{pro} models using the program eHITs with a structure file for residues P1 and P2 used to define the target cavity (S1 and S2). eHITs was performed using several virtual libraries, with optimal scoring obtained from the ChemBridge CORE virtual screening library. Therefore top-ranked compounds (as determined by eHITS) from the ChemBridge CORE collection were pooled as part of an *in silico* enriched library.
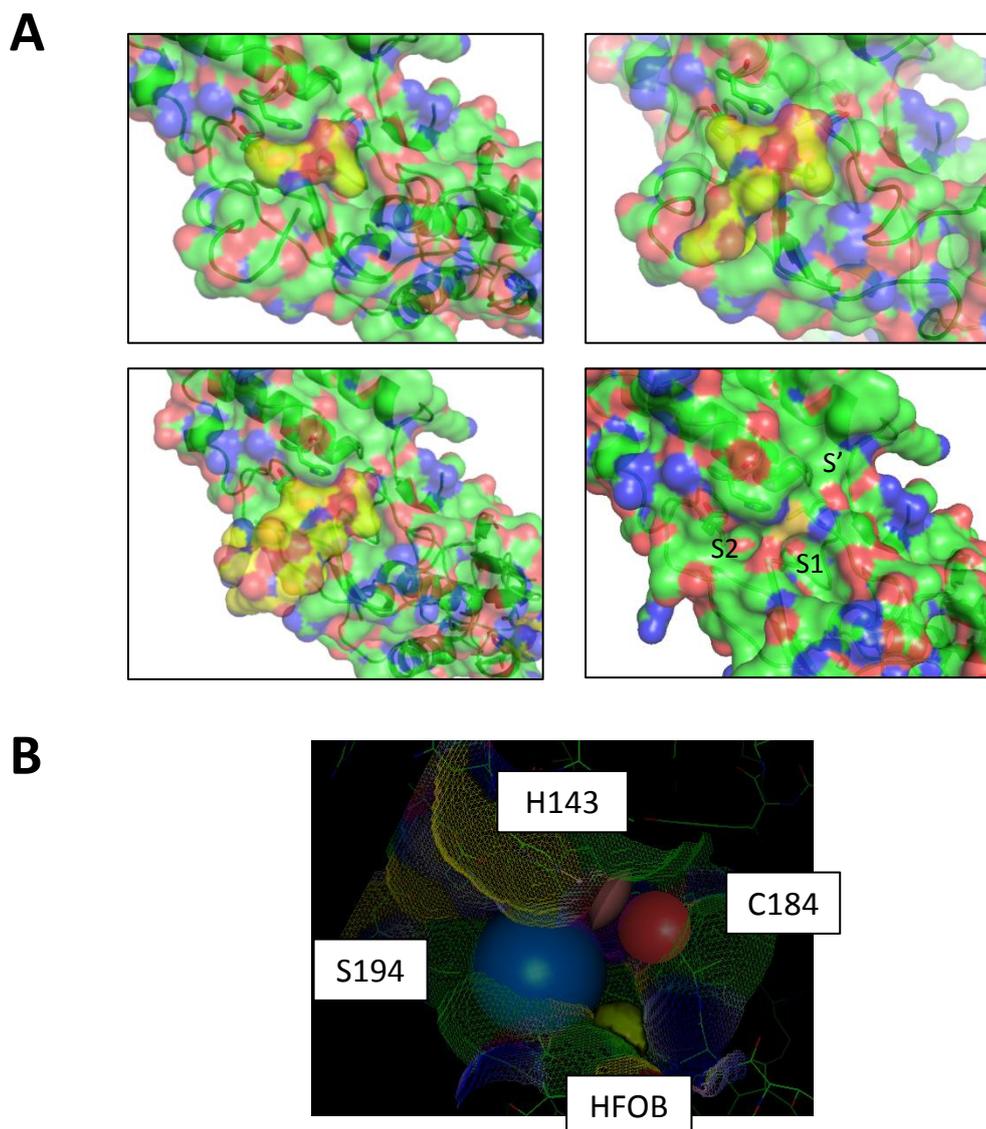
Figure 4.8: NS2^{pro} models used for structure-guided vHTS.

A) Structures of NS2^{pro} models used for vHTS. Top left; NS2^{pro}ΔP1-P2, Top right; NS2^{pro}ΔP1-P5, Bottom left; NS2^{pro}ΔP1-P10. Surface representation is shown with carbon atoms in green, oxygen atoms in red and nitrogen atoms in blue. Surface representation of removed substrate sequence residues are shown with yellow carbons. Bottom right; NS2^{pro}ΔP1-P10 with S1, S2 and S' sites indicated. B) SPROUT analysis of the S1 cavity of the NS2^{pro} active site. Selected hydrogen bond donors are shown in red, selected hydrogen bond acceptors in blue and selected hydrophobic regions (HFOB) in green.

The *in silico* enriched library was also complemented with small molecules targeted to certain 'target sites' using a combination of *de novo* design (SPROUT) and structural and chemical similarity searches (ROCS). Identification of key 'target sites' within the NS2[pro] active site focussed on the S1 site using the SPROUT program. The selected sites comprised two hydrogen bond donors; the backbone nitrogen of C184 and a H143 sidechain nitrogen. Both of these residues form part of the catalytic triad and comparison to other protease crystal structures suggested they occupy a physiologically relevant orientation in the NS2[pro] structure (Figure 1.9), with the backbone nitrogen of C184 also postulated to contribute to the NS2 autoprotease oxyanion hole. In addition, the C-terminal residue L217 of NS2[pro] is proposed to produce an inhibitory effect by sitting in the active site following proteolysis (Lorenz et al., 2006). Within the NS2[pro] structure the carboxylate of L217 forms a hydrogen bond contact with the backbone nitrogen of C184 (Figure 1.8B). A hydrophobic pocket (HFOB) was also identified within the S1 site. In the NS2[pro] structure this site is occupied by the hydrophobic side chain of L217. Additionally, the backbone carboxylate of S194 was identified as a hydrogen bond acceptor.

Appropriate 'starting templates' were docked to each target site using SPROUT before connecting with a variety of 'spacer templates' to produce numerous scaffolds. Relatively simple scaffolds making the required contacts were selected and the pose of these scaffolds were used as the input for a ROCS structural similarity search using a number of virtual libraries, to identify a pool of molecules predicted to be capable of forming some or all of the key contacts present in the initial scaffold. Optimal similarity scoring was also achieved with the ChemBridge CORE virtual library, unsurprising as this virtual library contains the highest number of molecules (> 5.2×10^5).

Top ranked hits from the ChemBridge CORE screening library that scored well in either the eHITs vHTS, or in terms of structural similarity to SPROUT *de novo* designed molecule(s) (as determined by ROCS), were pooled into an *in silico* enriched library. This library was further pruned to remove molecules with unfavourable solubility characteristics (clogP > 5), to remove duplicates, to remove molecules with an unfavourable binding pose from vHTS docking and to maximise chemical diversity within the context of a library of around 200 molecules. The resulting *in silico* enriched library consisted of 213 compounds and was ordered as powder in 96 well format from ChemBridge Corporation. Compounds were resuspended in DMSO solvent to a concentration of 40 mM for use in subsequent screening assays.

4.2.3 Screening for inhibitors of the NS2 autoprotease.

The *in silico* enriched screening library described in Section 4.2.2 underwent biological testing against both *in vitro* NS2-NS3 processing and for the ability to specifically perturb NS2-dependent HCV genome replication using the screening assays optimised in Chapter 3.

Compounds were screened for the ability to block NS2-dependent HCV replication using the medium-high throughout cell based screening assay described in Section 3.2.2. This involved testing each compound in parallel against a SGR cell line dependent on the NS2 autoprotease (NS2-5B) and a negative control cell line capable of replication independent of NS2 autoprotease activity and entirely lacking the NS2 coding region (NS3-5B). The epoxide based compounds described in Section 4.2.1 appear to inhibit the NS2 autoprotease with different genotypic sensitivity. Therefore, as the library of molecules to be screened had been enriched by vHTS against structural information of a genotype 1 isolate of the NS2 catalytic domain, the biological testing was performed against the genotype 1b SGR cell lines derived from the Con1 isolate. To minimise the variability within the assay compounds were screened at 10 µM, allowing a final DMSO concentration of 0.025%.

Most screening programmes produce a number of false positives. To strengthen the evidence in the event of identification of an inhibitor of the NS2 autoprotease, the *in silico* enriched library was also screened for the ability to block production of NS3-FLAG in the NS2-3 refolding assay. Due to the low throughput nature of the assay, each compound was tested at a single concentration. As this *in vitro* assay can tolerate reasonable levels of DMSO solvent, compounds could be screened at 100 µM (final DMSO 0.75%) in a 16 hour refolding reaction. Reactions were performed in batches of 12 compounds for comparative analysis by SDS-PAGE, with each batch of reactions including a 0 hour and DMSO only control. Again, this screening was performed using a genotype 1b isolate to correlate both with the cell-based SGR assay and the crystal structure utilised for vHTS.

SGR-feo-Con1 [NS2-5B]
SGR-feo-Con1 [NS3-5B]
◀ NS3-FLAG
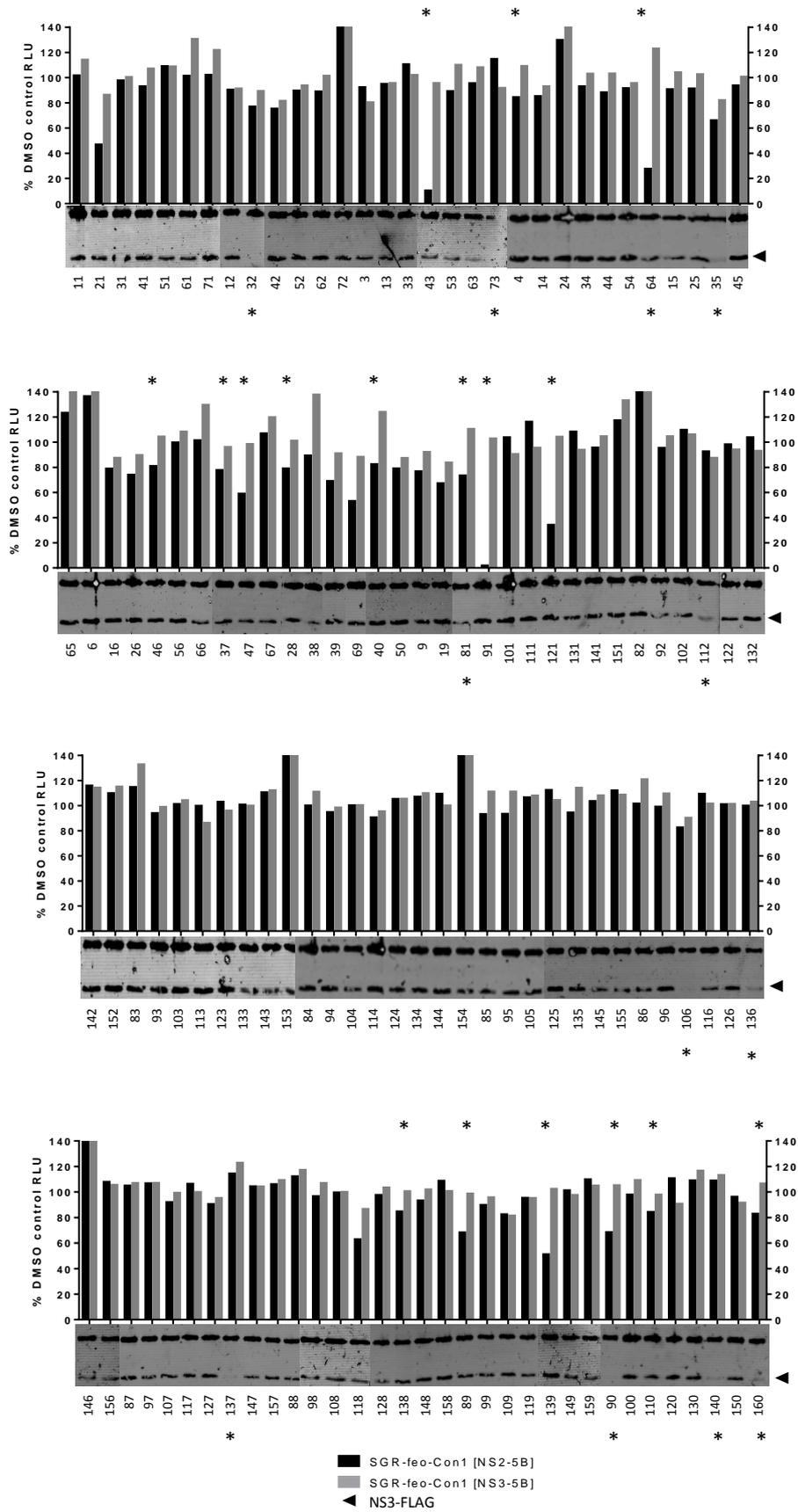
Figure 4.9: Screening of an *in silico* enriched library for inhibitors of the NS2 autoprotease.

Representative data for compounds which had < 20% effect on luciferase signal from the control cell line (SGR-feo-Con1 [NS3-5B]). Luciferase activity from both cell lines is normalised to DMSO control (n=4). The normalised luciferase signal from both cell lines is plotted following a 72 hour treatment with each compound at 10 μM (n=1). Hits (*) were defined as exerting < 10% effect against SGR-feo-Con1 [NS3-5B] while reducing SGR-feo-Con1 [NS2-5B] luciferase to < 85% of DMSO control (> 15% effect). Below, data for a 16 hour treatment of a J4 NS2-3 refolding reaction with 100 μM of the same compound is shown. Reactions were resolved by SDS-PAGE and analysed by western blot using αFLAG antibody. Hits (*) were defined by a visual reduction or complete loss of NS3-FLAG (arrowhead).

Luciferase signal from SGR-feo-Con1 [NS2-5B] and SGR-feo-Con1 [NS3-5B] following parallel treatment with each compound was normalised to DMSO control for each plate, though luciferase signal was similar between the two cell lines. Average luciferase activity (± SD) from DMSO treated SGR-feo-Con1 [NS2-5B] was $8.35 \times 10^5 \pm 3.86 \times 10^4$ RLU, while luciferase activity from SGR-feo-Con1 [NS3-5B] was $8.31 \times 10^5 \pm 0.95 \times 10^4$. Compounds that reduced luciferase signal from the control SGR-feo-Con1 [NS3-5B] cell line to below 80% of DMSO control were excluded from analysis. A representative sample of data for the majority of the remaining compounds is shown in Figure 4.9.

For analysis of cell based SGR screening data, a 'hit' was defined as any compound that did not reduce the SGR-feo-Con1 [NS3-5B] luciferase signal to < 90% of DMSO control but did reduce the SGR-feo-Con1 [NS2-5B] luciferase signal to < 85% of DMSO control (error within the DMSO control for the assay was 2.2 ± 2.0%). The *in vitro* NS2-3 autoproteolysis assay was monitored visually for loss or significant reduction of the NS3-FLAG proteolysis product band. Hits that meet the criteria for each assay are indicated in Figure 4.9 (*).

Screening for specific effects against NS2-dependent HCV replication using the cell based SGR assay identified 17 hits that match the above criteria. 12 compounds were considered to have significantly affected NS2-mediated proteolysis *in vitro*. Of these compounds, 4 were identified as a hit in both the cell based SGR assay and the *in vitro* NS2 autoprotease assay; JS-64, JS-81, JS-90 and JS-160 (Figure 4.9).

Compounds identified as a hit in the cell based SGR assay but not the *in vitro* assay most likely represent false positives. However, as the cell based assay will identify inhibitors of NS2-dependent HCV genome replication, it was considered that these hits could inhibit NS2-dependent replication without directly blocking the NS2 autoprotease. The most likely cause of this could be due to identification of an antagonist of a host factor essential for NS2-NS3 proteolysis in cell culture. A subset of the best hits from the cell-based assay were therefore re-tested for a specific effect against NS2-dependent HCV replication. JS-43, JS-91 and JS-121 were tested against both SGR-feo-Con1 [NS2-5B] and SGR-feo-Con1 [NS3-5B] for 72 hours at 10 µM in duplicate (Figure 4.10A). These compounds failed to reproduce a specific reduction of luciferase from SGR-feo-Con1 [NS2-5B] and therefore most likely represent false positives.

Similarly, hits from the *in vitro* NS2 autoprotease assay that do not impact on NS2-dependent replication could be false positives or could exert a true effect on the NS2 autoprotease but lack sufficient potency in cell culture at 10 µM. Inactivity in cell culture could include an

inability to cross the cell membrane or instability under cell culture conditions. Again, a subset of these molecules were re-tested against the J4 NS2-3 refolding assay at 100 µM, with JS-106 and JS-137 most likely false positives due to subsequent inactivity (Figure 4.10B).

Compounds identified as a hit in both screening assays were re-tested in the same manner. JS-64, JS-81 and JS-90 all appeared inactive when re-tested against SGR-feo-Con1 [NS2-5B] and SGR-feo-Con1 [NS3-5B] in triplicate (10 µM). In contrast, a repeat of the SGR based assay with JS-160 in triplicate (10 µM) reproduced the initial cell-based screening data. JS-160 produced a significant decrease in luciferase activity from SGR-feo-Con1 [NS2-5B] to 67.6 ± 13.2% of DMSO control with no effect on SGR-feo-Con1 [NS3-5B] (98.8 ± 7.2% of DMSO control).

Of the four compounds identified as a hit in both assays, JS-90 and JS-160 had the greatest effect on the *in vitro* NS2-3 refolding assay. J4 NS2-3 refolding reactions were treated with JS-90 and JS-160 to validate these hits. Addition of JS-90 (100 µM) appeared to cause a slight reduction in levels of NS3-FLAG, though an additional repeat revealed no effect against NS2-NS3 proteolysis. However, the activity of JS-160 against the *in vitro* NS2-3 autoprotease assay was reproducible. Addition of JS-160 (100 µM) to the J4 NS2-3 refolding reaction significantly reduced levels of NS3-FLAG from two technical repeats (Figure 4.10B).
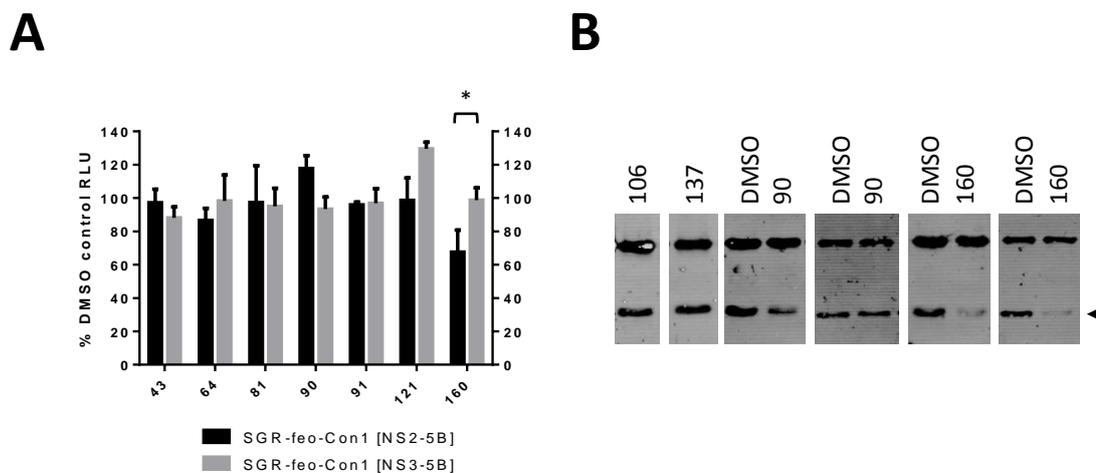
Figure 4.10: Verification of screening 'hits'.

A) A subset of 'hits' from the cell based SGR screen from Figure 4.9 re-tested against SGR-feo-Con1 [NS2-5B] and SGR-feo-Con1 [NS3-5B] for 72 hours at 10 μM. Luciferase activity is normalised to DMSO control. Statistical significance was determined by unpaired Student's t test (* p < 0.006) B) A subset of 'hits' from the in vitro NS2-3 autoprotease assay shown in Figure 4.9 re-tested against a J4 NS2-3 refolding assay at 100 μM. Reactions were analysed by western blot, with NS3-FLAG indicated by the arrowhead.

Analysis of the combination of these two assays for the identification of inhibitors of the NS2 autoprotease revealed that of the 213 compounds within the *in silico* enriched library, 63 (29.6%) were ruled out due to a significant effect against SGR-feo-Con1 [NS3-5B] (most likely cellular toxicity). 4 of these molecules were defined as hits in the *in vitro* NS2-3 assay but were not pursued or validated. Data for a representative sample (containing all of the 'hits') of 127 from the remaining 150 compounds is shown in Figure 4.9. The SGR based assay recorded 16 false positives (7.5%). The NS2-3 refolding assay recorded 11 false positives (5.1%) within the context of compounds not affecting SGR-feo-Con1 [NS3-5B] luciferase activity. When combined, the two assays produced fewer false positives (1.4%). Validation experiments identified a single hit, JS-160, a hit rate of 0.47%. The chemical structure of JS-160 and physiochemical properties are shown in Figure 4.11.

JS-160

ChemBridge CORE # 38490315
(2R, 3R)-3-amino-1'-[(1-methylcyclohexyl) carbonyl]-2,3-dihydrospiro[indene-1.4'-piperidin]-2-ol

$C_{21}H_{30}N_2O_2$
Mw 342
LogP 2.98
Racemic stereochemistry
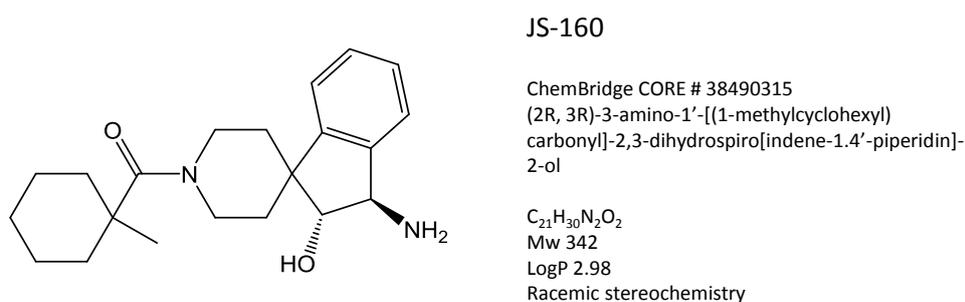
Figure 4.11: Structural and physiochemical properties of JS-160.

Integrity of JS-160 was confirmed by mass spectrometry (Appendix I).

4.3 Discussion

4.3.1 Epoxide warhead based compounds inhibit the NS2 autoprotease *in vitro* and block HCV genome replication

To date no specific small molecule inhibitors of the NS2 autoprotease have been reported which are capable of blocking HCV genome replication by perturbing NS2-NS3 processing. Even inhibiting the NS2 autoprotease *in vitro* has proven difficult, with a wide variety of protease inhibitors and substrate peptides showing no activity or weak activity. Other than zinc chelators, the only small molecules identified as capable of blocking NS2-NS3 processing are highly reactive halomethyl ketones such as TPCK which are unlikely to exhibit much specificity.

A common route for inhibition of proteases is conjugation of a reactive electrophilic warhead to a substrate-derived peptide to form a mechanism-based covalent inhibitor. Of the most common electrophilic warheads, epoxides arguably offer the greatest degree of selectivity based on the requirement of a non-covalent binding pose at the active site to allow opening of the epoxide. Due to the increased specificity and reduced reactivity of epoxide-derived mechanism-based inhibitors they have a greater range of pharmaceutical applications and as such an epoxide based inhibitor of the NS2 autoprotease is more likely to be capable of exerting an effect against NS2-dependent genome replication within a cell-based system.

Mechanism-based inhibitors conjugated to NS2 substrate peptides have previously been reported to show no activity (Pallaoro et al., 2001), and a range of NS2 substrate peptides and different peptide-conjugated electrophilic warheads are also inactive. Within the context of the NS2-3 *in vitro* refolding assay similar observations were made, with the peptide-derived, epoxide warhead-based inhibitor E64 showing no inhibitory activity (Figure 4.1). Taken together these findings support that the NS2 autoprotease active site is not amenable to the binding of peptides *in trans*. This could be due to the size of such peptides (or their derivatives), an idea supported by the inactivity against the NS2 autoprotease of the epoxide-based natural product cerulenin which contains an extended hydrocarbon chain.

If substrate-derived peptides are unable to bind into the active site of NS2, alternative backbone scaffolds may be required to incorporate an electrophilic warhead as a mechanism-based inhibitor. In this regard, a series of aromatic groups conjugated to an epoxide warhead show varying activity against the NS2 autoprotease *in vitro*. Variation of the backbone of this series alters activity against NS2-mediated proteolysis, supporting a contribution of the

backbone to a non-covalent binding pose before potential opening of the epoxide to form a covalent interaction. Additional evidence for this comes from the inactivity of the epoxide warhead alone. The SAR reported in Table 4.1 demonstrates that activity does not correlate with size of the backbone scaffold. Additional SAR around this series should be performed to refine the backbone scaffold for both activity against the NS2 autoprotease and for selectivity.

The current range of epoxide based inhibitors offers the first evidence that a small molecule inhibitor of the NS2 autoprotease can exert effects against the virus lifecycle. The ability of compound **1** to reduce luciferase activity upon addition to SGR-feo-JFH1 [NS2-5B] with no significant effect on the levels of cellular metabolism is indicative of a block on genome replication. The proposed mode of action, through blocking essential NS2-mediated NS2-NS3 processing, is supported by the observation that a similar activity is not observed upon treatment of SGR-feo-JFH1 [NS3-5B]. In this setting, where NS3 does not require release from NS2 and hence genome replication does not depend on the NS2 autoprotease, the epoxide based inhibitor did not significantly affect luciferase activity. This observation also confirms that compound **1** does not exert any inhibitory effect on the luciferase reporter.

As further evidence that inhibitors that block the NS2 autoprotease *in vitro* can inhibit HCV genome replication via the NS2 autoprotease, observations made in the NS2-3 refolding assay were also made in the context of the SGR. Firstly, *in vitro* SAR analysis suggested compound **2** exerted no activity against NS2, while toxicity analysis revealed a comparable toxicity profile to compound **1**. Treatment of SGRs with compound **2** within this concentration range did not affect luciferase activity; hence compound **2** also appears inactive against NS2-dependent genome replication. Secondly, analysis of the activity of compound **1** *in vitro* demonstrated an enhanced inhibition of genotype 1b NS2 (J4 NS2-3). Again, a similar observation was made in the context of the SGR, where a stable cell line containing a SGR derived from a genotype 1b isolate (Con1) showed a greater reduction in luciferase activity upon treatment with compound **1**. The correlation between *in vitro* activity against the NS2 autoprotease and activity against NS2-dependent genome replication supports a mode of action whereby genome replication is blocked by inhibiting the NS2 autoprotease. The fact that blocking the NS2 autoprotease can reduce reporter signal adds additional evidence that perturbing the action of the NS2 autoprotease with a small molecule inhibitor is a viable route to exert an anti-viral effect.

4.3.2 Identification of JS-160 as a lead-like small molecule inhibitor of the NS2 autoprotease

by screening an *in silico* enriched library.

To attempt to identify a novel, non-covalent small molecule inhibitor with improved 'lead-like' properties a library of compounds was screened for activity against the NS2 autoprotease. With the aim of improving hit rates, structure-guided CADD was applied to a series of models of the NS2 autoprotease derived from the post-cleaved NS2^{pro} crystal structure (Lorenz et al., 2006). These models lacked sequential residues from the C-terminus as these residues comprise the substrate sequence for the NS2 autoprotease and preclude the active site. Removal of these residues yielded a relatively shallow active site cavity with defined S1, S2 and S' sites (Figure 4.8B). These sites were targeted by vHTS using eHITs, with top-ranked molecules pooled into an enriched libarary. Additionally the *de novo* design program SPROUT was implemented to identify good target sites within the active site and to build simple scaffolds predicted to be capable of forming contacts with these sites. Through chemical and structural similarity searches using ROCS, commercially available molecules were selected which were also predicted to form such contacts.

Biological testing of this *in silico* enriched library identified a single confirmed hit. The validated hit, JS-160, was identified in both the cell-based and *in vitro* assays, validating each independently. However, the observation of several false positives from each of these assays demonstrated the effectiveness of the use of both assays in parallel to rapidly eliminate the majority of false positives. The true hit rate following completion of the assay was 0.47% (1/213), a fairly typical hit rate for a CADD screening programme and higher than the usual hit rate for HTS (Section 1.9).

The NS2-3 refolding assay is clearly not amenable to a more high-throughput system in its current form. Due to the nature of the endpoint of the assay (western blot), compounds could only be analysed 12 at a time with each batch requiring controls and refolding assays having to be initiated manually. By contrast, the cell-based screening assay should be amenable to HTS as the readout requires addition of a luciferase reagent following treatment with compound so as to monitor luciferase activity from two cell lines in parallel. Indeed, the cell based screening assay has been reformatted to 384-well plate with automated addition of compounds before monitoring luciferase activity. This assay, which is amenable to HTS, is currently being employed to screen a library of > 10,000 compounds (Rachel Trowbridge, Richard Foster, University of Leeds). Such an assay may be of use for rapid identification of additional small

molecule inhibitors of the NS2 autoprotease which are capable of blocking NS2-dependent genome replication.

While the hit from the initial assays, JS-160, was validated in both the *in vitro* refolding assay and the cell-based SGR assay, further analysis was performed to confirm anti-viral activity against HCV and support an NS2-mediated mode of action. Results of these experiments will be the subject of Chapter 5.

Chapter 5 - Analysis of NS2 autoprotease inhibitors

5.1 Introduction

The NS2 autoprotease has been shown to be essential for HCV in both tissue culture models and *in vivo* (Jones et al., 2007, Kolykhalov et al., 2000) leading to the suggestion that pharmalogical inhibition of NS2-mediated polyprotein processing is a viable route to exert an anti-viral effect. Data presented in Chapter 4 supports the use of NS2 inhibitors to block genome replication by preventing proteolysis at the NS2-NS3 junction. The lead-like small molecule JS-160 identified from an *in silico* enriched library of compounds represents an attractive starting point for the development of an inhibitor targeting the NS2 autoprotease. While the identification of JS-160 as a 'hit' has been verified, the potency against the NS2 autoprotease and against NS2-dependent genome replication remains to be determined.

JS-160 causes only a partial reduction in luciferase activity against SGR-feo-Con1 [NS2-5B] at 10 µM (Figure 4.10) indicating activity in the mid to high µM range against HCV genome replication. The potency of this range of NS2 autoprotease inhibitors would therefore most likely require significant improvement before consideration as a drug molecule. In order to improve activity through a SAR programme it is important first to gain an understanding of the features of JS-160 which are essential for or contribute to activity against the NS2 autoprotease. The first section of this chapter (Section 5.2.1) focuses upon the determination of the potency of JS-160 against the NS2 autoprotease using the *in vitro* refolding assay. In addition an analysis of structurally similar small molecules is performed *in vitro* to begin to map the requirements for activity of JS-160. Structurally similar compounds were obtained both by the synthesis of novel molecules and as 'SAR by inventory' using commercially available compounds.

Section 5.2.2 will explore the activity of JS-160 and its analogues against NS2-dependent genome replication. In particular the confirmation of a therapeutic window is required to confirm a concentration range that inhibits reporter signal from a SGR containing NS2-5B in the absence of cytotoxic effects.

Subsequent experiments will explore the mode of action of JS-160. The NS2 autoprotease performs an essential proteolysis reaction during HCV polyprotein processing which if blocked is likely to negatively regulate the activity of the NS3 protease. As such a NS2 autoprotease inhibitor would be expected to act by a similar mode of action to the established NS3 protease inhibitors, albeit through an alternative molecular target. The profile of SGR inhibition upon

treatment with JS-160 should therefore be compared to the profile of an established NS3 protease inhibitor.

While SGR-based systems are well established for the identification and analysis of DAAs targeting HCV genome replication they do not represent the complete virus lifecycle. To confirm an anti-viral effect of inhibitors of the NS2 autoprotease, JS-160 was tested against infectious HCV in a cell culture model (HCVcc). These studies will be the subject of Section 5.2.3, where the ability of NS2 autoprotease inhibitors to reduce the number of infectious HCV virions will be assessed. Furthermore, the use of HCVcc allows the effects of NS2 inhibitors on different stages of the virus lifecycle to be analysed.

5.2 Results

5.2.1 *In vitro* analysis of small molecule inhibitors of NS2 mediated proteolysis

To examine the potency of JS-160 against NS2-3 autoprocessing, the compound was added to refolding reactions at a range of concentrations from 1.7 - 300 µM before quantification of the NS3-FLAG proteolysis product by western blot. All reactions contained standard levels of DMSO vector (0.75%), and DMSO only and 0 hour controls were included. In line with an inhibitory effect of JS-160 against the J4 NS2-3 refolding assay (Figure 4.10B), treatment with JS-160 reduced levels of NS3-FLAG proteolysis product in a dose responsive manner above 54 µM, though inhibition was not complete at 300 µM (Figure 5.1).
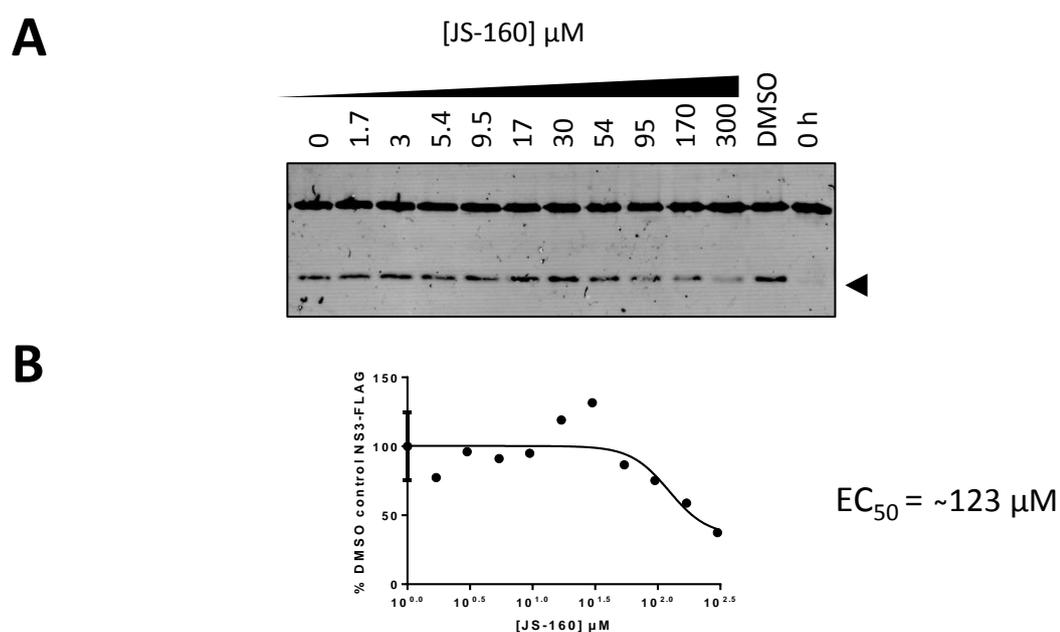

Figure 5.1: Activity of JS-160 against J4 NS2-3 refolding.

A) J4 NS2-3 refolding reactions treated with indicated concentrations of JS-160 alongside controls before analysis by western blot. NS3-FLAG proteolysis product is indicated by the arrowhead. B) Quantification of NS3-FLAG from A. Proteolysis is reduced in a dose-responsive manner above 95 µM yielding an EC_{50} = ~123 µM.

Previous observations with the epoxide based inhibitors described in Section 4.2.1 suggested differing sensitivity between genotype 1b and genotype 2a isolates. Additionally, as the JFH1 NS2-3 refolding reaction is more efficient it provides a more accurate assay to test the potency of different compounds. To test whether JS-160 was capable of inhibiting JFH1 NS2-NS3 proteolysis, the same concentrations were added to JFH1 NS2-3 refolding reactions. As shown in Figure 5.2, JS-160 appears to be more potent against the JFH1 NS2 autoprotease. To confirm this observation, three technical repeats were performed. Each of these demonstrated a dose-responsive reduction in NS3-FLAG, yielding an average EC_{50} of 44.6 ± 6.2 µM (Figure 5.2). In line with the epoxide-warhead based compound **1**, JS-160 shows different sensitivity between genotype 1b and genotype 2a isolates. However, while compound **1** showed enhanced potency against genotype 1b, JS-160 appears more active against genotype 2a JFH1.

To confirm the activity of JS-160, a synthetic route was devised and undertaken. Intermediates of this synthetic route were also tested for activity against the NS2 autoprotease to define the requirements for activity within JS-160 by SAR analysis. Additional SAR was performed using molecules structurally similar to JS-160 which were commercially available from the same source as JS-160. Compounds tested as part of a SAR analysis of JS-160 are shown in Table 5.1.
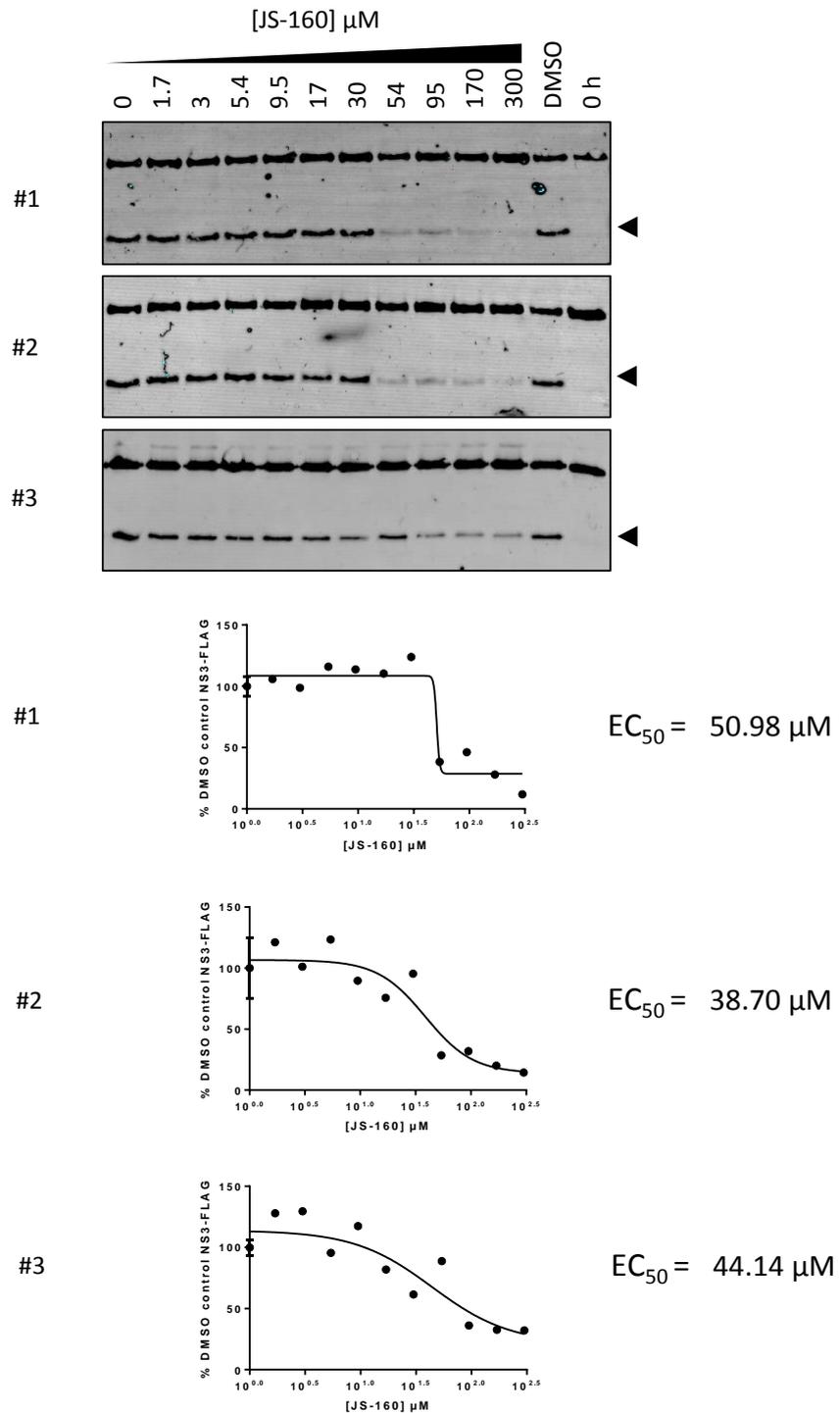
Figure 5.2: Activity of JS-160 against JFH1 NS2-3 autoprocessing.

JFH1 NS2-3 refolding reactions treated with indicated concentrations of JS-160 and analysed by western blot. Three technical repeats are shown. NS3-FLAG (arrowhead) was quantified to calculate $EC_{50} = 44.6 \pm 6.2\ \mu M$ (AVG ± SD, n=3).

Table 5.1: Structures tested for activity against the NS2 autoprotease as part of SAR analysis of JS-160.

Compound	Structure	Source
JS-160		ChemBridge CORE
JS-160-1		In house synthesis
JS-160-2		In house synthesis
JS-160-3		ChemBridge CORE
JS-160-4		ChemBridge CORE
JS-160-5		ChemBridge CORE
JS-160-6		ChemBridge CORE
JS-160-7		ChemBridge CORE
JS-160-8		ChemBridge CORE
JS-160-9		ChemBridge CORE
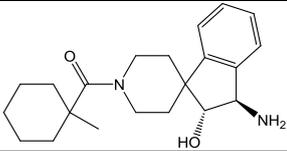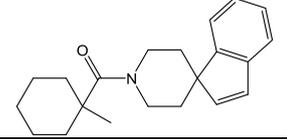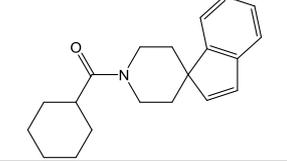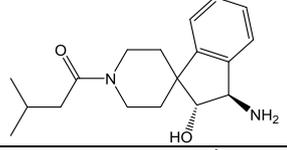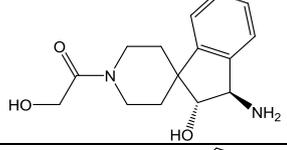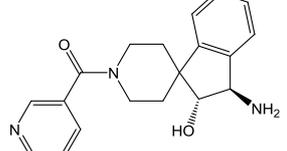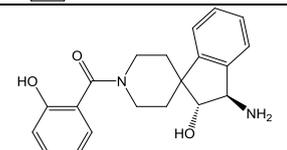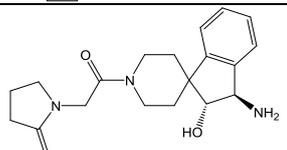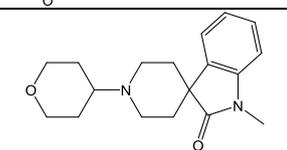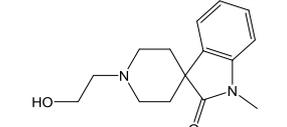
Synthesis of JS-160 was performed by amide coupling of 1-methyl-1-cyclohexanecarboxylic acid (SM-2) and 4-spiroindene-piperidine hydrochloride (SM-1) to yield compound **JS-160-1** (for details of synthetic reactions see Appendix I). JS-160-1 was treated with meta-chloroperbenzoic acid to yield the epoxide **JS-160-1.2**. However, this molecule proved insoluble in all tested solvents. Reaction of JS-160-1.2 with ammonia under heat and pressure did produce a small quantity of **JS-160-1.3** but the low yield, most likely due to the majority of JS-160-1.2 remaining insoluble, prevented purification and complete analysis. Although the presence of JS-160-1.3 was confirmed by mass spectrometry (Appendix I), due to the insoluble nature of JS-160-1.2, synthesis of sufficient JS-160 for complete purification and analysis may require an alternative synthetic route.

To explore the activity of JS-160-1, the same range of concentrations were added to JFH1 NS2-3 refolding reactions. Interestingly, JS-160-1 also caused a dose-responsive reduction in levels of NS3-FLAG. Three technical repeats produced an average EC_{50} with no significant difference from JS-160 (Figure 5.4), suggesting the primary alcohol and amine groups of JS-160 are not essential for activity. The activity of JS-160-1 further indicates that the 1-methyl cyclohexane group of JS-160 may be important for activity. In order to further test the contributions of this group it was replaced by a cyclohexane group. Amide coupling of cyclohexane carboxylic acid to 4-spiroindene-piperidine hydrochloride (SM-1) yielded **JS-160-2** (Figure 5.3).

JS-160-1

JS-160-1.2
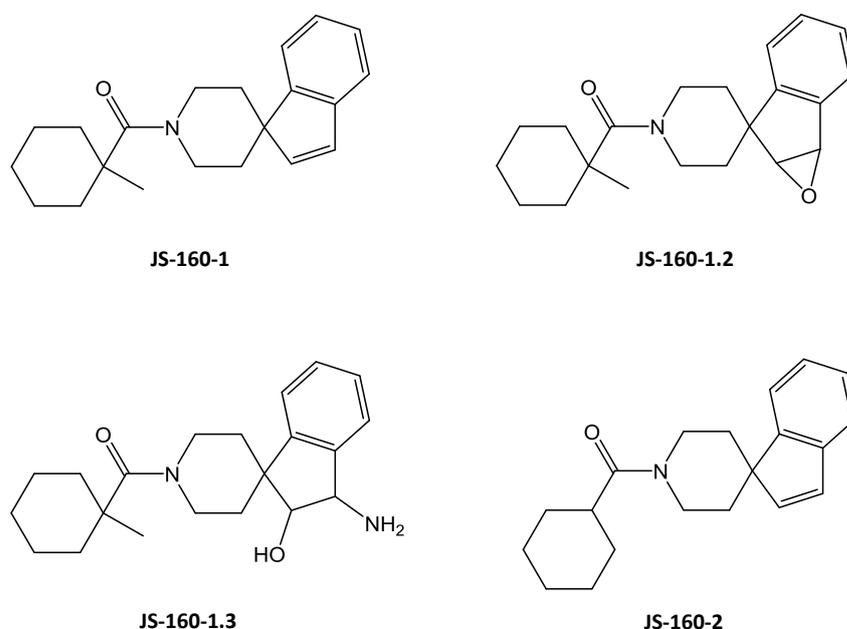
JS-160-1.3

JS-160-2

Figure 5.3: Chemical structures of synthetic products.

For details of synthesis see Appendix I.
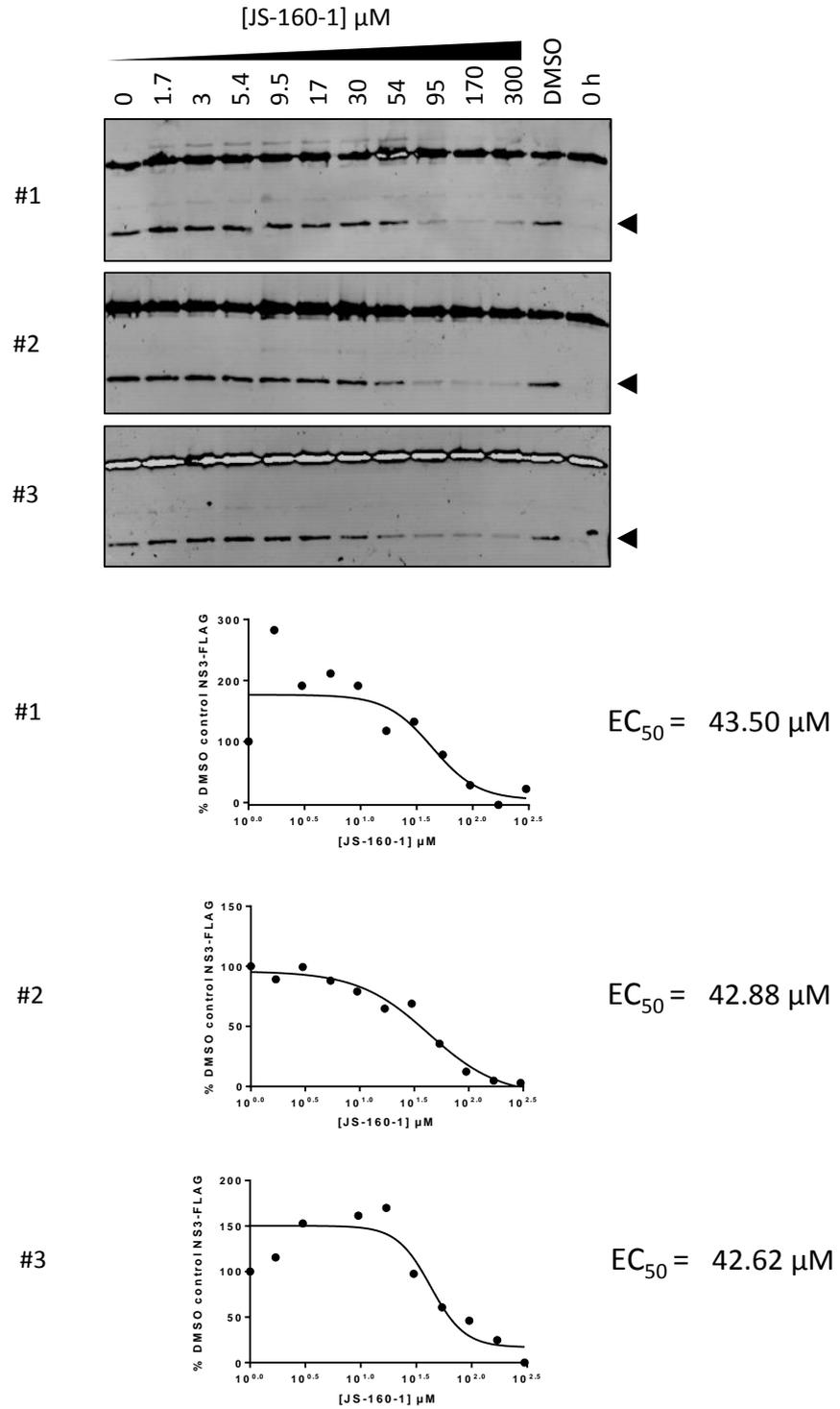
**Figure 5.4: Activity of JS-160-1 against JFH1 NS2-3 autoprocessing.**

JFH1 NS2-3 refolding reactions treated with indicated concentrations of JS-160-1 and analysed by western blot. Three technical repeats are shown. NS3-FLAG (arrowhead) was quantified to calculate EC_{50} = 43.0 ± 0.5 μM (AVG ± SD, n=3).
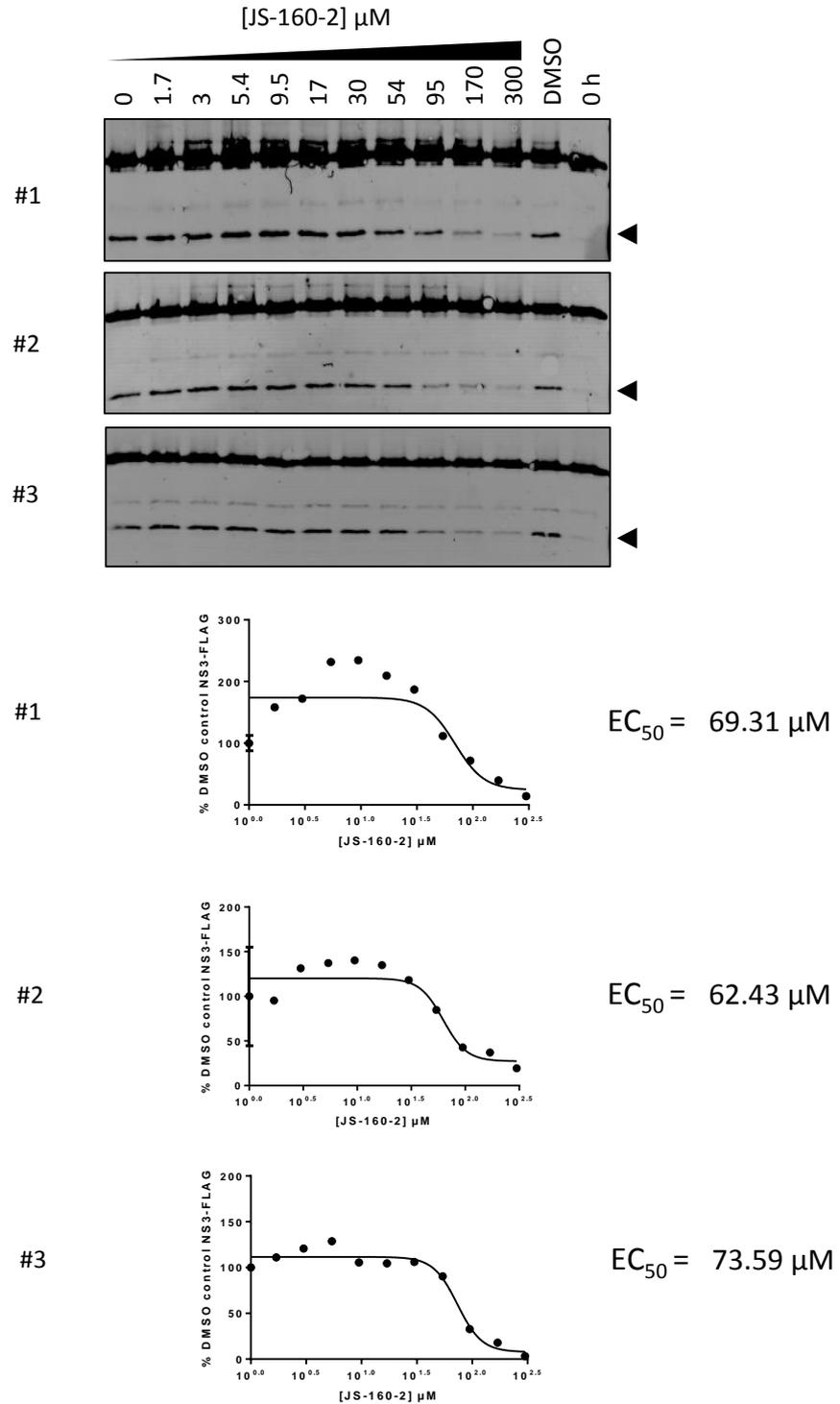
Figure 5.5: Activity of JS-160-2 against JFH1 NS2-3 autoprocessing.

JFH1 NS2-3 refolding reactions treated with indicated concentrations of JS-160-2 and analysed by western blot. Three technical repeats are shown. NS3-FLAG (arrowhead) was quantified to calculate $EC_{50} = 68.4 \pm 5.6\ \mu M$ (AVG ± SD, n=3).

164

As evidenced by three technical repeats, while JS-160-2 did retain activity, it appeared less potent against JFH1 NS2-mediated proteolysis with an average EC_{50} of 68.4 ± 5.6 µM (Figure 5.5). Therefore, while the cyclohexane group of JS-160 is sufficient for activity against the NS2 autoprotease, the 1-methyl group also contributes to activity.

Treatment of JFH1 NS2-3 refolding reactions with the starting materials 4-spiroindene-piperidine hydrochloride (**SM-1**) or 1-methyl-1-cyclohexanecarboxylic acid (**SM-2**) confirmed that while the 1-methyl-cyclohexane is required for activity (Figure 5.6B), it is not sufficient for activity (Figure 5.6A). In addition, the NS2-3 refolding reaction could be treated simultaneously with up to 300 µM SM-1 and 300 µM SM-2 without affecting levels of proteolysis (Figure 5.6C). Yet when these two molecules were coupled together they produce activity with an average EC_{50} of 43 µM (JS-160-1).

To further explore the chemical features of JS-160 important for activity a range of structurally similar compounds were tested. These molecules (JS-160-3 to JS-160-9) are commercially available from the same source as JS-160. Compounds were again tested at a range of concentrations to allow calculation of EC_{50} against the JFH1 NS2-3 refolding reaction. However, none of the tested compounds showed any activity up to 300 µM (Figure 5.7, Figure 5.8). Variation of the 1-methyl cyclohexane group from JS-160, as in JS-160-3 to JS-160-7, abolished all activity. More significant changes (JS-160-8 and JS-160-9) to the structure also showed no activity against the NS2 autoprotease. This SAR analysis adds additional evidence that the 1-methyl cyclohexane moiety is essential for the activity of JS-160. As similar compounds from commercial sources have proven inactive, further SAR will likely rely on the synthesis of novel molecules.

Table 5.2 shows SAR analysis of the current series of compounds for activity against the JFH1 NS2 autoprotease *in vitro*, summarising the data from Figure 5.2, Figure 5.4, Figure 5.5, Figure 5.6, Figure 5.7 and Figure 5.8.
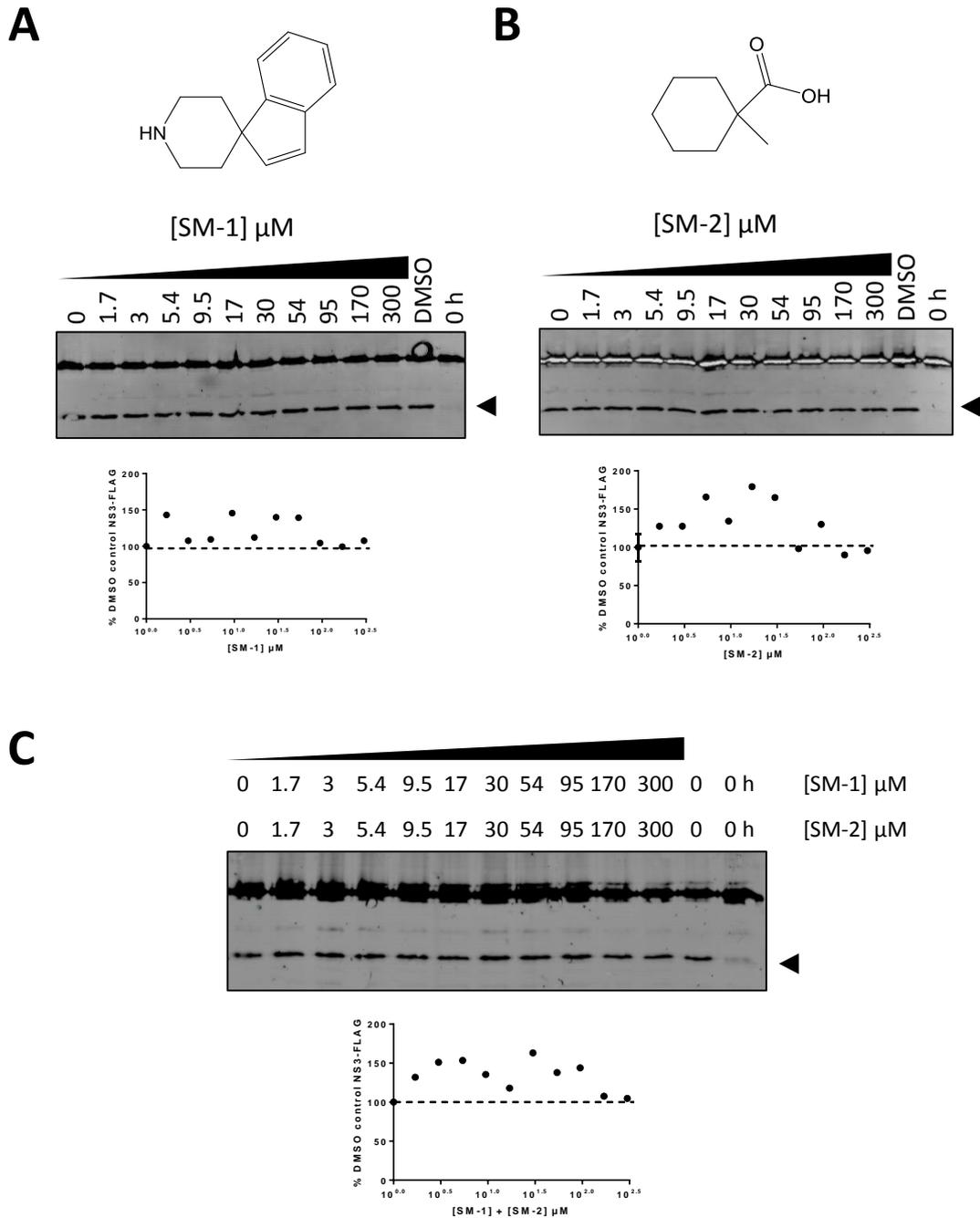
Figure 5.6: Activity of JS-160-1 fragments against JFH1 NS2-3 autoprocessing.

A) Chemical structure of JS-160-1 fragment SM-1. JFH1 NS2-3 refolding reactions treated with indicated concentrations of SM-1 alongside controls were analysed by western blot. Quantified NS3-FLAG (arrowhead) is plotted normalised to DMSO control. B) The same experiment as in A with JS-160-1 fragment SM-2. C) JFH1 NS2-3 refolding reactions treated with indicated concentrations of both SM-1 and SM-2 and analysed by western blot. Quantified NS3-FLAG is normalised to DMSO (1.5%) control.
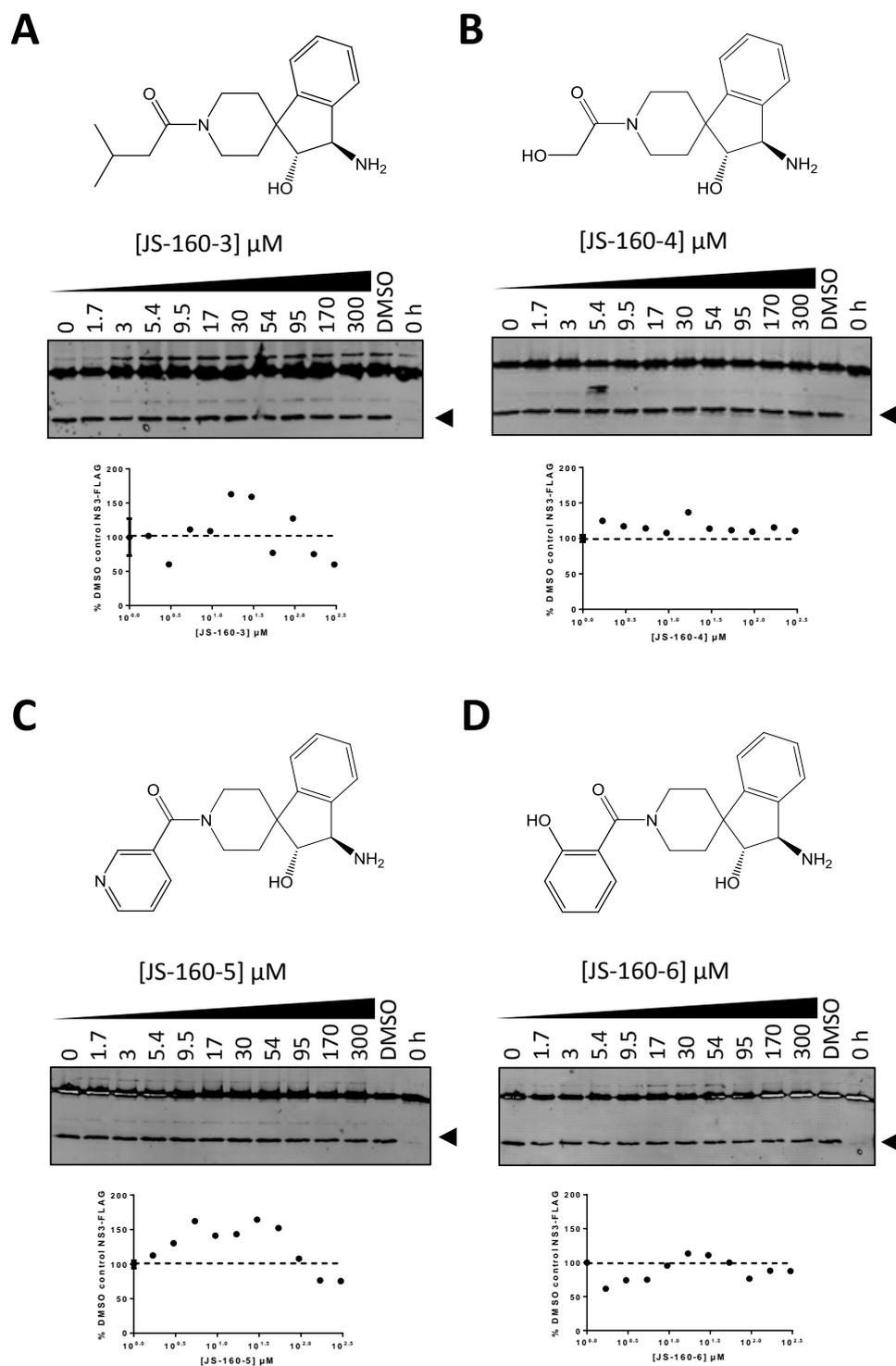
Figure 5.7: Activity of molecules structurally similar to JS-160 on JFH1 NS2-3 autoprocessing.

A) JFH1 NS2-3 refolding reactions treated with indicated concentrations of JS-160 structural analogue JS-160-3 alongside controls. Quantified NS3-FLAG (arrowhead) is plotted below normalised to DMSO control. B) The same experiment as in A with compound JS-160-4. C) The same experiment with JS-160-5. D) The same experiment with JS-160-6.

167
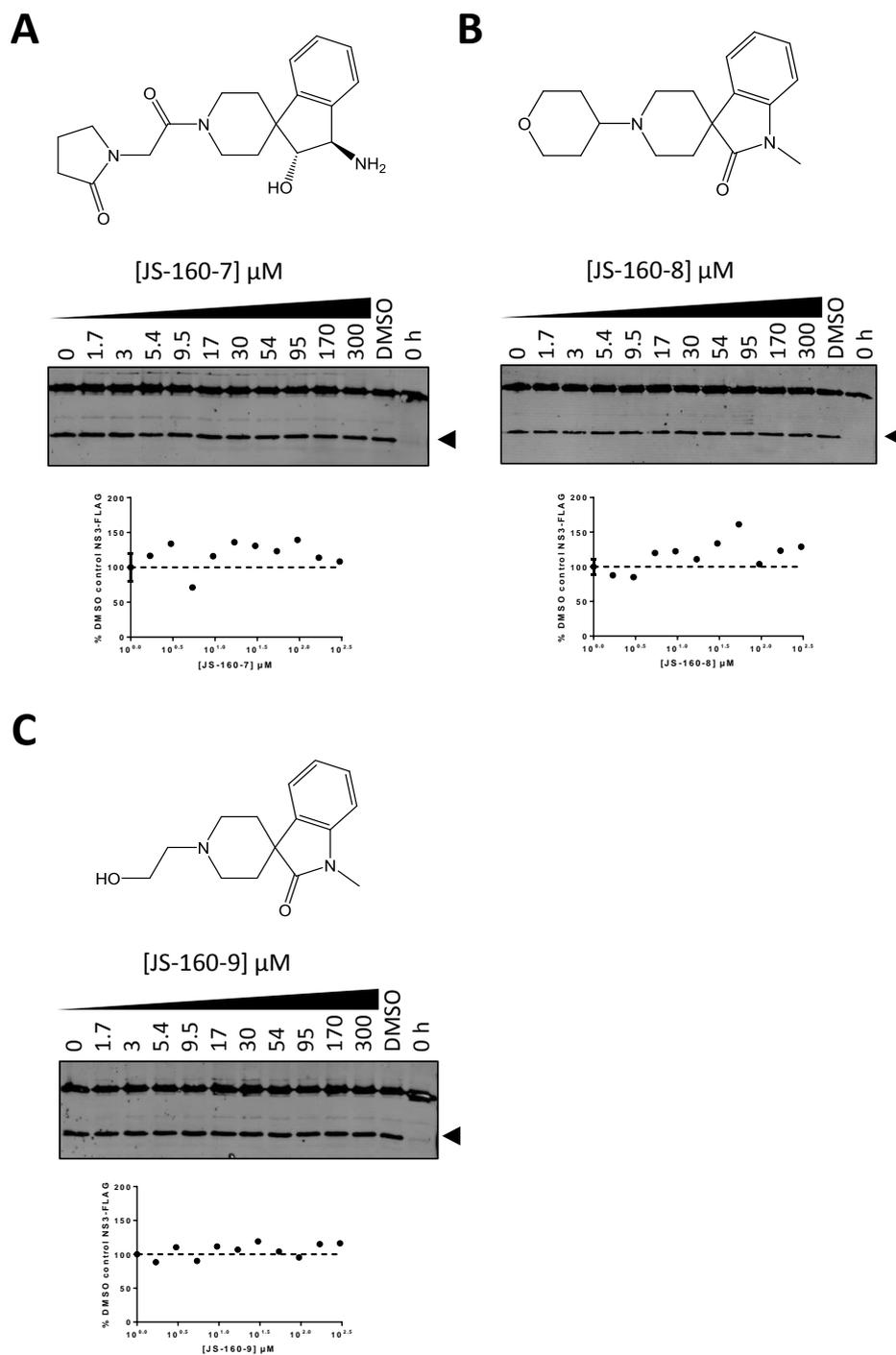
Figure 5.8: Activity of additional molecules structurally similar to JS-160 on JFH1 NS2-3 autoprocessing.

A) JFH1 NS2-3 refolding reactions treated with indicated concentrations of JS-160 structural analogue JS-160-7 alongside controls. Quantified NS3-FLAG (arrowhead) is plotted below normalised to DMSO control. B) The same experiment as in A with compound JS-160-8. C) The same experiment with JS-160-9.

168

Table 5.2: Summary of *in vitro* SAR of JS160. *Average of three technical repeats ± SD.*

Compound	Structure	JFH1 NS2-3 EC$_{50}$ (µM)
JS-160		44.6 ± 6.2 *
JS-160-1		43.0 ± 0.5 *
JS-160-2		68.4 ± 5.6 *
JS-160-3		>300
JS-160-4		>300
JS-160-5		>300
JS-160-6		>300
JS-160-7		>300
JS-160-8		>300
JS-160-9		>300
SM-1		>300
SM-2		>300
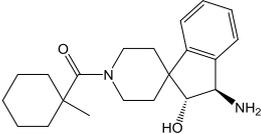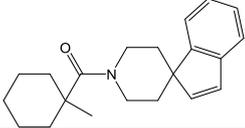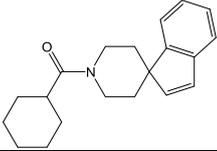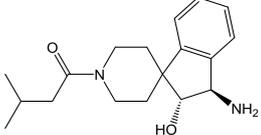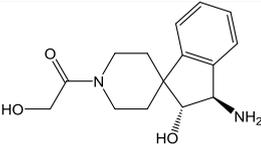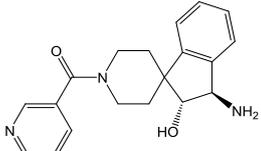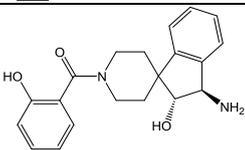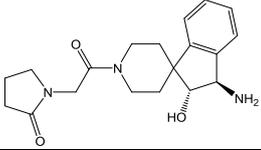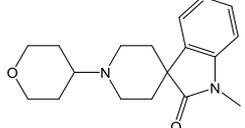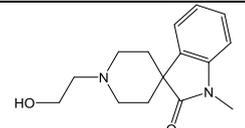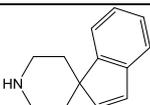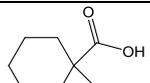
To explore the kinetic profile of NS2 inhibition *in vitro*, JFH1 NS2-3 refolding reactions were treated with an intermediate concentration of JS-160-1 (100 µM) or with DMSO alone and reactions halted every 30 minutes for 4.5 hours. Quantified NS3-FLAG is plotted without normalisation (raw NS3-FLAG intensity), demonstrating reduced levels of the NS3-FLAG proteolysis product. However, JS-160-1 does not appear to have any dramatic effect on the rate of NS2-mediated proteolysis. Both untreated reactions and reactions in the presence of the inhibitor were complete by 3 h (Figure 5.9A).

As an additional control to demonstrate that addition of JS-160-1 blocks NS2-mediated proteolysis rather than affecting detection of the NS3-FLAG product, a range of refolding reactions were allowed to proceed untreated for 16 hours. JS-160-1 was then added to these reactions at a range of concentrations and incubated for 30 minutes before quantifying NS3-FLAG by western blot. In this setting, where NS2-mediated processing of NS2-3 has already reached completion, the addition of JS-160-1 up to 300 µM had no effect on levels of the NS3-FLAG proteolysis product (Figure 5.9B).
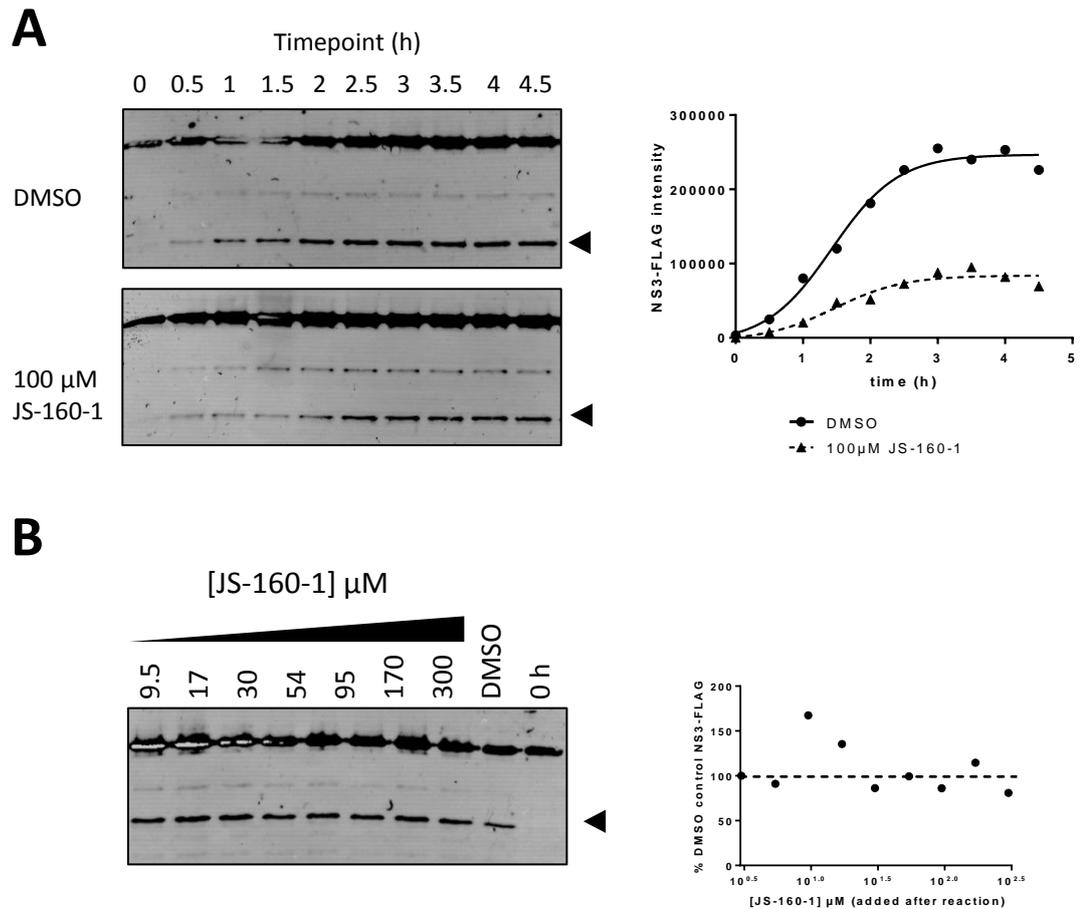
Figure 5.9: Effects of JS-160-1 on the NS2-3 refolding reaction.

A) JFH1 NS2-3 refolding reactions halted at indicated timepoints in the presence or absence of 100 µM JS-160-1 before analysis by western blot. Quantified NS3-FLAG (arrowhead) intensity is plotted. B) JFH1 NS2-3 refolding reactions were allowed to proceed untreated for 16 hours then treated with indicated concentrations of JS-160-1 for a further 30 minutes before quantifying NS3-FLAG by western blot analysis.

5.2.2 Effects of NS2 autoprotease inhibitors against NS2-dependent genome replication

As described in Section 4.3, JS-160 was also identified as a hit in a cell based screening assay. 10 µM JS-160 appeared to have a reproducible effect against luciferase activity from SGR-feo-Con1 [NS2-5B] without affecting the luciferase activity from SGR-feo-Con1 [NS3-5B].

As JS-160 appeared more potent against the JFH1 NS2 autoprotease *in vitro*, effects on the SGR were further evaluated using JFH1 isolates. Such an approach was also necessary as only JFH1 derived genotype 2a isolates can sufficiently reproduce the entire HCV lifecycle in cell culture, hence the effect of JS-160 against infectious virus can only be examined for JFH1.

Before analysing compounds for effects against the SGR, the toxicity profile of key compounds was determined. Huh7.5 cells were treated with a range of concentrations of JS-160 and the inactive derivatives JS-160-3, JS-160-4, JS-160-5, JS-160-6 and JS-160-7. After 48 hours levels of cellular metabolism were monitored by MTT assay. JS-160 showed some effects on cellular metabolism at concentrations greater than 100 µM. A similar profile was observed with the inactive derivative JS-160-3. By contrast, JS-160-4, JS-160-5, JS-160-6 and JS-160-7 showed no significant toxic effects up to 300 µM (Figure 5.10).

These findings suggest that a hydrophobic group in position of the 1-methyl-cyclohexane of JS-160 is associated with some toxic effects in cell culture. However, while the hydrophobic di-methyl groups at this position in JS-160-3 produce a similar toxicity profile, these groups are not sufficient for any activity against the NS2 autoprotease *in vitro* (Table 5.2). As such JS-160-3 may prove useful as a negative control as an inactive compound with a comparable toxicity profile to JS-160 in Huh7.5 cells.

To explore the activity of these compounds against the SGR, a range of concentrations were added to SGR-feo-JFH1 [NS2-5B] for 48 hours before measuring luciferase activity. In parallel, cells were analysed for cytotoxic effects.

Treatment of SGR-feo-JFH1 [NS2-5B] with JS-160 produced a dose-responsive reduction in luciferase activity with an EC_{50} of 43.4 µM (Figure 5.11A). As toxic effects were not apparent at lower concentrations (CC_{50} = > 100 µM) reduced luciferase activity in this range most likely represent activity against genome replication. In contrast, reduced luciferase activity at higher concentrations upon addition of JS-160-3 correlated with toxicity analysis (ffLuc EC_{50} = >100 µM, CC_{50} = >100 µM) (Figure 5.11C). This data supports the use of JS-160-3 as a negative control wherein a similar toxicity profile is observed to that of JS-160, but no additional effect

on replication (as measured by luciferase activity) is apparent, in line with the inactivity of JS-160-3 against the NS2 autoprotease *in vitro*. As additional evidence that the SAR reported in Table 5.2 is observed in the context of HCV replication, treatment of SGR-feo-JFH1 [NS2-5B] with JS-160-4, which exerts no toxic effects below 300 µM, had no effect on luciferase activity (Figure 5.11E).
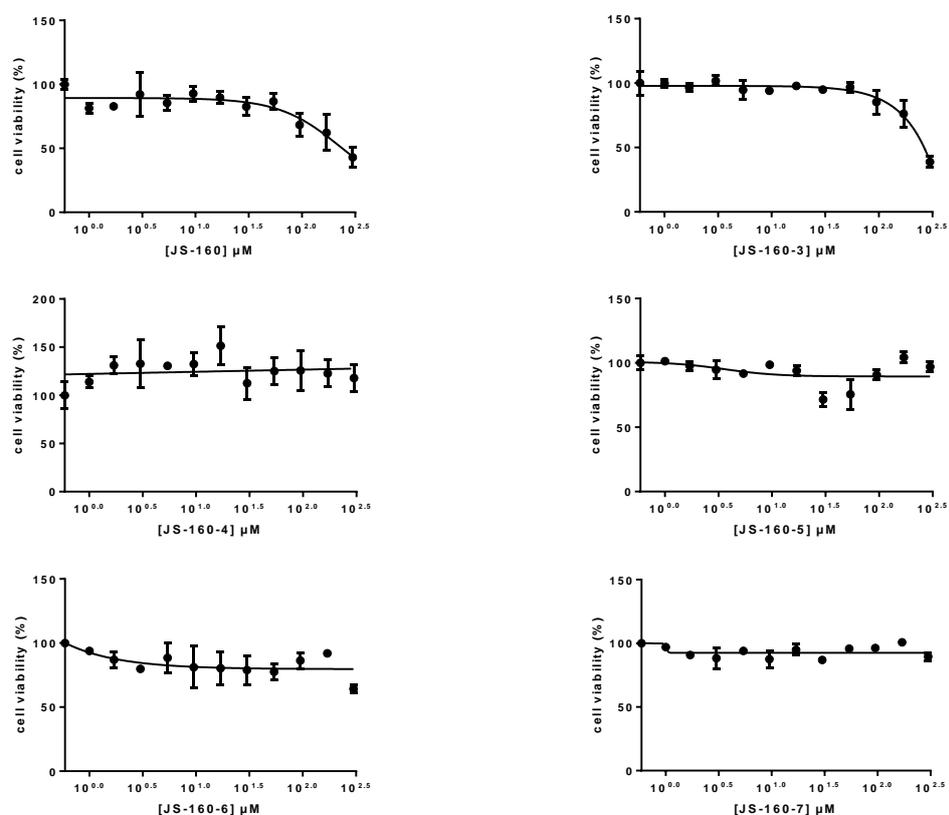
Figure 5.10: Toxicity profiles of JS-160 and inactive structural analogues.

Huh7.5 cells were treated with a range of concentrations of indicated compounds for 48 hours before monitoring cellular metabolism by MTT assay. Data normalised to DMSO control (plotted as $10^{-0.25}$ µM).
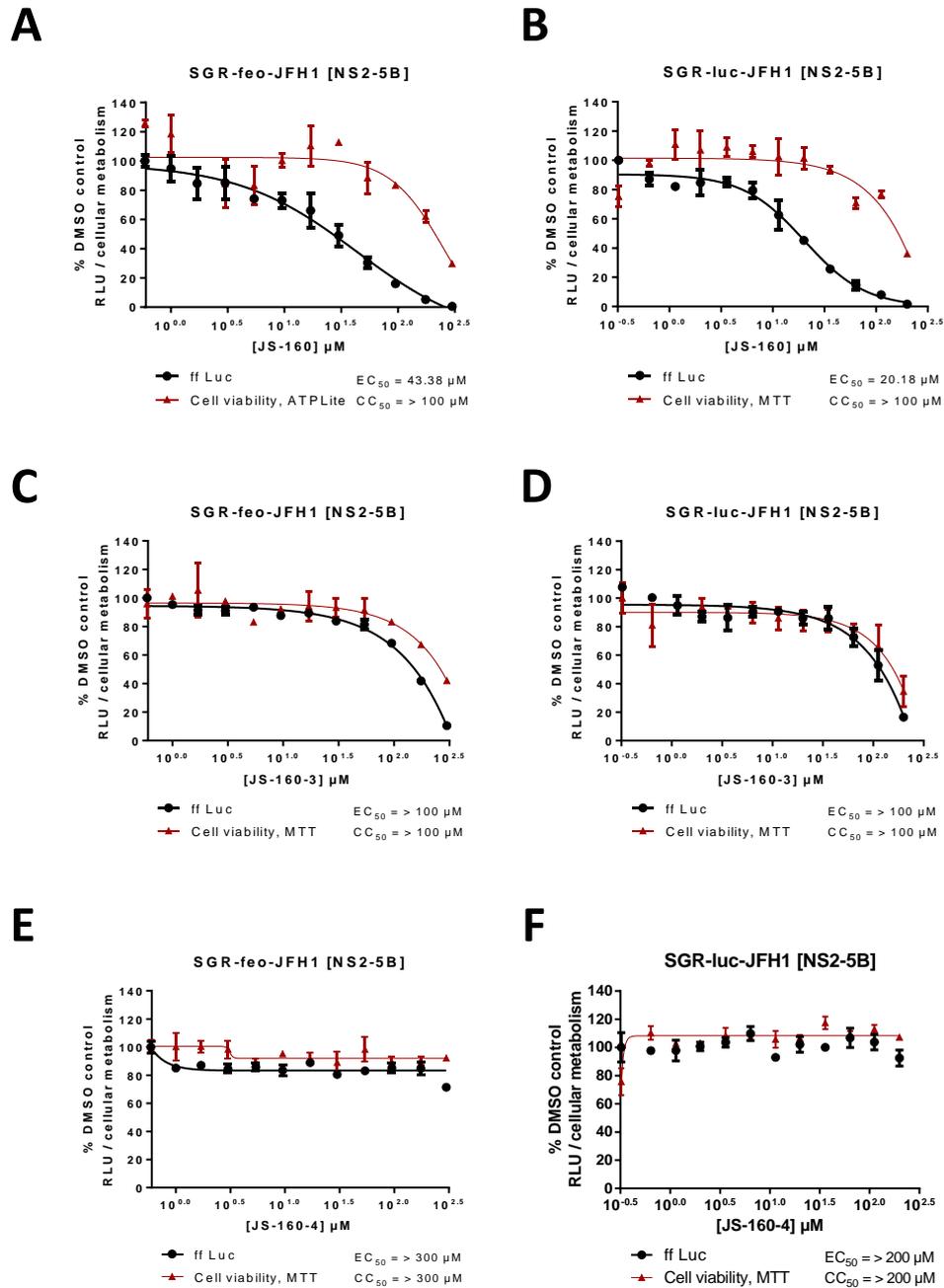
Figure 5.11: Activity of JS-160 and analogues against NS2-containing replicons.

A) SGR-feo-JFH1 [NS2-5B] was treated with a range of concentrations of JS-160 from 1 μM to 300 μM for 48 hours before quantifying luciferase (ffLuc). Cell viability was analysed in parallel. Data normalised to DMSO control. B) SGR-luc-JFH1 [NS2-5B] was electroporated into Huh7.5 cells and treated with the same concentrations of JS-160 at 4 h.p.e. for 48 hours. Cell viability was analysed in parallel. C) The same experiment as in A treated with JS-160-3 D) The same experiment as in B treated with JS-160-3 E) The same experiment as in A treated with JS-160-4. F) The same experiment as in B treated with JS-160-4.

To confirm the effect of JS-160 against NS2-dependent genome replication, the same experiments were repeated in the context of a transient electroporation into Huh7.5 cells. A transcript of SGR-luc-JFH1 [NS2-5B] was electroporated into Huh7.5 cells and treated with the same range of compound concentrations at 4 h.p.e.. The observed data showed a similar effect as in SGR-feo-JFH1 [NS2-5B]. JS-160 produced a dose-responsive reduction in luciferase activity in concentration ranges that exhibited no effect on cellular metabolism; EC_{50} = 20.2 µM, CC_{50} = > 100 µM (Figure 5.11B). Treatment with JS-160 in a transient SGR assay appeared to produce a lower EC_{50}. This effect could be due to enhanced sensitivity to inhibition of the NS2 autoprotease over 48 hours when all NS proteins and RCs have to be established *de novo* in a transient system, an effect previously observed for other classes of HCV DAAs (Lee et al., 2011).

Treatment of SGR-luc-JFH1 [NS2-5B] with JS-160-3 and JS-160-4 produced similar results as experiments in SGR-feo-JFH1 [NS2-5B] stable cell lines. Loss of luciferase activity upon treatment with JS-160-3 above 100 µM correlated with reduced cellular metabolism while addition of JS-160-4 had no effect on luciferase activity or cellular metabolism up to 300 µM (Figure 5.11F).

As the loss of luciferase activity upon addition of JS-160-3 correlates with reduced cellular metabolism, the MTT assay (as a measure of cell viability) and luciferase activity (as a measure of genome replication) appear to show similar sensitivity, supporting a direct effect of JS-160 against genome replication at concentrations which are not cytotoxic. To further validate this effect, an additional readout of cell viability was employed. Cell viability in the experiment in Figure 5.11A was analysed using the ATPLite assay to quantify levels of ATP. This system has the advantage of quantifying the number of healthy cells using a luciferase based assay and hence should show comparable assay sensitivity to the luciferase based readout of genome replication. However, because this assay uses luciferase activity as a measure of healthy cells, and SGR containing cells produce luciferase, ATPLite analysis was performed in parallel in naïve Huh7.5 cells. The toxicity profile of JS-160 in Huh7.5 cells using the ATPLite assay (Figure 5.11A) was comparable to the toxicity profile in Huh7.5 cells by MTT assay (Figure 5.10A) and to the toxicity profile in SGR containing cells using the MTT assay (Figure 5.11B). These data demonstrate an accurate measure of cytotoxicity caused by JS-160 and support an additional effect against the NS2-containing SGRs at concentrations below 100 µM.

5.2.3 Mode of action studies of NS2 autoprotease inhibitors

To explore the mode of action by which JS-160 inhibits HCV replication, the kinetics of inhibition at an intermediate concentration were investigated. Inhibitors of the NS2 autoprotease should act by perturbing polyprotein processing in a similar manner to NS3 protease inhibitors. The profile of JS-160 mediated inhibition was therefore compared to that of the NS3 protease inhibitor Telaprevir. SGR-luc-JFH1 [NS2-5B] was electroporated into Huh7.5 cells and at 4 h.p.e. an EC_{50} concentration of JS-160 (25 μM) or Telaprevir (100 nM) was added. Luciferase activity was assayed every 4 hours for the next 48 hours (52 h.p.e.). As an additional control, SGR-luc-JFH1 [NS3-5B] GND was electroporated and analysed in parallel.

As shown in Figure 5.12A, Huh7.5 cells electroporated with SGR-luc-JFH1 [NS2-5B] and treated with a DMSO control gave a gradual increase in luciferase activity over time, while the GND polymerase defective control produced no significant luciferase signal. Addition of Telaprevir (100 nM) led to a reduction in luciferase activity at later timepoints and a lag in detectable luciferase activity. Treatment with JS-160 (25 μM) produced a similar effect, reducing the luciferase signal and again delaying the onset of detectable luciferase activity (Figure 5.12A), indicative of a similar mechanism of action.

Based on the dose response curves shown in Figure 5.11 this concentration of JS-160 should not exert any cytotoxic effects. To confirm this within the context of the current experiment, cellular lysates at various hours post electroporation were analysed for the presence of the endogenous cellular protein GAPDH by western blot. To allow comparison of the total number of cells protein levels were not normalised. GAPDH could be detected from all tested lysates from DMSO treated-, JS-160 treated- or Telaprevir treated-cells, and in all cases levels of GAPDH increased over time (Figure 5.12B). Previous calculation of EC_{50} and CC_{50} for Telaprevir in SGR containing Huh7.5 cells had confirmed no toxic effects with up to 10 μM Telaprevir, while the CC_{50} of Telaprevir in hepatoma cell lines has been reported as greater than 10 μM (Kwong et al., 2011). As such treatment with 100 nM Telaprevir can be considered a similar control to DMSO only.

The observation of GAPDH in lysates confirmed the presence of significant cell numbers up to 48 h.p.e. in the presence of JS-160. In addition, to confirm that the addition of JS-160 has no effect on cell proliferation, GAPDH was quantified and plotted as a relative fold increase in intensity over 4 h.p.e.. A similar increase in GAPDH intensity over time to that of the DMSO control was observed in the presence of JS-160 (25 μM) or Telaprevir (100 nM) (Figure 5.12C).

Furthermore, a lack of cytotoxic effects in the presence of JS-160 at 25 µM was confirmed using the ATPLite assay in Huh7.5 cells. The average of three technical repeats revealed no effect on the levels of ATP as a measure of living cells upon treatment with JS-160 (Figure 5.13A). To confirm that inhibition of genome replication occurred in the absence of cytotoxic effects cellular lysates (at 52 h.p.e. from the experiment shown in Figure 5.12) were analysed by western blot for both GAPDH and NS5A. Both JS-160 (25 µM) and Telaprevir (100 nM) treated cells had reduced levels of NS5A relative to DMSO control, correlating with luciferase activity from Figure 5.11A (Figure 5.12B).
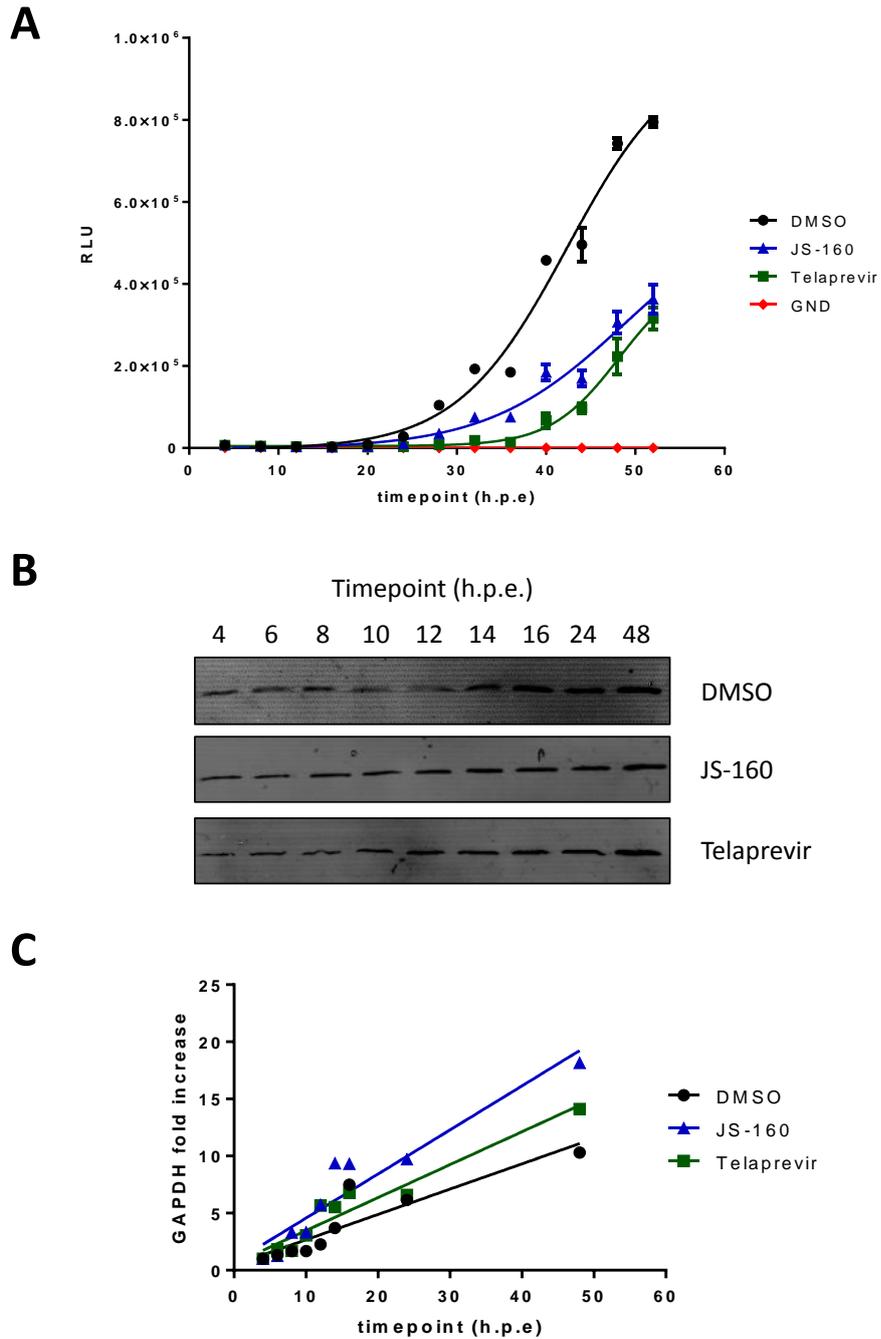
Figure 5.12: JS-160 inhibition kinetics.

A) SGR-luc-JFH1 [NS2-5B] was electroporated into Huh7.5 cells and treated at 4 h.p.e. with either JS-160 (25 μM), Telaprevir (100 nM) or DMSO control. As a negative control SGR-luc-JFH1 [NS3-5B] GND was also electroporated into Huh7.5 cells and remained untreated. Luciferase was monitored at indicated timepoints (n=3). B) Indicated timepoints from A analysed for levels of endogenous GAPDH by western blot. C) Quantified GAPDH from B normalised to 4 h.p.e. (n=1).
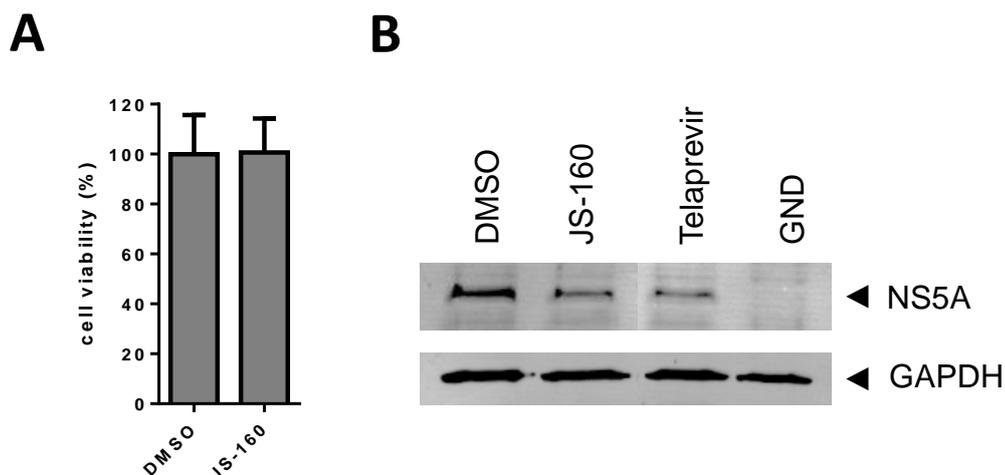
Figure 5.13: Activity of JS-160 in cell based assays.

A) JS-160 (25 µM) was added to Huh7.5 cells for 48 hours before analysing cell viability by quantifying ATP in a luciferase based system using the ATPLite assay. Luciferase activity is normalised to DMSO control. Data represents the average of three technical repeats, each in triplicate. B) SGR-luc-JFH1 [NS2-5B] was electroporated into Huh7.5 cells which were then treated with DMSO control, JS-160 (25 µM), or Telaprevir (100 nM). SGR-luc-JFH1 [NS3-5B] GND was also electroporated into Huh7.5 cells. After 48 hours lysates were analysed for NS5A and GAPDH by western blot.

Having confirmed that JS-160 exerts an inhibitory effect on NS2-dependent HCV replication in the absence of cytotoxicity, the same concentration (25 µM) was used to treat both SGR-feo-JFH1 [NS2-5B] and the SGR-feo-JFH1 [NS3-5B] cell line. Similarly, the anti-viral effects of JS-160 were assessed following electroporation of a Jc1-NLuc transcript of infectious HCVcc into Huh7.5 cells. In a similar manner to a transient SGR experiment, electroporation of a Jc1-Nluc transcript allows sensitive quantification of the degree of genome replication by quantifying Nanoluciferase (Nluc) activity from cellular lysates. However, Jc1 encodes a complete HCV genome along with the reporter (Figure 1.7) and is therefore capable of the assembly and release of infectious virions and the subsequent re-infection of naïve cells (Pietschmann et al., 2006). Therefore, Nluc activity is representative of genome replication in the context of the entire HCV lifecycle.

While neither JS-160 (25 µM) nor JS-160-3 (25 µM) had any significant effect on the luciferase activity from SGR-feo-JFH1 [NS3-5B], JS-160 did produce a 58.4 ± 7.2% reduction in luciferase activity from SGR-feo-JFH1 [NS2-5B] (Figure 5.14A). This data supports the observation from the cell-based screening that the inhibitory effect of JS-160 is only apparent when replication is dependent on the NS2 autoprotease, as with SGRs comprising NS2-5B. JS-160 produced a similar effect on Nluc activity in the context of Jc1 infectious virus, reducing signal to 38.5 ± 7.8% of DMSO control (Figure 5.14A). Comparative MTT assays performed in parallel again confirmed no significant effects on cellular metabolism upon addition of JS-160 (25 µM) (Figure 5.14B).
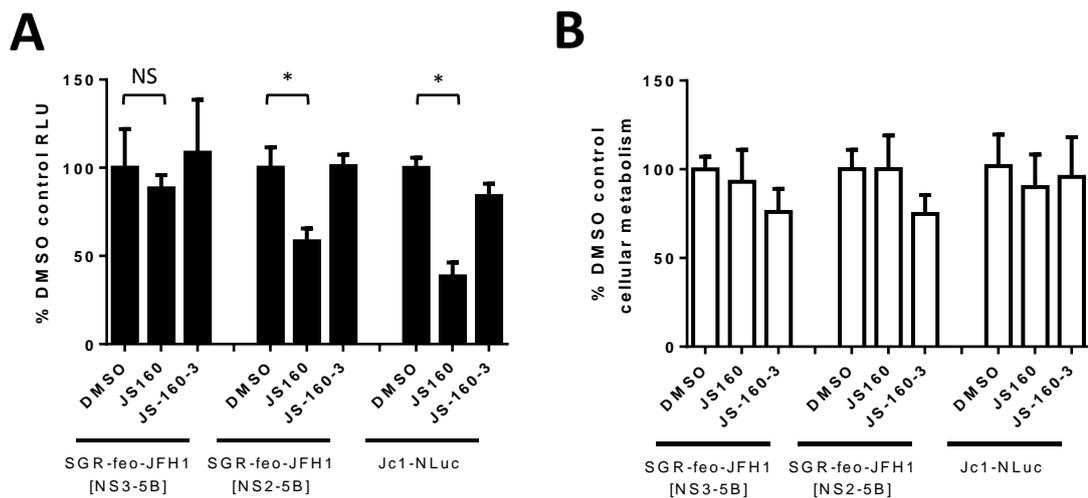
Figure 5.14: Activity of JS-160 against SGRs and infectious virus.

A) SGR-feo-JFH1 [NS3-5B] was treated with JS-160 (25µM) or JS-160-3 (25 µM) for 48 hours. Luciferase activity is normalised to DMSO control for each experiment (one technical repeat, n=3). The experiment was performed in parallel against SGR-feo-JFH1 [NS2-5B]. Jc1-NLuc was electroporated into Huh7.5 cells and treated at 4 h.p.e. with JS-160 (25 µM) or JS-160-3 (25 µM) for 48 hours alongside DMSO control. Statistical significance was determined using an unpaired Student's t test (* p < 0.005) B) Analysis of cellular metabolism by MTT assay from experiments in A.

To explore the sensitivity to JS-160 in the context of infectious HCV, a Jc1-NLuc transcript was electroporated into Huh7.5 cells and treated with a range of concentrations of either JS-160 or JS-160-3. After 48 hours, NLuc activity was quantified. The data obtained was similar to that observed in the context of SGRs (Figure 5.11). JS-160 had an inhibitory effect on NLuc activity with an EC_{50} of 49.4 µM, while cytotoxic effects were observed at higher concentrations, with CC_{50} = > 100 µM (Figure 5.15A). JS-160-3 exhibited a similar toxicity profile with a closer correlation between loss of NLuc activity and loss of viable cells such that both EC_{50} and CC_{50} were above 100 µM (Figure 5.15B). Toxicity analysis for this assay was performed in the same Huh7.5 cells electroporated with Jc1-NLuc transcript, hence these toxicity profiles are in the context of infectious virus.
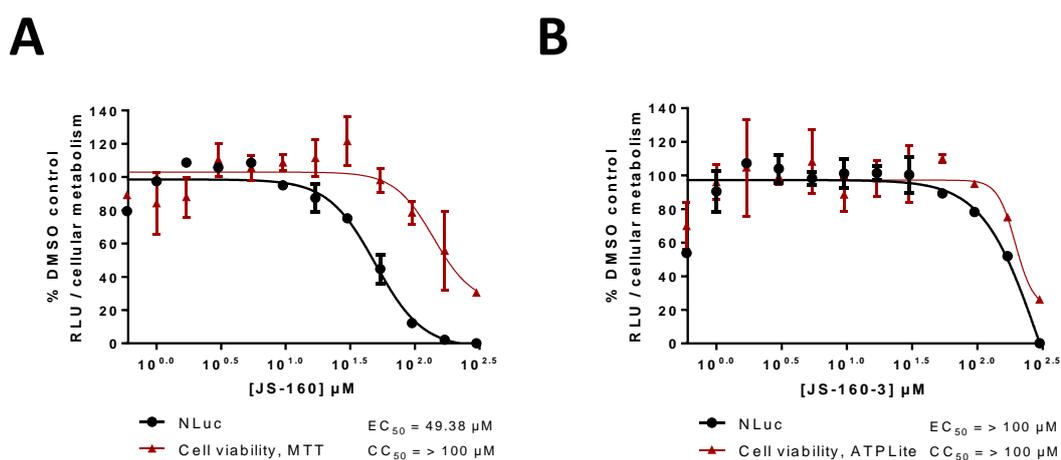
Figure 5.15: Activity of JS-160 and analogues against Jc1.

Jc1-NLuc was electroporated into Huh7.5 cells and treated with a range of concentrations of JS-160 from 1 µM – 300 µM at 4 h.p.e. for 48 hours. Quantified Nanoluciferase (NLuc) activity is normalised to untreated control. Identical JS-160 treated, Jc1 harbouring Huh7.5 cells were analysed for cellular metabolism by MTT assay normalised to untreated. B) Identical experiment as in A treated with a range of concentrations of JS-160-3.

As described in Section 1.5.4, NS2 is known to play a role in virus assembly. While this activity is thought to be independent of the NS2 protease activity, it may still require mature NS2 and hence be dependent on NS2-mediated processing at the NS2-NS3 junction.

JS-160 was therefore tested for additional effects against alternative stages of the HCV lifecycle. Huh7.5 cells were electroporated with Jc1-NLuc and treated with a range of inhibitor concentrations. At 48 h.p.e., virus-containing supernatant was harvested and cells lysed for NLuc quantification, indicative of genome replication. Virus was titrated onto naïve Huh7.5 cells, which were lysed 16 hours post infection to reduce the degree of re-infection and cell-to-cell transmission. NLuc activity of these lysates was measured and reflects the initial infectious titre of the treated producer cells.

NLuc activity from the producer cells was reduced in a dose-responsive manner by JS-160 (Figure 5.16B) to a similar degree as previously observed in Figure 5.15, while a comparable profile was observed upon treatment with JS-160-3 where NLuc activity was reduced at higher concentrations (Figure 5.16C). Previous results have demonstrated this effect to be due to cytotoxicity. Typical sensitivity to Telaprevir was also observed, with an EC_{50} of 121.4 nM (Figure 5.16D).

Analysis of infected cells following 16 hour incubation with virus-containing supernatant revealed significant NLuc activity, demonstrating infection of Huh7.5 cells with Jc1-NLuc. The infection experiment produced a strong NLuc signal approximately 1 $\log_{10}$ lower than the signal from producer cells 48 h.p.e. (Figure 5.16A). To simplify comparison between the two stages of the experiment, NLuc activity from both producer cells and infected cells was normalised to the relative DMSO control shown in Figure 5.16A.

NLuc activity from infected cells was reduced when replication was reduced (as monitored by producer cells NLuc activity), demonstrating that decreased genome replication over 48 hours leads to a reduction in the number of infectious virions released. As such the first conclusion drawn was that JS-160 mediated inhibition of HCV replication (at concentrations previously shown not to exert cytotoxic effects) manifests as a reduction into the number of infectious virions released and hence a decrease in infectious titre. This finding demonstrates an anti-viral effect of JS-160 and supports the use of NS2 autoprotease inhibitors as a viable class of DAA.

An additional observation was that NLuc activity from infected cells was more sensitive to inhibition by JS-160 than genome replication (producer cells NLuc activity) (Figure 5.16B). The

same effect was apparent to some degree with Telaprevir treatment (Figure 5.16D), though the effect was most pronounced for JS-160, where concentrations appearing to exert no effect against replication 48 h.p.e. reduced NLuc activity in the subsequent infection assay. This indicates that despite these concentrations of JS-160 having no effect on HCV replication, they reduce the number of infectious virions released, implicating an effect on virus assembly or release.

The design of this experiment is such that NLuc activity from infected cells may not directly correlate with the number of infectious virions in the media; for example there may have been compound carried over in the media which could further inhibit virus replication in the infected cells. This possibility and other limiting factors to these conclusions are discussed in Section 5.3.3. However, it is encouraging that this effect is not observed upon treatment with JS-160-3. NLuc activity from producer cells and infected cells correlates well in the presence of a range of JS-160-3 concentrations (Figure 5.16C), indicating no additional effect on virus release and that virus release is only reduced as a consequence of reduced genome replication. This reduction in genome replication upon JS-160-3 treatment has been previously shown to be due to cytotoxic effects. As this is not a specific anti-viral effect, no additional effect on virus release would be expected. Therefore JS-160, which inhibits the NS2 autoprotease *in vitro* and inhibits NS2-dependent genome replication in SGRs, appears to exert an additional effect on the release of infectious virus, while JS-160-3, which is inactive against the NS2 autoprotease *in vitro* and does not inhibit NS2-dependent genome replication in SGRs other than as a product of cytotoxicity, causes no additional effect on the release of infectious virus.
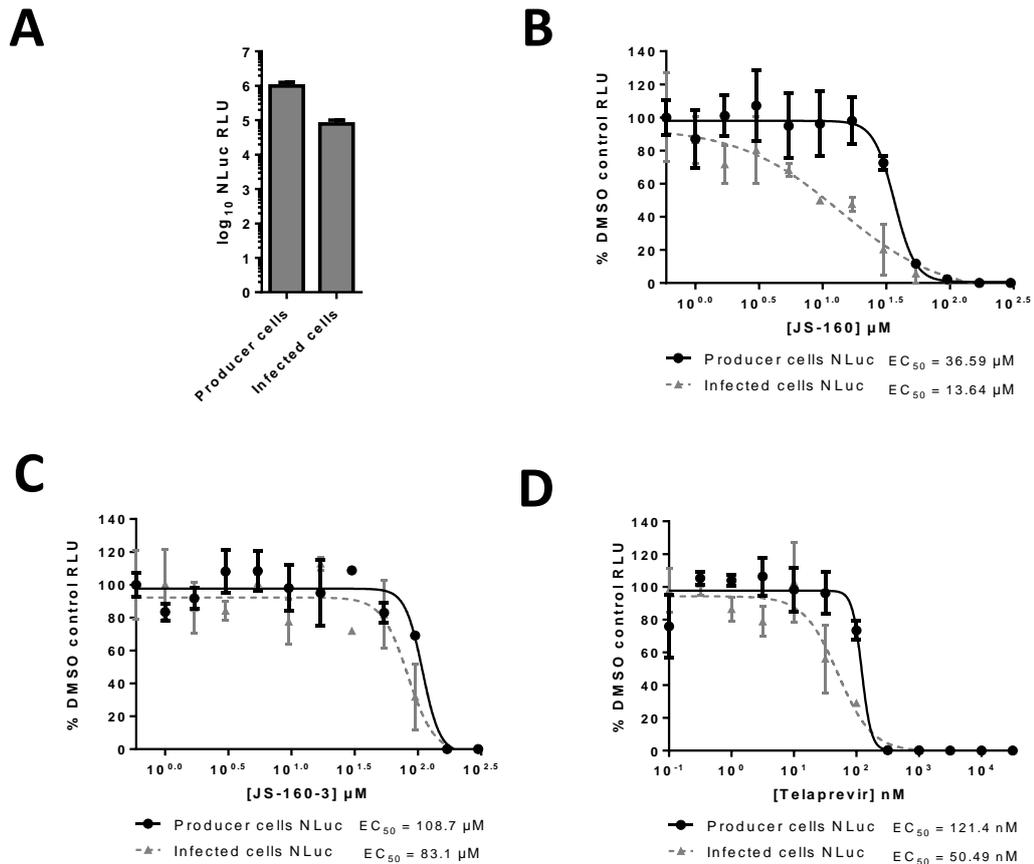
Figure 5.16: Effect of NS2 inhibitors on different stages of the virus lifecycle.

A) Nanoluciferase (NLuc) activity from Huh7.5 cells 48 h.p.e. of Jc1-NLuc RNA (producer cells) and NLuc activity from Huh7.5 cells following 16 hour incubation in virus containing supernatant (infected cells). Data the mean and standard deviation represented on a logarithmic scale of one experiment (n=6). B) Jc1-NLuc was electroporated into Huh7.5 cells and treated with a range of concentrations of JS-160 from 1 μM to 300 μM at 4 h.p.e. until 48 h.p.e.. At 48 h.p.e. NLuc was quantified (producer cells). Virus containing supernatant was titrated onto naïve Huh7.5 cells in the absence of inhibitor. NLuc activity was quantified 16 hours later (infected cells). NLuc activity from producer cells and infected cells was normalised to the relevant DMSO control in A. C) Identical experiment as in B treated with a range of concentrations of JS-160-3. D) Identical experiment as in A treated with a range of concentrations of Telaprevir from 316 pM to 31.6 μM.

5.2.4 Modelling NS2 autoprotease inhibitors

To model an interaction of this series of compounds with the NS2 autoprotease, compound structures were built using the Maestro interface (Schrodinger) and underwent energy minimisation. JS-160 was modelled against NS2^{pro}ΔP1-P2 using Autodock, with a single binding pose produce (Figure 5.17A). JS-160 was predicted to sit in the S1 pocket at the active site of the NS2 autoprotease. An identical binding pose was produced using the alternative software Glide. In addition, this JS-160 binding pose was reproduced using the NS2^{pro}ΔP1-P5 and NS2^{pro}ΔP1-P10 models. To explore whether features of this model matched experimental evidence, JS-160-1, JS-160-2 and JS-160-3 were modelled in the same manner. While JS-160-1 was able to reproduce this binding pose, JS-160-2 had a slightly altered pose, as did JS-160-3 (Figure 5.17A).

While these models matched the experimentally observed SAR to some degree, they could not explain all the observations made. The primary amine of JS-160 is proposed to form hydrogen bond contacts with backbone carboxylates from A179 and A182 (Figure 5.17A). JS-160-1 is predicted to bind in a similar manner, but lacks this amine group and hence cannot form these hydrogen bonds. This is reflected in a less favourable docking score for JS-160-1 relative to JS-160. However, *in vitro* SAR suggests JS-160-1 had comparable activity to JS-160 against NS2-mediated proteolysis. Hence the experimental evidence suggests this amine group does not contribute to JS-160 activity.

Further experimental evidence derived from SAR analysis revealed the cyclohexane of JS-160 to be essential for activity, yet the model binding pose of JS-160 predicts this region to be relatively solvent exposed with little obvious contribution to the ligand interaction. However, the models for both JS-160 and JS-160-1 (which showed comparable activity *in vitro*) place the 1-methyl cyclohexane group in an identical orientation. JS-160-2, which showed a slight reduction in activity in the absence of the 1-methyl group, is modelled with a modified cyclohexane position despite the 'core' of the compound docking in a similar manner to JS-160 and JS-160-1 (Figure 5.17A). Furthermore, JS-160-3, which was inactive *in vitro*, produced a pose that varied more significantly at this position (Figure 5.17A). Hence variability in the position of the cyclohexane group (or alternative groups in this position) from modelling ligand-protein interactions correlates with reduced or abolished activity observed experimentally. As the contribution of this cyclohexane region of the series is not apparent from the current models, its importance may be mediated through an alternative element of the NS2-3 precursor, perhaps through a region encoded within the NS3 N-terminal domain

which forms an interaction near the NS2 active site or within a cavity formed at such an interface. Structural data for the NS2-NS3 pre-cursor would help to further validate and refine the model for the current series of NS2 autoprotease inhibitors.

Residues of NS2 contributing to the S1 cavity where JS-160 is predicted to bind are highlighted in Figure 5.17A (amino acids within 4Å of the JS-160 docked pose). These residues appear highly conserved across all HCV genotypes (Figure 5.17B). Observations made against the NS2 autoprotease *in vitro* suggested enhanced potency against genotype 2a (JFH1) compared to genotype 1b (J4), yet the residues surrounding the NS2 active site differ by only a single substitution between these two isolates (D180E), while across a subset of other genotypes only residue 201 is not conserved. Such conservation is indicative of a strong selective pressure on residues contributing to the active site of NS2 and suggests that should inhibitors of the NS2 autoprotease bind at the active site they may exhibit a high barrier to resistance.

A

B

```
2HD0        QASLLKVPYFVRVQGLLRICALARKIAGGHYVQMAIIKLGALTGTYVYNHLTPLRDWAHN 153
gt1a_H77    QASLLKVPYFVRVQGLLRICALARKIAGGHYVQMAIIKLGALTGTYVYNHLTPLRDWAHN 153
gt1b_Con1   QAGITKVPYFVRAHGLIRACMLVRKVAGGHYVQMALMKLAALTGTYVYDHLTPLRDWAHA 153
gt1b_J4     QAGITRVPYFVRAQGLIRACMLVRKVAGGHYVQMVFMKLGALTGTYVYHHLTPLRDWAHA 153
gt2a_JFH1   RAALTHVPYFVRAHALIRVCALVKQLAGGRYVQVALLALGRWTGTYIYDHLTPMSDWAAS 153
gt3a_NZL1   QATITTTPYFVRAHVLVRLCMLVRSVIGGKYFQMIILSIGRWFNTYLYDHLAPMQHWAAA 153
gt4a_ED43   QASLLRIPYFVRAQALVKICSLLRGVVYGKYFQMVVLKSRGLTGTYIYDHLTPMSDWPPY 153
gt5a_SA13   QYSLIKLPYFIRARALLRACLLAKHLACGRYVQAALLHLGRLTGTYIYDHLAPMKDWAAS 153
gt6a_EUHK2  QYSLLKVPYFVRAHILLRACLLVRRLAGGKYVQACLLRLGAWTGTFVYDHLAPLSDWASD 153
            :  :   ***:*.: *:: * * : :  *:*.*  .:     .*::*:**:*: .*.
```

```
2HD0        GLRDLAVAVEPVVFSRMETKLITWGADTAACGDIINGLPVSARRGQEILLGPADGMVSKG 213
gt1a_H77    GLRDLAVAVEPVVFSRMETKLITWGADTAACGDIINGLPVSARRGQEILLGPADGMVSKG 213
gt1b_Con1   GLRDLAVAVEPVVFSDMETKVITWGADTAACGDIILGLPVSARRGREIHLGPADSLEGQG 213
gt1b_J4     GLRDLAVAVEPVVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEGQG 213
gt2a_JFH1   GLRDLAVAVEPIIFSPMEKKVIVWGAETAACGDIILGLPVSARLGQEILLGPADGYTSKG 213
gt3a_NZL1   GLKDLAVATEPVIFSPMEIKVITWGADTAACGDILCGLPVSARLGREVLLGPADDYREMG 213
gt4a_ED43   GLRDLAVALEPVVFTPMEKKVIVWGADTAACGDIIRGLPVSARLGNEILLGPADTETSKG 213
gt5a_SA13   GLRDLAVATEPIIFSPMEKKVIVWGADTAACGDILAGLPVSARRGHEIFLGPADDIREAG 213
gt6a_EUHK2  GLRDLAVAVEPVIFSPMEKKIITWGADTAACGDILSGLPVSARLGNLVLLGPADDMQRGG 213
            **:*****  **::*: ** *:*.***:******: ******  *.  :  *****    *
```

```
2HD0        WRLL 217
gt1a_H77    WRLL 217
gt1b_Con1   WRLL 217
gt1b_J4     WRLL 217
gt2a_JFH1   WKLL 217
gt3a_NZL1   WRLL 217
gt4a_ED43   WRLL 217
gt5a_SA13   WRLL 217
gt6a_EUHK2  WKLL 217
            *:**
```

Figure 5.17: Model of JS-160 – NS2^{pro} interaction.

A) Top left, predicted binding pose of JS-160 (blue carbons) docked to various models of NS2^{pro} (NS2ΔP1-P10 shown). The NS2 catalytic triad are highlighted with orange carbons. Residues within 4Å are shown. Top right, binding poses of active molecules JS-160 (blue carbons), JS-160-1 (pink carbons) and JS-160-2 (yellow carbons) and the inactive derivative JS-160-3 (grey carbons). Bottom left, space fill model of JS-160 in the S1 pocket of the NS2 active site. Bottom right, 2D ligand interaction diagram created using the Maestro interface highlighting interactions and residues within 4Å of the JS-160 binding pose. B) Sequence alignment of the NS2^{pro} C-terminal catalytic domain from the crystal structure (2HD0) and selected HCV genotype 1-6 isolates. Residues forming the catalytic triad are highlighted in green. Residues within 4Å of the S1 site are highlighted in yellow. Alignment performed using EMBL-EBI (Clustal W2).

5.3 Discussion

5.3.1 JS-160 is a small molecule inhibitor of NS2 mediated proteolysis *in vitro*

Results reported in Section 5.2.1 confirmed the identification of JS-160 as a small molecule capable of blocking NS2-mediated proteolysis *in vitro*. While the activity against J4 NS2-3 described in Chapter 4 was reproduced in a dose-responsive manner, JS-160 appears more potent against the JFH1 isolate of NS2. The enhanced potency against JFH1, and the greater signal to noise from refolding reactions using JFH1 NS2-3, allowed the calculation of the EC_{50} for blocking production of the NS3-FLAG cleavage product, which is produced as a measure of NS2 autoprotease activity. As such in this *in vitro* system JS-160 is active in the μM range with an EC_{50} of 44.6 ± 6.2 μM.

The activity of JS-160 described throughout Section 5.2 was verified for three separate batches of JS-160 from its original commercial source, while the structure shown is consistent with *m/z* peaks identified during mass spectrometry analysis (Appendix I). Synthesis of this molecule was attempted to confirm that the structure reported for JS-160 is responsible for inhibitory activity against the NS2 autoprotease. Amide coupling of commercially available starting materials (SM-1 and SM-2) yielded JS-160-1. Treatment of this product with an epoxidisation agent converted the alkene to an epoxide which could subsequently be opened with ammonia to yield a form of JS-160 without chiral selectivity (JS-160-1.3). However, as the epoxide intermediate JS-160-1.2 was insoluble in all tested solvents, poor yields were obtained for the final target molecule JS-160-1.3 preventing adequate purification or complete analysis. While the synthesis of some JS-160-1.3 was confirmed by mass spectrometry, an alternative synthetic route may be required to produce greater quantities of pure JS-160 (or derivatives thereof).

A degree of SAR surrounding this series has been performed *in vitro*. The observation of inhibition of the NS2 autoprotease upon treatment with JS-160-1, with no significant difference in potency from JS-160, reveals that the primary alcohol and amine groups at the chiral centres of the indene do not directly contribute to activity *in vitro*. While these chiral centres are not racemic in JS-160, the fact that NS2 inhibitors do not require these heteroatoms in an R, R conformation should simplify future synthesis exploring this series.

While JS-160-1 does not meet the criteria for lead-likeness as it only contains one hydrogen bond acceptor, impacting on cLogP, it retains features of JS-160, most notably an interesting non-planar shape. Such distinctive three-dimensional shapes are a favourable property of lead-

likeness (Nadin et al., 2012). The reduced activity of JS-160-2 revealed that the 1-methyl group at the cyclohexane of JS-160 contributes to, but is not required for, activity. These findings, along with the inactivity of both of the aforementioned starting materials (SM-1 and SM-2) indicate that the cyclohexane moiety makes a significant contribution to the activity of JS-160. Further SAR of this portion of JS-160 was performed using commercially available molecules JS-160-3 to JS-160-7. Replacing the 1-methyl cyclohexane portion of JS-160 with any of an isobutyl (JS-160-3), a hydroxyl-methyl (JS-160-4), a 3-pyridile (JS-160-5), a 2-hydroxy phenyl (JS-160-6) or a methyl-pyrrolidine-2-one (JS-160-7) all abolished activity below 300 μM, further confirming the contribution of the cyclohexane portion of JS-160. Compounds with greater structural diversity from JS-160 with a spiro-indolone in place of the spiro-indene core (JS-160-8 and JS-160-9) also lacked any activity against the NS2 autoprotease, though as both of these molecules also lack the cyclohexane moiety the precise effects of changes to the indene cannot be determined.

Current SAR, while failing to improve the activity of JS-160, has begun to map the molecular features that are required for and that contribute to activity. A particular structure consisting of a spiro-indene-piperidine core linked to a cyclohexane appears important for activity, further suggesting that the shape of this non-planar scaffold is paramount.

In agreement with this, the NS2-3 refolding reaction could tolerate 300 μM SM-1 and 300 μM SM-2 without affecting the levels of NS2-mediated proteolysis. These two starting materials each represent a slightly modified 'half' of the active molecule JS-160-1. Hence both 'halves' of JS-160-1 added individually do not block the NS2 autoprotease, yet when these two components are coupled to form a similar structure to JS-160 (JS-160-1), the resulting product acts as an inhibitor of the NS2 autoprotease. At an intermediate concentration, JS-160-1 has little detectable effect on the rate of NS2-mediated proteolysis but reduces the amount of NS2-mediated processing for the duration of the assay (Figure 5.9A).

The current series contains three compounds active against the NS2 autoprotease *in vitro*.

5.3.2 Effects of NS2 autoprotease inhibitors on HCV genome replication

While activity in the μM range in an *in vitro* system is unlikely to translate into remarkable activity against HCV in cell-based systems it is important to determine that this series displays anti-viral activity at an early stage. Unfortunately, JS-160 showed a degree of cytotoxicity at higher concentrations. The observation of reduced cellular metabolism in the presence of ≥ 100 μM JS-160 complicated analysis of anti-viral effects. Therefore significant efforts were

made to establish a therapeutic window between cytotoxic effects and anti-viral effects. The structural analogue JS-160-3 was used to help define this window. JS-160-3 was inactive against the NS2 autoprotease *in vitro* but showed a comparable toxicity profile to JS-160 in Huh7.5 cells. All other inactive analogues of JS-160 tested did not affect cellular metabolism below 300 µM. This suggests that the presence of a hydrophobic group in position of the cyclohexane of JS-160 is associated with cytotoxic effects, though only a cyclohexane group at this position confers activity against the NS2 autoprotease. JS-160-3 was therefore used as a comparative tool in cell-based assays to add additional evidence for an effect against genome replication greater than that inferred solely due to cytotoxic effects.

JS-160 caused a reproducible reduction in luciferase activity from SGR-feo-Con1 [NS2-5B] in Chapter 4 (Figure 4.10). Upon addition of a range of concentrations to SGR-feo-JFH1 [NS2-5B], JS-160 caused a dose-responsive reduction in luciferase activity, with an EC_{50} of 43.4 µM. Comparative analysis of cytotoxic effects through two independent assays demonstrated reduced luciferase activity at concentrations which did not affect cell viability, most notably upon addition of 54 µM, 30 µM, 17 µM and 9.5 µM JS-160 (Figure 5.11A). An identical experiment treated with JS-160-3 only reduced luciferase activity at higher concentrations such that an EC_{50} could not be accurately fitted, with a correlation between cell viability and luciferase activity apparent. Upon treatment with JS-160-3 both the EC_{50} for effect against the luciferase reporter and the CC_{50} were greater than 100 µM (Figure 5.11C).

Data reported in Figure 5.11 demonstrates that JS-160, which inhibits the NS2 autoprotease *in vitro*, has an inhibitory effect against NS2-dependent genome replication (at concentrations lower than those determined as cytotoxic) while JS-160-3, which lacks activity against the NS2 autoprotease *in vitro*, does not affect genome replication other than as a consequence of cytotoxicity. Hence SAR observed *in vitro* is also apparent in SGR based assays. As further evidence of a correlation between *in vitro* activity and reduced luciferase activity in cell-based SGR assays, a JS-160 analogue lacking activity *in vitro* and devoid of cytotoxicity (JS-160-4) had no effect on luciferase activity up to 300 µM (Figure 5.11E).

The data obtained using SGR-feo-JFH1 [NS2-5B] was replicated in a transient SGR system using a previously reported construct (SGR-luc-JFH1 [NS2-5B]) (Figure 5.13B, D and E). However it was noted that JS-160 appeared more potent in this transient system with an EC_{50} of 20.2 µM. Such a shift in sensitivity might be due to the different SGR systems. As a stable cell line harbouring constitutive genome replication, SGR-feo-JFH1 [NS2-5B] contains numerous active RCs at the time of addition of the compound. In contrast, the transient assay requires all RCs to

be produced *de novo*, with fewer RCs established within these cells at the time of adding compound. Hence a transient system may be more sensitive to inhibition as it is more reliant on the production of new RCs. Such a phenomenon has been previously reported for other classes of DAA (Lee et al., 2011), thus further supporting an inhibitory action of JS-160 against genome replication.

The mode of action of JS-160 is proposed to be a block of polyprotein processing, a very similar mode of action to NS3 protease inhibitors such as Telaprevir. In support of this JS-160 has a similar inhibition profile to Telaprevir at an intermediate concentration (Figure 5.12A). This effect is apparent at a concentration of JS-160 that does not exert any cytotoxicity by MTT assay in Huh7.5 cells or SGR containing Huh7.5 cells (Figure 5.10, Figure 5.11), by ATPLite assay in Huh7.5 cells (Figure 5.11A, Figure 5.13A) or through detection of endogenous GAPDH (Figure 5.12B, C). Furthermore, the reduction in luciferase activity upon addition of this dose of JS-160 correlates with reduced levels of HCV NS5A by western blot analysis (Figure 5.13B).

As previously observed for Con1 derived SGRs in the cell based screening assay in Section 4.2.3, the reduction of luciferase activity from SGR-feo-JFH1 [NS2-5B] upon treatment with a non-cytotoxic dose of JS-160 was not observed upon treatment of SGR-feo-JFH1 [NS3-5B], indicating an effect against replication mediated through the NS2 protein. The current data therefore supports an inhibitory effect of JS-160 against genome replication mediated through the NS2 protein with a similar mode of action to an NS3 protease inhibitor which blocks polyprotein processing.

5.3.3 Anti-viral effects of NS2 autoprotease inhibitors on the HCV lifecycle

In addition to blocking NS2-dependent genome replication in a SGR system, JS-160 reduces reporter activity as a measure of viral replication in the context of infectious HCV. Jc1-Nluc is a chimeric HCV genome containing a Nanoluciferase reporter which is capable of numerous rounds of infection in cell culture (Pietschmann et al., 2006). Addition of JS-160 following electroporation of Jc1-Nluc into Huh7.5 cells reduced reporter activity in a similar manner to that observed with SGRs that contain NS2-5B (Figure 5.14), indicating a similar block on genome replication within the context of infectious virus. This observation was confirmed at several doses of JS-160 where, as in SGR based experiments, luciferase activity was reduced in a dose responsive manner, with cytotoxicity only observed at higher concentrations. Similarly,

the profile of the inactive analogue JS-160-3 was comparable to experiments using SGRs, indicating this analogue has no anti-viral effect.

Data presented in Figure 5.16 demonstrates that this effect on HCV replication leads to a reduction or loss of infectious HCV particles secreted into the media. JS-160, within a range of concentrations previously shown not to exert cytotoxicity, affected not only the activity of the reporter in producer cells over the first 48 hours following electroporation (a measure of genome replication following artificial introduction of the HCV genome) but also reporter signal from naive cells subjected to virus containing supernatant.

Huh7.5 cells with significant Nluc activity within this re-infection assay, such as when untreated (Figure 5.16A) represent infection by Jc1 virus particles which were released into the media within the first 48 hours following electroporation. Cells not infected with Jc1-Nluc virus particles would not otherwise contain the Nluc reporter. As would be expected, where genome replication is reduced or abolished in producer cells, the effect on replication manifests as a reduction in the number of released infectious virions. Hence upon addition of this media to naïve Huh7.5 cells the Nluc activity (as a measure of infection) is reduced relative to DMSO control, demonstrating a reduction in infectious titre.

This demonstrates that addition of non-cytotoxic concentrations of JS-160 (e.g. 54 µM and 30 µM) inhibits genome replication and reduces the degree of Nluc activity as a measure of infected cells. JS-160 therefore reduces the infectious titre of HCV, most likely by diminishing the number of infectious virions released as a consequence of blocking genome replication. This data therefore supports the use of inhibitors of the NS2 autoprotease to exert an anti-viral effect.

This assay also offers insights into the effect of NS2 autoprotease inhibitors on different stages of the virus lifecycle. NLuc activity from infected cells appears more sensitive to inhibition by JS-160 than NLuc from producer cells, hence the effect of JS-160 on infectious titre is greater than the effect on genome replication. Not only was the calculated EC_{50} from infected cells lower than the calculated EC_{50} for producer cells, but concentrations of JS-160 that did not affect genome replication reduced Nluc activity in the subsequent infection assay. This could suggest an additional effect of JS-160 on other stages of the HCV lifecycle. Significant evidence links NS2 with a role in release, possibly at an early stage of virion assembly through interactions with other HCV proteins, most notably p7, E2 and NS3. JS-160 may exert an inhibitory effect on the process of virus assembly and release through the NS2 protein at doses

that do not significantly perturb genome replication. Further studies will be required to confirm such an effect as the current assay contains a number of caveats.

Firstly, the virus containing supernatant added to infected cells will still contain compound. Given the enhanced potency of JS-160 in a transient SGR assay compared to a stable SGR assay, JS-160 may exhibit greater potency still in the context of an infection, as the cell is subjected to much less infectious genome than in the artificial situation of electroporation. The presence of compound with the virus before infection also allows for the possibility of an additional effect on HCV entry. Given that NS2 is not thought to be present in the virion or to play any role in entry this seems unlikely.

Secondly, Nluc activity only serves as an indirect measure of the number of infectious particles. This reporter quantifies the amount Nluc that is present in the cell as a consequence of infection and translation from the genome and is amplified further by replication within the infected cells over 16 hours. With numerous variables the Nluc activity may not correlate directly with the number of infectious virions (i.e. a 50% reduction in the number of released infectious particles may not necessarily correlate with a 50% reduction in Nluc signal from the infected cells). However, such differences in sensitivity are unlikely to explain why treatment of producer cells with a concentration of JS-160 that does not affect replication (9.5 μM) leads to a 50% reduction in Nluc activity in infected cells.

Controls included in this experiment support an additional effect of JS-160 against virus release. Firstly, a shift in sensitivity of Nluc activity between replication and infection assays is not observed to the same degree with JS-160-3. This molecule is inactive against the NS2 autoprotease and the reduced Nluc activity at higher concentrations has been shown to be a consequence of cytotoxicity. Both the degree of Jc1 replication and the number of infectious Jc1 virions released (as monitored by NLuc activity in infected cells) correlate well, suggesting both are a consequence of the loss of viable cells.

In addition, inhibitors of the NS3 protease have recently been reported to exert effects on HCV release. A kinetic analysis of a range of DAAs revealed reductions in virus release at early timepoints before a later effect on genome replication, an effect observed for both NS5A inhibitors and the 1st generation NS3 protease inhibitor Boceprevir (McGivern et al., 2014). The NS3 protease inhibitor Telaprevir appeared to exert additional effects on release as measured by NLuc activity from infected cells in a similar manner to JS-160. This data supports an additional effect of the NS3 protease inhibitor Telaprevir on virus release and suggests that

inhibitors of the NS2 autoprotease may exert a similar effect. As previously reported, the effect of JS-160 on virus release should be tested at earlier timepoints, before a significant effect on replication has occurred.

While both NS2 and NS3 are documented to play roles in HCV assembly and release, an intriguing hypothesis might be that release is tightly linked to polyprotein processing. Given the important roles for several non-structural proteins in this process, most notably NS5A, the process of HCV release may be more sensitive to a block of polyprotein processing than the process of genome replication. It certainly seems feasible that perturbing polyprotein processing, either through an NS3 protease inhibitor or an NS2 autoprotease inhibitor, would negatively impact both replication and release.

While further studies will be required to explore the effects of these compounds on different stages of the virus lifecycle, the data presented in this chapter demonstrates an inhibitory effect of JS-160 against HCV genome replication where polyprotein processing relies on the NS2 autoprotease, and that this activity produces an anti-viral effect capable of reducing the number of infectious virions released from HCV infected cells.

Chapter 6 – Conclusions and future perspectives

Prior to this study NS2 represented the only HCV-encoded enzyme activity that had not been targeted with small molecule inhibitors, despite extensive evidence that the NS2 autoprotease activity is essential in the virus lifecycle. This study set out both to test the hypothesis that a small molecule inhibitor of the NS2 autoprotease could exert an anti-viral effect and to identify a lead-like compound as a suitable starting point for the development of a DAA targeting NS2.

Following the observation that chaperone antagonists, which were previously reported to perturb NS2-NS3 proteolysis, do not mediate an effect against NS2-dependent genome replication, it was apparent that the ability to block HCV replication by pharmalogical inhibition of NS2 had yet to be demonstrated. The observation that an epoxide-based small molecule which blocks NS2-mediated proteolysis *in vitro* is able to reduce reporter signal from a SGR that requires NS2-NS3 processing for replication offered the first evidence that an anti-viral effect could be exerted through an inhibitor of the NS2 autoprotease.

Subsequently, the combination of two assays which measure NS2 autoprotease function allowed for the identification of a lead-like small molecule inhibitor of the NS2 autoprotease; JS-160. This compound blocks NS2-mediated processing at the NS2-NS3 junction *in vitro* and exerts a specific effect on SGRs that require NS2 autoprotease function for constitutive replication. Furthermore, JS-160 blocks replication in the context of infectious virus in cell culture, which manifests as a reduction in the number of infectious virions released from a cell population. Hence inhibitors of the NS2 autoprotease are able to exert an anti-viral effect.

These findings suggest NS2 autoprotease inhibitors could represent an additional class of DAA to complement combination therapies for the treatment of HCV. Inhibitors of the NS2 autoprotease have the advantage of targeting an entirely novel molecular target yet acting through a somewhat established mode of action, in that perturbing HCV polyprotein processing has already been shown to produce a beneficial effect in the clinic through inhibitors of the NS3 protease.

A degree of SAR analysis has been performed for the current series, beginning to map the requirements for activity. An obvious next step would be a more extensive SAR programme to attempt to improve activity. Ideally such a programme would benefit from further CADD, though current models of an interaction at the NS2 active site may need further refinement and ideally, experimental evidence, as at present they do not correlate with all the experimental observations. Indeed the precise impact of using structure-guided CADD prior to screening for an NS2 autoprotease inhibitor is hard to gauge. While circumstantial evidence

suggests a hit-rate higher than would be expected from unbiased HTS, no direct evidence exists as yet that JS-160 interacts with NS2 as modelled *in silico*. A more feasible route for future SAR may be to focus on assays amenable to more high-throughput systems. This is unlikely to be achieved using the *in vitro* NS2-3 refolding assay, but could be performed in the cell-based SGR screen, as evidenced by re-formatting of this assay and its application to screen a library of > 10,000 compounds.

The current series should undergo further study to confirm that inhibition of genome replication is a consequence of a block on polyprotein processing. While these compounds have been shown to block NS2-mediated proteolysis *in vitro* and show comparable inhibition kinetics to an NS3 protease inhibitor against replicons, a build-up of unprocessed polyprotein would be difficult to observe in the current systems. Where polyprotein processing is blocked, the inhibition of replication (due to the absence of mature viral proteins to form RCs) prevents sufficient protein accumulating for detection (and may be further influenced by the rapid degradation of unprocessed NS2-NS3). Detection of unprocessed polyprotein in the presence of JS-160 would require the use of expression constructs or constitutive expression through a vaccinia system, systems previously used to explore inhibitor mode-of-action for other classes of HCV DAAs (Targett-Adams et al., 2011).

Experiments using the current series should also further explore effects on different stages of the virus lifecycle. Additional effects on the assembly and release of virus are beginning to become apparent for several DAA classes including NS5A inhibitors and NS3 protease inhibitors (McGivern et al., 2014), with obvious advantages for progression to the clinic. Confirming effects of NS2 autoprotease inhibitors on multiple stages of the virus lifecycle would support their progression as a DAA.

The ability of HCV to establish resistance to an NS2 autoprotease inhibitor should be examined, and the identification of resistance mutations encoded within the NS2 protein would further validate the molecular target. While this experiment is challenging at present due to the narrow therapeutic window of JS-160 in cell culture, the high degree of conservation at the NS2 active site across all HCV genotypes suggests a strong selective pressure and indicates that inhibitors which bind at the NS2 active site may exhibit a high barrier to resistance. As with many of the investigations described above, the identification of resistant variants would benefit from a more potent inhibitor of the NS2 autoprotease.

It will also be important to investigate the selectivity of this series with regards to other viruses, other proteases and in particular, to the NS2 proteins of related viruses which encode a protease. For example, the catalytic domain of the NS2 protein from the *Pestivirus* BVDV shows considerable homology to HCV NS2, and has been shown to act as a stand-alone cysteine protease (Lackner et al., 2004). While HCV NS2 protease activity is enhanced by the zinc binding domain of NS3 (Schregel et al., 2009), NS2 of BVDV encodes a series of additional residues proposed to coordinate zinc and replace the contribution of NS3 to NS2-NS3 processing required in HCV (Lackner et al., 2004). It would therefore be intriguing to test the current series for activity against BVDV or other members of the *Pestivirus* genera. In contrast, for members of the *Flavivirus* genera (e.g. YFV), the NS3 protease performs all the non-structural proteolysis events with NS2A-2B acting as a co-factor (Chambers et al., 1991, Chambers et al., 1990). Given that JS-160 appears inactive against the HCV NS3 protease (as it does not affect a SGR containing NS3-5B), this series would be expected to show no activity against members of the *Flavivirus* genus.

More broadly, to the best of our knowledge this series represents the first small molecule inhibitors of a viral autoprotease. Other viral autoproteases have been postulated as attractive therapeutic targets, such as the HIV-1 autoprotease activity encoded within the Gag-Pol precursor (PR) (Huang, 2013). As an aspartyl autoprotease formed from a homo-dimer this HIV-encoded target could benefit from identification of inhibitors of the HCV NS2 autoprotease, at the least as evidence that a viral autoprotease can be blocked with a small molecule inhibitor. Other virally encoded autoproteases that may be of interest for this series include the N^{pro} protein of BVDV, a cysteine autoprotease for which structural information has recently become available (Gottipati et al., 2013).

Finally, regardless of its viability as a DAA, JS-160 and other inhibitors of the NS2 autoprotease may be of use for further study of the HCV NS2-NS3 precursor. The enhancing role of the NS3 N-terminus on the NS2 autoprotease remains poorly understood. Given the evidence for a regulatory role of NS2-NS3 processing on both viral replication and the pathogenicity of BVDV (Lackner et al., 2004), processing at the HCV NS2-NS3 junction may have important implications for the mechanism by which HCV self-regulates its lifecycle, with direct implications for the disease state and ability of HCV to establish a persistent, chronic infection. The use of an NS2 autoprotease inhibitor may aid structural studies on the pre-processed form of NS2-3 which would further inform on the molecular basis by which the NS2 autoprotease functions.

Chapter 7 – References

AGARWAL, A. K. & FISHWICK, C. W. 2010. Structure-based design of anti-infectives. *Ann N Y Acad Sci,* 1213**,** 20-45.

ALIGO, J., JIA, S., MANNA, D. & KONAN, K. V. 2009. Formation and function of hepatitis C virus replication complexes require residues in the carboxy-terminal domain of NS4B protein. *Virology,* 393**,** 68-83.

APPEL, N., ZAYAS, M., MILLER, S., KRIJNSE-LOCKER, J., SCHALLER, T., FRIEBE, P., KALLIS, S., ENGEL, U. & BARTENSCHLAGER, R. 2008. Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS Pathog,* 4**,** e1000035.

APPLEBY, T. C., ANDERSON, R., FEDOROVA, O., PYLE, A. M., WANG, R., LIU, X., BRENDZA, K. M. & SOMOZA, J. R. 2011. Visualizing ATP-dependent RNA translocation by the NS3 helicase from HCV. *J Mol Biol,* 405**,** 1139-53.

BALDICK, C. J., WICHROSKI, M. J., PENDRI, A., WALSH, A. W., FANG, J., MAZZUCCO, C. E., POKORNOWSKI, K. A., ROSE, R. E., EGGERS, B. J., HSU, M., ZHAI, W., ZHAI, G., GERRITZ, S. W., POSS, M. A., MEANWELL, N. A., COCKETT, M. I. & TENNEY, D. J. 2010. A novel small molecule inhibitor of hepatitis C virus entry. *PLoS Pathog,* 6**,** e1001086.

BANKWITZ, D., STEINMANN, E., BITZEGEIO, J., CIESEK, S., FRIESLAND, M., HERRMANN, E., ZEISEL, M. B., BAUMERT, T. F., KECK, Z. Y., FOUNG, S. K., PECHEUR, E. I. & PIETSCHMANN, T. 2010. Hepatitis C virus hypervariable region 1 modulates receptor interactions, conceals the CD81 binding site, and protects conserved neutralizing epitopes. *J Virol,* 84**,** 5751-63.

BARREIRO, P., VISPO, E., POVEDA, E., FERNANDEZ-MONTERO, J. V. & SORIANO, V. 2013. Hepatitis C therapy: highlights from the 2012 annual meeting of the European Association for the Study of the Liver. *Clin Infect Dis,* 56**,** 560-6.

BARTENSCHLAGER, R., PENIN, F., LOHMANN, V. & ANDRE, P. 2011. Assembly of infectious hepatitis C virus particles. *Trends Microbiol,* 19**,** 95-103.

BEHRENS, S. E., TOMEI, L. & DE FRANCESCO, R. 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J,* 15**,** 12-22.

BERAN, R. K., LINDENBACH, B. D. & PYLE, A. M. 2009. The NS4A protein of hepatitis C virus promotes RNA-coupled ATP hydrolysis by the NS3 helicase. *J Virol,* 83**,** 3268-75.

BERGER, K. L. & RANDALL, G. 2009. Potential roles for cellular cofactors in hepatitis C virus replication complex formation. *Commun Integr Biol,* 2**,** 471-3.

BHATTACHARYA, D., ANSARI, I. H., MEHLE, A. & STRIKER, R. 2012. Fluorescence resonance energy transfer-based intracellular assay for the conformation of hepatitis C virus drug target NS5A. *J Virol,* 86**,** 8277-86.

BIHOVSKY, R., POWERS, J. C., KAM, C. M., WALTON, R. & LOEWI, R. C. 1993. Further evidence for the importance of free carboxylate in epoxysuccinate inhibitors of thiol proteases. *J Enzyme Inhib,* 7**,** 15-25.

BINDER, M., KOCHS, G., BARTENSCHLAGER, R. & LOHMANN, V. 2007. Hepatitis C virus escape from the interferon regulatory factor 3 pathway by a passive and active evasion strategy. *Hepatology,* 46**,** 1365-74.

BLANCHARD, E., BELOUZARD, S., GOUESLAIN, L., WAKITA, T., DUBUISSON, J., WYCHOWSKI, C. & ROUILLE, Y. 2006. Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J Virol,* 80**,** 6964-72.

BLIGHT, K. J., MCKEATING, J. A. & RICE, C. M. 2002. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J Virol,* 76**,** 13001-14.

BOSON, B., GRANIO, O., BARTENSCHLAGER, R. & COSSET, F. L. 2011. A concerted action of hepatitis C virus p7 and nonstructural protein 2 regulates core localization at the endoplasmic reticulum and virus assembly. *PLoS Pathog,* 7**,** e1002144.

BRADLEY, D. W., MAYNARD, J. E., POPPER, H., COOK, E. H., EBERT, J. W., MCCAUSTLAND, K. A., SCHABLE, C. A. & FIELDS, H. A. 1983. Posttransfusion non-A, non-B hepatitis: physicochemical properties of two distinct agents. *J Infect Dis,* 148**,** 254-65.

BRADLEY, D. W., MCCAUSTLAND, K. A., COOK, E. H., SCHABLE, C. A., EBERT, J. W. & MAYNARD, J. E. 1985. Posttransfusion non-A, non-B hepatitis in chimpanzees. Physicochemical evidence that the tubule-forming agent is a small, enveloped virus. *Gastroenterology,* 88**,** 773-9.

BRASS, V., BERKE, J. M., MONTSERRET, R., BLUM, H. E., PENIN, F. & MORADPOUR, D. 2008. Structural determinants for membrane association and dynamic organization of the hepatitis C virus NS3-4A complex. *Proc Natl Acad Sci U S A,* 105**,** 14545-50.

BURBELO, P. D., DUBOVI, E. J., SIMMONDS, P., MEDINA, J. L., HENRIQUEZ, J. A., MISHRA, N., WAGNER, J., TOKARZ, R., CULLEN, J. M., IADAROLA, M. J., RICE, C. M., LIPKIN, W. I. & KAPOOR, A. 2012. Serology-enabled discovery of genetically diverse hepaciviruses in a new host. *J Virol,* 86**,** 6171-8.

CATANESE, M. T., URYU, K., KOPP, M., EDWARDS, T. J., ANDRUS, L., RICE, W. J., SILVESTRY, M., KUHN, R. J. & RICE, C. M. 2013. Ultrastructural analysis of hepatitis C virus particles. *Proc Natl Acad Sci U S A,* 110**,** 9505-10.

CHAMBERS, T. J., GRAKOUI, A. & RICE, C. M. 1991. Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 proteinase domain and NS2B are required for cleavages at dibasic sites. *J Virol,* 65**,** 6042-50.

CHAMBERS, T. J., WEIR, R. C., GRAKOUI, A., MCCOURT, D. W., BAZAN, J. F., FLETTERICK, R. J. & RICE, C. M. 1990. Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proc Natl Acad Sci U S A,* 87**,** 8898-902.

CHATTERJI, U., BOBARDT, M., SELVARAJAH, S., YANG, F., TANG, H., SAKAMOTO, N., VUAGNIAUX, G., PARKINSON, T. & GALLAY, P. 2009. The isomerase active site of cyclophilin A is critical for hepatitis C virus replication. *J Biol Chem,* 284**,** 16998-7005.

CHATTERJI, U., LIM, P., BOBARDT, M. D., WIELAND, S., CORDEK, D. G., VUAGNIAUX, G., CHISARI, F., CAMERON, C. E., TARGETT-ADAMS, P., PARKINSON, T. & GALLAY, P. A. 2010. HCV resistance to cyclosporin A does not correlate with a resistance of the NS5A-cyclophilin A interaction to cyclophilin inhibitors. *J Hepatol,* 53**,** 50-6.

CHEN, S. L. & MORGAN, T. R. 2006. The natural history of hepatitis C virus (HCV) infection. *Int J Med Sci,* 3**,** 47-52.

CHENG, J. C., CHANG, M. F. & CHANG, S. C. 1999. Specific interaction between the hepatitis C virus NS5B RNA polymerase and the 3' end of the viral RNA. *J Virol,* 73**,** 7044-9.

CHENG, K. C., GUPTA, S., WANG, H., USS, A. S., NJOROGE, G. F. & HUGHES, E. 2011. Current drug discovery strategies for treatment of hepatitis C virus infection. *J Pharm Pharmacol,* 63**,** 883-92.

CHOO, Q. L., KUO, G., WEINER, A. J., OVERBY, L. R., BRADLEY, D. W. & HOUGHTON, M. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science,* 244**,** 359-62.

CHOO, Q. L., WEINER, A. J., OVERBY, L. R., KUO, G., HOUGHTON, M. & BRADLEY, D. W. 1990. Hepatitis C virus: the major causative agent of viral non-A, non-B hepatitis. *Br Med Bull,* 46**,** 423-41.

CIESEK, S., STEINMANN, E., WEDEMEYER, H., MANNS, M. P., NEYTS, J., TAUTZ, N., MADAN, V., BARTENSCHLAGER, R., VON HAHN, T. & PIETSCHMANN, T. 2009. Cyclosporine A inhibits hepatitis C virus nonstructural protein 2 through cyclophilin A. *Hepatology,* 50**,** 1638-45.

COATS, S. J., GARNIER-AMBLARD, E. C., AMBLARD, F., EHTESHAMI, M., AMIRALAEI, S., ZHANG, H., ZHOU, L., BOUCLE, S. R., LU, X., BONDADA, L., SHELTON, J. R., LI, H., LIU, P., LI, C., CHO, J. H., CHAVRE, S. N., ZHOU, S., MATHEW, J. & SCHINAZI, R. F. 2014. Chutes and ladders in hepatitis C nucleoside drug development. *Antiviral Res,* 102**,** 119-47.

COBURN, G. A., FISCH, D. N., MOORJI, S. M., DE MUYS, J. M., MURGA, J. D., PAUL, D., PROVONCHA, K. P., ROTSHTEYN, Y., HAN, A. Q., QIAN, D., MADDON, P. J. & OLSON, W.

C. 2012. Novel small-molecule inhibitors of hepatitis C virus entry block viral spread and promote viral clearance in cell culture. *PLoS One,* 7**,** e35351.

COELMONT, L., HANOULLE, X., CHATTERJI, U., BERGER, C., SNOECK, J., BOBARDT, M., LIM, P., VLIEGEN, I., PAESHUYSE, J., VUAGNIAUX, G., VANDAMME, A. M., BARTENSCHLAGER, R., GALLAY, P., LIPPENS, G. & NEYTS, J. 2010. DEB025 (Alisporivir) inhibits hepatitis C virus replication by preventing a cyclophilin A induced cis-trans isomerisation in domain II of NS5A. *PLoS One,* 5**,** e13687.

COUNIHAN, N. A., RAWLINSON, S. M. & LINDENBACH, B. D. 2011. Trafficking of hepatitis C virus core protein during virus particle assembly. *PLoS Pathog,* 7**,** e1002302.

CRISTINA, J., LOPEZ, F., MORATORIO, G., LOPEZ, L., VASQUEZ, S., GARCIA-AGUIRRE, L. & CHUNGA, A. 2005. Hepatitis C virus F protein sequence reveals a lack of functional constraints and a variable pattern of amino acid substitution. *J Gen Virol,* 86**,** 115-20.

DAITO, T., WATASHI, K., SLUDER, A., OHASHI, H., NAKAJIMA, S., BORROTO-ESODA, K., FUJITA, T. & WAKITA, T. 2014. Cyclophilin Inhibitors Reduce Phosphorylation of RNA-Dependent Protein Kinase to Restore Expression of IFN-Stimulated Genes in HCV-Infected Cells. *Gastroenterology,* 147**,** 463-72.

DI MARCO, S., VOLPARI, C., TOMEI, L., ALTAMURA, S., HARPER, S., NARJES, F., KOCH, U., ROWLEY, M., DE FRANCESCO, R., MIGLIACCIO, G. & CARFI, A. 2005. Interdomain communication in hepatitis C virus polymerase abolished by small molecule inhibitors bound to a novel allosteric site. *J Biol Chem,* 280**,** 29765-70.

DORNER, M., HORWITZ, J. A., DONOVAN, B. M., LABITT, R. N., BUDELL, W. C., FRILING, T., VOGT, A., CATANESE, M. T., SATOH, T., KAWAI, T., AKIRA, S., LAW, M., RICE, C. M. & PLOSS, A. 2013. Completion of the entire hepatitis C virus life cycle in genetically humanized mice. *Nature,* 501**,** 237-41.

DUFNER-BEATTIE, J., O'GUIN, A., O'GUIN, S., BRILEY, A., WANG, B., BALSAROTTI, J., ROTH, R., STARKEY, G., SLOMCZYNSKA, U., NOUEIRY, A., OLIVO, P. D. & RICE, C. M. 2014. Identification of AP80978, a Novel Small-Molecule Inhibitor of Hepatitis C Virus Replication That Targets NS4B. *Antimicrob Agents Chemother,* 58**,** 3399-3410.

EGGER, D., WOLK, B., GOSERT, R., BIANCHI, L., BLUM, H. E., MORADPOUR, D. & BIENZ, K. 2002. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J Virol,* 76**,** 5974-84.

EINAV, S., GERBER, D., BRYSON, P. D., SKLAN, E. H., ELAZAR, M., MAERKL, S. J., GLENN, J. S. & QUAKE, S. R. 2008. Discovery of a hepatitis C target and its pharmacological inhibitors by microfluidic affinity analysis. *Nat Biotechnol,* 26**,** 1019-27.

EL OMARI, K., IOURIN, O., HARLOS, K., GRIMES, J. M. & STUART, D. I. 2013. Structure of a pestivirus envelope glycoprotein E2 clarifies its role in cell entry. *Cell Rep,* 3**,** 30-5.

EVANS, M. J., VON HAHN, T., TSCHERNE, D. M., SYDER, A. J., PANIS, M., WOLK, B., HATZIIOANNOU, T., MCKEATING, J. A., BIENIASZ, P. D. & RICE, C. M. 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature,* 446**,** 801-5.

FARQUHAR, M. J., HU, K., HARRIS, H. J., DAVIS, C., BRIMACOMBE, C. L., FLETCHER, S. J., BAUMERT, T. F., RAPPOPORT, J. Z., BALFE, P. & MCKEATING, J. A. 2012. Hepatitis C virus induces CD81 and claudin-1 endocytosis. *J Virol,* 86**,** 4305-16.

FEINSTONE, S. M., HU, D. J. & MAJOR, M. E. 2012. Prospects for prophylactic and therapeutic vaccines against hepatitis C virus. *Clin Infect Dis,* 55 Suppl 1**,** S25-32.

FEINSTONE, S. M., KAPIKIAN, A. Z., PURCELL, R. H., ALTER, H. J. & HOLLAND, P. V. 1975. Transfusion-associated hepatitis not due to viral hepatitis type A or B. *N Engl J Med,* 292**,** 767-70.

FENEANT, L., LEVY, S. & COCQUEREL, L. 2014. CD81 and hepatitis C virus (HCV) infection. *Viruses,* 6**,** 535-72.

FERNANDES, F., POOLE, D. S., HOOVER, S., MIDDLETON, R., ANDREI, A. C., GERSTNER, J. & STRIKER, R. 2007. Sensitivity of hepatitis C virus to cyclosporine A depends on nonstructural proteins NS5A and NS5B. *Hepatology,* 46**,** 1026-33.

FEUERSTEIN, S., SOLYOM, Z., ALADAG, A., FAVIER, A., SCHWARTEN, M., HOFFMANN, S., WILLBOLD, D. & BRUTSCHER, B. 2012. Transient structure and SH3 interaction sites in an intrinsically disordered fragment of the hepatitis C virus protein NS5A. *J Mol Biol,* 420**,** 310-23.

FLORES, M. V., STRAWBRIDGE, J., CIARAMELLA, G. & CORBAU, R. 2009. HCV-NS3 inhibitors: determination of their kinetic parameters and mechanism. *Biochim Biophys Acta,* 1794**,** 1441-8.

FOSTER, T. L., BELYAEVA, T., STONEHOUSE, N. J., PEARSON, A. R. & HARRIS, M. 2010a. All three domains of the hepatitis C virus nonstructural NS5A protein contribute to RNA binding. *J Virol,* 84**,** 9267-77.

FOSTER, T. L., GALLAY, P., STONEHOUSE, N. J. & HARRIS, M. 2011. Cyclophilin A interacts with domain II of hepatitis C virus NS5A and stimulates RNA binding in an isomerase-dependent manner. *J Virol,* 85**,** 7460-4.

FOSTER, T. L., TEDBURY, P. R., PEARSON, A. R. & HARRIS, M. 2010b. A comparative analysis of the fluorescence properties of the wild-type and active site mutants of the hepatitis C virus autoprotease NS2-3. *Biochim Biophys Acta,* 1804**,** 212-22.

FOSTER, T. L., THOMPSON, G. S., KALVERDA, A. P., KANKANALA, J., BENTHAM, M., WETHERILL, L. F., THOMPSON, J., BARKER, A. M., CLARKE, D., NOERENBERG, M., PEARSON, A. R., ROWLANDS, D. J., HOMANS, S. W., HARRIS, M., FOSTER, R. & GRIFFIN, S. 2014. Structure-guided design affirms inhibitors of hepatitis C virus p7 as a viable class of antivirals targeting virion release. *Hepatology,* 59**,** 408-22.

FRANCK, N., LE SEYEC, J., GUGUEN-GUILLOUZO, C. & ERDTMANN, L. 2005. Hepatitis C virus NS2 protein is phosphorylated by the protein kinase CK2 and targeted for degradation to the proteasome. *J Virol,* 79**,** 2700-8.

FRANK, C., MOHAMED, M. K., STRICKLAND, G. T., LAVANCHY, D., ARTHUR, R. R., MAGDER, L. S., EL KHOBY, T., ABDEL-WAHAB, Y., ALY OHN, E. S., ANWAR, W. & SALLAM, I. 2000. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet,* 355**,** 887-91.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., MAINZ, D. T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M., PERRY, J. K., SHAW, D. E., FRANCIS, P. & SHENKIN, P. S. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem,* 47**,** 1739-49.

FROMENTIN, R., MAJEAU, N., LALIBERTE GAGNE, M. E., BOIVIN, A., DUVIGNAUD, J. B. & LECLERC, D. 2007. A method for in vitro assembly of hepatitis C virus core protein and for screening of inhibitors. *Anal Biochem,* 366**,** 37-45.

GAO, M., NETTLES, R. E., BELEMA, M., SNYDER, L. B., NGUYEN, V. N., FRIDELL, R. A., SERRANO-WU, M. H., LANGLEY, D. R., SUN, J. H., O'BOYLE, D. R., 2ND, LEMM, J. A., WANG, C., KNIPE, J. O., CHIEN, C., COLONNO, R. J., GRASELA, D. M., MEANWELL, N. A. & HAMANN, L. G. 2010. Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature,* 465**,** 96-100.

GAO, Y., YU, X., XUE, B., ZHOU, F., WANG, X., YANG, D., LIU, N., XU, L., FANG, X. & ZHU, H. 2014. Inhibition of hepatitis C virus infection by DNA aptamer against NS2 protein. *PLoS One,* 9**,** e90333.

GARRY, R. F. & DASH, S. 2003. Proteomics computational analyses suggest that hepatitis C virus E1 and pestivirus E2 envelope glycoproteins are truncated class II fusion proteins. *Virology,* 307**,** 255-65.

GAWLIK, K. & GALLAY, P. A. 2014. HCV core protein and virus assembly: what we know without structures. *Immunol Res*.

GILLET, V. J., NEWELL, W., MATA, P., MYATT, G., SIKE, S., ZSOLDOS, Z. & JOHNSON, A. P. 1994. SPROUT: recent developments in the de novo design of molecules. *J Chem Inf Comput Sci,* 34**,** 207-17.

GORZIN, A. A., RAMSLAND, P. A., TACHEDJIAN, G. & GOWANS, E. J. 2012. Identification of residues involved in NS2 homodimerization and elucidation of their impact on the HCV life cycle. *J Viral Hepat,* 19**,** 189-98.

GOTTIPATI, K., RUGGLI, N., GERBER, M., TRATSCHIN, J. D., BENNING, M., BELLAMY, H. & CHOI, K. H. 2013. The structure of classical swine fever virus N(pro): a novel cysteine Autoprotease and zinc-binding protein involved in subversion of type I interferon induction. *PLoS Pathog,* 9**,** e1003704.

GOUTTENOIRE, J., CASTET, V., MONTSERRET, R., ARORA, N., RAUSSENS, V., RUYSSCHAERT, J. M., DIESIS, E., BLUM, H. E., PENIN, F. & MORADPOUR, D. 2009a. Identification of a novel determinant for membrane association in hepatitis C virus nonstructural protein 4B. *J Virol,* 83**,** 6257-68.

GOUTTENOIRE, J., MONTSERRET, R., KENNEL, A., PENIN, F. & MORADPOUR, D. 2009b. An amphipathic alpha-helix at the C terminus of hepatitis C virus nonstructural protein 4B mediates membrane association. *J Virol,* 83**,** 11378-84.

GRAKOUI, A., MCCOURT, D. W., WYCHOWSKI, C., FEINSTONE, S. M. & RICE, C. M. 1993. A second hepatitis C virus-encoded proteinase. *Proc Natl Acad Sci U S A,* 90**,** 10583-7.

GRIFFIN, S. D., BEALES, L. P., CLARKE, D. S., WORSFOLD, O., EVANS, S. D., JAEGER, J., HARRIS, M. P. & ROWLANDS, D. J. 2003. The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett,* 535**,** 34-8.

GRISE, H., FRAUSTO, S., LOGAN, T. & TANG, H. 2012. A conserved tandem cyclophilin-binding site in hepatitis C virus nonstructural protein 5A regulates Alisporivir susceptibility. *J Virol,* 86**,** 4811-22.

GU, M. & RICE, C. M. 2013. Structures of hepatitis C virus nonstructural proteins required for replicase assembly and function. *Curr Opin Virol,* 3**,** 129-36.

GUO, J. T., BICHKO, V. V. & SEEGER, C. 2001. Effect of alpha interferon on the hepatitis C virus replicon. *J Virol,* 75**,** 8516-23.

HAN, Q., MANNA, D., BELTON, K., COLE, R. & KONAN, K. V. 2013. Modulation of hepatitis C virus genome encapsidation by nonstructural protein 4B. *J Virol,* 87**,** 7409-22.

HAN, W., HU, Z., JIANG, X. & DECICCO, C. P. 2000. Alpha-ketoamides, alpha-ketoesters and alpha-diketones as HCV NS3 protease inhibitors. *Bioorg Med Chem Lett,* 10**,** 711-3.

HARRIS, H. J., FARQUHAR, M. J., MEE, C. J., DAVIS, C., REYNOLDS, G. M., JENNINGS, A., HU, K., YUAN, F., DENG, H., HUBSCHER, S. G., HAN, J. H., BALFE, P. & MCKEATING, J. A. 2008. CD81 and claudin 1 coreceptor association: role in hepatitis C virus entry. *J Virol,* 82**,** 5007-20.

HASHEM, Y., DES GEORGES, A., DHOTE, V., LANGLOIS, R., LIAO, H. Y., GRASSUCCI, R. A., PESTOVA, T. V., HELLEN, C. U. & FRANK, J. 2013. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature,* 503**,** 539-43.

HAWKINS, P. C., SKILLMAN, A. G., WARREN, G. L., ELLINGSON, B. A. & STAHL, M. T. 2010. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model,* 50**,** 572-84.

HEIKKILA, T., THIRUMALAIRAJAN, S., DAVIES, M., PARSONS, M. R., MCCONKEY, A. G., FISHWICK, C. W. & JOHNSON, A. P. 2006. The first de novo designed inhibitors of Plasmodium falciparum dihydroorotate dehydrogenase. *Bioorg Med Chem Lett,* 16**,** 88-92.

HELLE, F., GOFFARD, A., MOREL, V., DUVERLIE, G., MCKEATING, J., KECK, Z. Y., FOUNG, S., PENIN, F., DUBUISSON, J. & VOISSET, C. 2007. The neutralizing activity of anti-hepatitis

C virus antibodies is modulated by specific glycans on the E2 envelope protein. *J Virol,* 81**,** 8101-11.

HELLEN, C. U., KRAUSSLICH, H. G. & WIMMER, E. 1989. Proteolytic processing of polyproteins in the replication of RNA viruses. *Biochemistry,* 28**,** 9881-90.

HIJIKATA, M., MIZUSHIMA, H., AKAGI, T., MORI, S., KAKIUCHI, N., KATO, N., TANAKA, T., KIMURA, K. & SHIMOTOHNO, K. 1993. Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J Virol,* 67**,** 4665-75.

HIROWATARI, Y., HIJIKATA, M. & SHIMOTOHNO, K. 1995. A novel method for analysis of viral proteinase activity encoded by hepatitis C virus in cultured cells. *Anal Biochem,* 225**,** 113-20.

HOOFNAGLE, J. H., MULLEN, K. D., JONES, D. B., RUSTGI, V., DI BISCEGLIE, A., PETERS, M., WAGGONER, J. G., PARK, Y. & JONES, E. A. 1986. Treatment of chronic non-A,non-B hepatitis with recombinant human alpha interferon. A preliminary report. *N Engl J Med,* 315**,** 1575-8.

HUANG, L. 2013. Inhibiting Protease Auto-processing: A Novel Strategy for Anti-HIV-1 Drug Development. *Biochem Physiol 2,* e115.

HUANG, L., HWANG, J., SHARMA, S. D., HARGITTAI, M. R., CHEN, Y., ARNOLD, J. J., RANEY, K. D. & CAMERON, C. E. 2005. Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J Biol Chem,* 280**,** 36417-28.

HUANG, Y., STASCHKE, K., DE FRANCESCO, R. & TAN, S. L. 2007. Phosphorylation of hepatitis C virus NS5A nonstructural protein: a new paradigm for phosphorylation-dependent viral RNA replication? *Virology,* 364**,** 1-9.

HUGHES, M., GRIFFIN, S. & HARRIS, M. 2009. Domain III of NS5A contributes to both RNA replication and assembly of hepatitis C virus particles. *J Gen Virol,* 90**,** 1329-34.

HWANG, J., HUANG, L., CORDEK, D. G., VAUGHAN, R., REYNOLDS, S. L., KIHARA, G., RANEY, K. D., KAO, C. C. & CAMERON, C. E. 2010. Hepatitis C virus nonstructural protein 5A: biochemical characterization of a novel structural class of RNA-binding proteins. *J Virol,* 84**,** 12480-91.

IVANOV, A. V., TUNITSKAYA, V. L., IVANOVA, O. N., MITKEVICH, V. A., PRASSOLOV, V. S., MAKAROV, A. A., KUKHANOVA, M. K. & KOCHETKOV, S. N. 2009. Hepatitis C virus NS5A protein modulates template selection by the RNA polymerase in in vitro system. *FEBS Lett,* 583**,** 277-80.

IVASHKINA, N., WOLK, B., LOHMANN, V., BARTENSCHLAGER, R., BLUM, H. E., PENIN, F. & MORADPOUR, D. 2002. The hepatitis C virus RNA-dependent RNA polymerase membrane insertion sequence is a transmembrane segment. *J Virol,* 76**,** 13088-93.

JENKINS, J. L., KAO, R. Y. & SHAPIRO, R. 2003. Virtual screening to enrich hit lists from high-throughput screening: a case study on small-molecule inhibitors of angiogenin. *Proteins,* 50**,** 81-93.

JIRASKO, V., MONTSERRET, R., APPEL, N., JANVIER, A., EUSTACHI, L., BROHM, C., STEINMANN, E., PIETSCHMANN, T., PENIN, F. & BARTENSCHLAGER, R. 2008. Structural and functional characterization of nonstructural protein 2 for its role in hepatitis C virus assembly. *J Biol Chem,* 283**,** 28546-62.

JIRASKO, V., MONTSERRET, R., LEE, J. Y., GOUTTENOIRE, J., MORADPOUR, D., PENIN, F. & BARTENSCHLAGER, R. 2010. Structural and functional studies of nonstructural protein 2 of the hepatitis C virus reveal its key role as organizer of virion assembly. *PLoS Pathog,* 6**,** e1001233.

JONES, C. T., MURRAY, C. L., EASTMAN, D. K., TASSELLO, J. & RICE, C. M. 2007. Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *J Virol,* 81**,** 8374-83.

KANDA, T., IMAZEKI, F. & YOKOSUKA, O. 2010. New antiviral therapies for chronic hepatitis C. *Hepatol Int,* 4**,** 548-61.

KAPOOR, A., SIMMONDS, P., GEROLD, G., QAISAR, N., JAIN, K., HENRIQUEZ, J. A., FIRTH, C., HIRSCHBERG, D. L., RICE, C. M., SHIELDS, S. & LIPKIN, W. I. 2011. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci U S A,* 108**,** 11608-13.

KAPOOR, A., SIMMONDS, P., SCHEEL, T. K., HJELLE, B., CULLEN, J. M., BURBELO, P. D., CHAUHAN, L. V., DURAISAMY, R., SANCHEZ LEON, M., JAIN, K., VANDEGRIFT, K. J., CALISHER, C. H., RICE, C. M. & LIPKIN, W. I. 2013. Identification of rodent homologs of hepatitis C virus and pegiviruses. *MBio,* 4**,** e00216-13.

KATO, T., DATE, T., MIYAMOTO, M., FURUSAKA, A., TOKUSHIGE, K., MIZOKAMI, M. & WAKITA, T. 2003. Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology,* 125**,** 1808-17.

KAUL, A., STAUFFER, S., BERGER, C., PERTEL, T., SCHMITT, J., KALLIS, S., ZAYAS, M., LOHMANN, V., LUBAN, J. & BARTENSCHLAGER, R. 2009. Essential role of cyclophilin A for hepatitis C virus replication and virus production and possible link to polyprotein cleavage kinetics. *PLoS Pathog,* 5**,** e1000546.

KITCHEN, D. B., DECORNEZ, H., FURR, J. R. & BAJORATH, J. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov,* 3**,** 935-49.

KNODELL, R. G., CONRAD, M. E., DIENSTAG, J. L. & BELL, C. J. 1975. Etiological spectrum of post-transfusion hepatitis. *Gastroenterology,* 69**,** 1278-85.

KOHLER, J. J., NETTLES, J. H., AMBLARD, F., HURWITZ, S. J., BASSIT, L., STANTON, R. A., EHTESHAMI, M. & SCHINAZI, R. F. 2014. Approaches to hepatitis C treatment and cure using NS5A inhibitors. *Infect Drug Resist,* 7**,** 41-56.

KOLYKHALOV, A. A., MIHALIK, K., FEINSTONE, S. M. & RICE, C. M. 2000. Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo. *J Virol,* 74**,** 2046-51.

KREY, T., D'ALAYER, J., KIKUTI, C. M., SAULNIER, A., DAMIER-PIOLLE, L., PETITPAS, I., JOHANSSON, D. X., TAWAR, R. G., BARON, B., ROBERT, B., ENGLAND, P., PERSSON, M. A., MARTIN, A. & REY, F. A. 2010. The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. *PLoS Pathog,* 6**,** e1000762.

KRIEGER, N., LOHMANN, V. & BARTENSCHLAGER, R. 2001. Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J Virol,* 75**,** 4614-24.

KWONG, A. D. 2014. The HCV Revolution Did Not Happen Overnight. *ACS Med Chem Lett,* 5**,** 214-220.

KWONG, A. D., KAUFFMAN, R. S., HURTER, P. & MUELLER, P. 2011. Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nat Biotechnol,* 29**,** 993-1003.

KWONG, A. D., RAO, B. G. & JEANG, K. T. 2005. Viral and cellular RNA helicases as antiviral targets. *Nat Rev Drug Discov,* 4**,** 845-53.

LACKNER, T., MULLER, A., PANKRAZ, A., BECHER, P., THIEL, H. J., GORBALENYA, A. E. & TAUTZ, N. 2004. Temporal modulation of an autoprotease is crucial for replication and pathogenicity of an RNA virus. *J Virol,* 78**,** 10765-75.

LANGE, C. M., BELLECAVE, P., DAO THI, V. L., TRAN, H. T., PENIN, F., MORADPOUR, D. & GOUTTENOIRE, J. 2014. Determinants for Membrane Association of the Hepatitis C Virus NS2 Protease Domain. *J Virol,* 88**,** 6519-23.

LAUER, G. M. & WALKER, B. D. 2001. Hepatitis C virus infection. *N Engl J Med,* 345**,** 41-52.

LEE, C., MA, H., HANG, J. Q., LEVEQUE, V., SKLAN, E. H., ELAZAR, M., KLUMPP, K. & GLENN, J. S. 2011. The hepatitis C virus NS5A inhibitor (BMS-790052) alters the subcellular localization of the NS5A non-structural viral protein. *Virology,* 414**,** 10-8.

LESBURG, C. A., CABLE, M. B., FERRARI, E., HONG, Z., MANNARINO, A. F. & WEBER, P. C. 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol,* 6**,** 937-43.

LEVRERO, M. 2006. Viral hepatitis and liver cancer: the case of hepatitis C. *Oncogene,* 25**,** 3834-47.

LI, Y., WANG, J., KANAI, R. & MODIS, Y. 2013. Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. *Proc Natl Acad Sci U S A,* 110**,** 6805-10.

LIM, P. J., CHATTERJI, U., CORDEK, D., SHARMA, S. D., GARCIA-RIVERA, J. A., CAMERON, C. E., LIN, K., TARGETT-ADAMS, P. & GALLAY, P. A. 2012. Correlation between NS5A dimerization and hepatitis C virus replication. *J Biol Chem,* 287**,** 30861-73.

LINDENBACH, B. D., EVANS, M. J., SYDER, A. J., WOLK, B., TELLINGHUISEN, T. L., LIU, C. C., MARUYAMA, T., HYNES, R. O., BURTON, D. R., MCKEATING, J. A. & RICE, C. M. 2005. Complete replication of hepatitis C virus in cell culture. *Science,* 309**,** 623-6.

LINDENBACH, B. D. & RICE, C. M. 2005. Unravelling hepatitis C virus replication from genome to function. *Nature,* 436**,** 933-8.

LLINAS-BRUNET, M., BAILEY, M., FAZAL, G., GOULET, S., HALMOS, T., LAPLANTE, S., MAURICE, R., POIRIER, M., POUPART, M. A., THIBEAULT, D., WERNIC, D. & LAMARRE, D. 1998. Peptide-based inhibitors of the hepatitis C virus serine protease. *Bioorg Med Chem Lett,* 8**,** 1713-8.

LOHMANN, V. & BARTENSCHLAGER, R. 2014. On the history of hepatitis C virus cell culture systems. *J Med Chem,* 57**,** 1627-42.

LOHMANN, V., KORNER, F., KOCH, J., HERIAN, U., THEILMANN, L. & BARTENSCHLAGER, R. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science,* 285**,** 110-3.

LORENZ, I. C. 2010. The Hepatitis C Virus Nonstructural Protein 2 (NS2): An Up-and-Coming Antiviral Drug Target. *Viruses,* 2**,** 1635-46.

LORENZ, I. C., MARCOTRIGIANO, J., DENTZER, T. G. & RICE, C. M. 2006. Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature,* 442**,** 831-5.

LOVE, R. A., BRODSKY, O., HICKEY, M. J., WELLS, P. A. & CRONIN, C. N. 2009. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *J Virol,* 83**,** 4395-403.

LUNDIN, M., MONNE, M., WIDELL, A., VON HEIJNE, G. & PERSSON, M. A. 2003. Topology of the membrane-associated hepatitis C virus protein NS4B. *J Virol,* 77**,** 5428-38.

MA, Y., ANANTPADMA, M., TIMPE, J. M., SHANMUGAM, S., SINGH, S. M., LEMON, S. M. & YI, M. 2011. Hepatitis C virus NS2 protein serves as a scaffold for virus assembly by interacting with both structural and nonstructural proteins. *J Virol,* 85**,** 86-97.

MA, Y., YATES, J., LIANG, Y., LEMON, S. M. & YI, M. 2008. NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J Virol,* 82**,** 7624-39.

MACDONALD, A. & HARRIS, M. 2004. Hepatitis C virus NS5A: tales of a promiscuous protein. *J Gen Virol,* 85**,** 2485-502.

MADAN, V., PAUL, D., LOHMANN, V. & BARTENSCHLAGER, R. 2014. Inhibition of HCV replication by cyclophilin antagonists is linked to replication fitness and occurs by inhibition of membranous web formation. *Gastroenterology,* 146**,** 1361-72 e1-9.

MANNS, M. P., WEDEMEYER, H. & CORNBERG, M. 2006. Treating viral hepatitis C: efficacy, side effects, and complications. *Gut,* 55**,** 1350-9.

MATSUDA, S. & KOYASU, S. 2000. Mechanisms of action of cyclosporine. *Immunopharmacology,* 47**,** 119-25.

MAZUMDAR, B., BANERJEE, A., MEYER, K. & RAY, R. 2011. Hepatitis C virus E1 envelope glycoprotein interacts with apolipoproteins in facilitating entry into hepatocytes. *Hepatology,* 54**,** 1149-56.

MCGIVERN, D. R., MASAKI, T., WILLIFORD, S., INGRAVALLO, P., FENG, Z., LAHSER, F., ASANTE-APPIAH, E., NEDDERMANN, P., DE FRANCESCO, R., HOWE, A. Y. & LEMON, S. M. 2014. Kinetic Analyses Reveal Potent and Early Blockade of Hepatitis C Virus Assembly by NS5A Inhibitors. *Gastroenterology*.

MCLAUCHLAN, J. 2000. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *J Viral Hepat,* 7**,** 2-14.

MCLAUCHLAN, J., LEMBERG, M. K., HOPE, G. & MARTOGLIO, B. 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *EMBO J,* 21**,** 3980-8.

MEERTENS, L., BERTAUX, C., CUKIERMAN, L., CORMIER, E., LAVILLETTE, D., COSSET, F. L. & DRAGIC, T. 2008. The tight junction proteins claudin-1, -6, and -9 are entry cofactors for hepatitis C virus. *J Virol,* 82**,** 3555-60.

MILLER, M. H., AGARWAL, K., AUSTIN, A., BROWN, A., BARCLAY, S. T., DUNDAS, P., DUSHEIKO, G. M., FOSTER, G. R., FOX, R., HAYES, P. C., LEEN, C., MILLSON, C., RYDER, S. D., TAIT, J., USTIANOWSKI, A. & DILLON, J. F. 2014. Review article: 2014 UK consensus guidelines - hepatitis C management and direct-acting anti-viral therapy. *Aliment Pharmacol Ther*.

MILLER, S. & KRIJNSE-LOCKER, J. 2008. Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol,* 6**,** 363-74.

MIYANARI, Y., ATSUZAWA, K., USUDA, N., WATASHI, K., HISHIKI, T., ZAYAS, M., BARTENSCHLAGER, R., WAKITA, T., HIJIKATA, M. & SHIMOTOHNO, K. 2007. The lipid droplet is an important organelle for hepatitis C virus production. *Nat Cell Biol,* 9**,** 1089-97.

MOK, N. Y., CHADWICK, J., KELLETT, K. A., CASAS-ARCE, E., HOOPER, N. M., JOHNSON, A. P. & FISHWICK, C. W. 2013. Discovery of biphenylacetamide-derived inhibitors of BACE1 using de novo structure-based molecular design. *J Med Chem,* 56**,** 1843-52.

MORICE, Y., RATINIER, M., MILADI, A., CHEVALIEZ, S., GERMANIDIS, G., WEDEMEYER, H., LAPERCHE, S., LAVERGNE, J. P. & PAWLOTSKY, J. M. 2009. Seroconversion to hepatitis C virus alternate reading frame protein during acute infection. *Hepatology,* 49**,** 1449-59.

MORIKAWA, K., LANGE, C. M., GOUTTENOIRE, J., MEYLAN, E., BRASS, V., PENIN, F. & MORADPOUR, D. 2011. Nonstructural protein 3-4A: the Swiss army knife of hepatitis C virus. *J Viral Hepat,* 18**,** 305-15.

MURPHY, D. G., WILLEMS, B., DESCHENES, M., HILZENRAT, N., MOUSSEAU, R. & SABBAH, S. 2007. Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5' untranslated region sequences. *J Clin Microbiol,* 45**,** 1102-12.

MURRAY, C. L., JONES, C. T. & RICE, C. M. 2008. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat Rev Microbiol,* 6**,** 699-708.

NADIN, A., HATTOTUWAGAMA, C. & CHURCHER, I. 2012. Lead-oriented synthesis: a new opportunity for synthetic chemistry. *Angew Chem Int Ed Engl,* 51**,** 1114-22.

NAKABAYASHI, H., TAKETA, K., MIYANO, K., YAMANE, T. & SATO, J. 1982. Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res,* 42**,** 3858-63.

NAKAGAWA, M., SAKAMOTO, N., ENOMOTO, N., TANABE, Y., KANAZAWA, N., KOYAMA, T., KUROSAKI, M., MAEKAWA, S., YAMASHIRO, T., CHEN, C. H., ITSUI, Y., KAKINUMA, S. & WATANABE, M. 2004. Specific inhibition of hepatitis C virus replication by cyclosporin A. *Biochem Biophys Res Commun,* 313**,** 42-7.

NEDDERMANN, P., CLEMENTI, A. & DE FRANCESCO, R. 1999. Hyperphosphorylation of the hepatitis C virus NS5A protein requires an active NS3 protease, NS4A, NS4B, and NS5A encoded on the same polyprotein. *J Virol,* 73**,** 9984-91.

NI, F., KOTA, S., TAKAHASHI, V., STROSBERG, A. D. & SNYDER, J. K. 2011. Potent inhibitors of hepatitis C core dimerization as new leads for anti-hepatitis C agents. *Bioorg Med Chem Lett,* 21**,** 2198-202.

NIELSEN, S. U., BASSENDINE, M. F., BURT, A. D., MARTIN, C., PUMEECHOCKCHAI, W. & TOMS, G. L. 2006. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. *J Virol,* 80**,** 2418-28.

NUNEZ, M. & SORIANO, V. 2005. Hepatitis C virus (HCV) genotypes and disease progression in HIV/HCV-coinfected patients. *J Infect Dis,* 191**,** 1-3.

OU-YANG, S. S., LU, J. Y., KONG, X. Q., LIANG, Z. J., LUO, C. & JIANG, H. 2012. Computational drug discovery. *Acta Pharmacol Sin,* 33**,** 1131-40.

PAESHUYSE, J., KAUL, A., DE CLERCQ, E., ROSENWIRTH, B., DUMONT, J. M., SCALFARO, P., BARTENSCHLAGER, R. & NEYTS, J. 2006. The non-immunosuppressive cyclosporin DEBIO-025 is a potent inhibitor of hepatitis C virus replication in vitro. *Hepatology,* 43**,** 761-70.

PALLAORO, M., LAHM, A., BIASIOL, G., BRUNETTI, M., NARDELLA, C., ORSATTI, L., BONELLI, F., ORRU, S., NARJES, F. & STEINKUHLER, C. 2001. Characterization of the hepatitis C virus NS2/3 processing reaction by using a purified precursor protein. *J Virol,* 75**,** 9939-46.

PAUL, D., ROMERO-BREY, I., GOUTTENOIRE, J., STOITSOVA, S., KRIJNSE-LOCKER, J., MORADPOUR, D. & BARTENSCHLAGER, R. 2011. NS4B self-interaction through conserved C-terminal elements is required for the establishment of functional hepatitis C virus replication complexes. *J Virol,* 85**,** 6963-76.

PAVLOS, R. & PHILLIPS, E. J. 2012. Individualization of antiretroviral therapy. *Pharmgenomics Pers Med,* 5**,** 1-17.

PENIN, F., DUBUISSON, J., REY, F. A., MORADPOUR, D. & PAWLOTSKY, J. M. 2004. Structural biology of hepatitis C virus. *Hepatology,* 39**,** 5-19.

PHAN, T., KOHLWAY, A., DIMBERU, P., PYLE, A. M. & LINDENBACH, B. D. 2011. The acidic domain of hepatitis C virus NS4A contributes to RNA replication and virus particle assembly. *J Virol,* 85**,** 1193-204.

PIERONI, L., SANTOLINI, E., FIPALDINI, C., PACINI, L., MIGLIACCIO, G. & LA MONICA, N. 1997. In vitro study of the NS2-3 protease of hepatitis C virus. *J Virol,* 71**,** 6373-80.

PIETSCHMANN, T., KAUL, A., KOUTSOUDAKIS, G., SHAVINSKAYA, A., KALLIS, S., STEINMANN, E., ABID, K., NEGRO, F., DREUX, M., COSSET, F. L. & BARTENSCHLAGER, R. 2006. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc Natl Acad Sci U S A,* 103**,** 7408-13.

PILERI, P., UEMATSU, Y., CAMPAGNOLI, S., GALLI, G., FALUGI, F., PETRACCA, R., WEINER, A. J., HOUGHTON, M., ROSA, D., GRANDI, G. & ABRIGNANI, S. 1998. Binding of hepatitis C virus to CD81. *Science,* 282**,** 938-41.

PLOSS, A., EVANS, M. J., GAYSINSKAYA, V. A., PANIS, M., YOU, H., DE JONG, Y. P. & RICE, C. M. 2009. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature,* 457**,** 882-6.

PLOSS, A., KHETANI, S. R., JONES, C. T., SYDER, A. J., TREHAN, K., GAYSINSKAYA, V. A., MU, K., RITOLA, K., RICE, C. M. & BHATIA, S. N. 2010. Persistent hepatitis C virus infection in microscale primary human hepatocyte cultures. *Proc Natl Acad Sci U S A,* 107**,** 3141-5.

PODEVIN, P., CARPENTIER, A., PENE, V., AOUDJEHANE, L., CARRIERE, M., ZAIDI, S., HERNANDEZ, C., CALLE, V., MERITET, J. F., SCATTON, O., DREUX, M., COSSET, F. L., WAKITA, T., BARTENSCHLAGER, R., DEMIGNOT, S., CONTI, F., ROSENBERG, A. R. & CALMUS, Y. 2010. Production of infectious hepatitis C virus in primary cultures of human adult hepatocytes. *Gastroenterology,* 139**,** 1355-64.

POPESCU, C. I., CALLENS, N., TRINEL, D., ROINGEARD, P., MORADPOUR, D., DESCAMPS, V., DUVERLIE, G., PENIN, F., HELIOT, L., ROUILLE, Y. & DUBUISSON, J. 2011. NS2 protein of

hepatitis C virus interacts with structural and non-structural proteins towards virus assembly. *PLoS Pathog,* 7**,** e1001278.

POWERS, J. C., ASGIAN, J. L., EKICI, O. D. & JAMES, K. E. 2002. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev,* 102**,** 4639-750.

PREMKUMAR, A., WILSON, L., EWART, G. D. & GAGE, P. W. 2004. Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride. *FEBS Lett,* 557**,** 99-103.

QIN, W., LUO, H., NOMURA, T., HAYASHI, N., YAMASHITA, T. & MURAKAMI, S. 2002. Oligomeric interaction of hepatitis C virus NS5B is critical for catalytic activity of RNA-dependent RNA polymerase. *J Biol Chem,* 277**,** 2132-7.

QIU, D., LEMM, J. A., O'BOYLE, D. R., 2ND, SUN, J. H., NOWER, P. T., NGUYEN, V., HAMANN, L. G., SNYDER, L. B., DEON, D. H., RUEDIGER, E., MEANWELL, N. A., BELEMA, M., GAO, M. & FRIDELL, R. A. 2011. The effects of NS5A inhibitors on NS5A phosphorylation, polyprotein processing and localization. *J Gen Virol,* 92**,** 2502-11.

RAI, R. & DEVAL, J. 2011. New opportunities in anti-hepatitis C virus drug discovery: targeting NS4B. *Antiviral Res,* 90**,** 93-101.

REED, K. E., GRAKOUI, A. & RICE, C. M. 1995. Hepatitis C virus-encoded NS2-3 protease: cleavage-site mutagenesis and requirements for bimolecular cleavage. *J Virol,* 69**,** 4127-36.

RICE, C. M. 2011. New insights into HCV replication: potential antiviral targets. *Top Antivir Med,* 19**,** 117-20.

ROMANO, K. P., ALI, A., AYDIN, C., SOUMANA, D., OZEN, A., DEVEAU, L. M., SILVER, C., CAO, H., NEWTON, A., PETROPOULOS, C. J., HUANG, W. & SCHIFFER, C. A. 2012. The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors. *PLoS Pathog,* 8**,** e1002832.

ROSENQUIST, A., SAMUELSSON, B., JOHANSSON, P. O., CUMMINGS, M. D., LENZ, O., RABOISSON, P., SIMMEN, K., VENDEVILLE, S., DE KOCK, H., NILSSON, M., HORVATH, A., KALMEIJER, R., DE LA ROSA, G. & BEUMONT-MAUVIEL, M. 2014. Discovery and development of simeprevir (TMC435), a HCV NS3/4A protease inhibitor. *J Med Chem,* 57**,** 1673-93.

ROSNOBLET, C., FRITZINGER, B., LEGRAND, D., LAUNAY, H., WIERUSZESKI, J. M., LIPPENS, G. & HANOULLE, X. 2012. Hepatitis C virus NS5B and host cyclophilin A share a common binding site on NS5A. *J Biol Chem,* 287**,** 44249-60.

ROSS-THRIEPLAND, D., AMAKO, Y. & HARRIS, M. 2013. The C terminus of NS5A domain II is a key determinant of hepatitis C virus genome replication, but is not required for virion assembly and release. *J Gen Virol,* 94**,** 1009-18.

ROSS-THRIEPLAND, D. & HARRIS, M. 2014. Insights into the complexity and functionality of hepatitis C virus NS5A phosphorylation. *J Virol,* 88**,** 1421-32.

RUSH, T. S., 3RD, GRANT, J. A., MOSYAK, L. & NICHOLLS, A. 2005. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J Med Chem,* 48**,** 1489-95.

SALAM, K. A. & AKIMITSU, N. 2013. Hepatitis C virus NS3 inhibitors: current and future perspectives. *Biomed Res Int,* 2013**,** 467869.

SCARSELLI, E., ANSUINI, H., CERINO, R., ROCCASECCA, R. M., ACALI, S., FILOCAMO, G., TRABONI, C., NICOSIA, A., CORTESE, R. & VITELLI, A. 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J,* 21**,** 5017-25.

SCHIEFER, I. T., TAPADAR, S., LITOSH, V., SIKLOS, M., SCISM, R., WIJEWICKRAMA, G. T., CHANDRASENA, E. P., SINHA, V., TAVASSOLI, E., BRUNSTEINER, M., FA, M., ARANCIO, O., PETUKHOV, P. & THATCHER, G. R. 2013. Design, synthesis, and optimization of

novel epoxide incorporating peptidomimetics as selective calpain inhibitors. *J Med Chem,* 56**,** 6054-68.

SCHMIDT-MENDE, J., BIECK, E., HUGLE, T., PENIN, F., RICE, C. M., BLUM, H. E. & MORADPOUR, D. 2001. Determinants for membrane association of the hepatitis C virus RNA-dependent RNA polymerase. *J Biol Chem,* 276**,** 44052-63.

SCHNEIDER, G. & BOHM, H. J. 2002. Virtual screening and fast automated docking methods. *Drug Discov Today,* 7**,** 64-70.

SCHREGEL, V., JACOBI, S., PENIN, F. & TAUTZ, N. 2009. Hepatitis C virus NS2 is a protease stimulated by cofactor domains in NS3. *Proc Natl Acad Sci U S A,* 106**,** 5342-7.

SCHUTZ-GESHWENDER, A., ZHANG, Y., HOLT, T., MCDERMITT, D., MICHAEL OLIVE, D. 2004. Quantitative, Two-Color Western Blot Detection With Infrared Fluorescence.

SHARMA, N. R., MATEU, G., DREUX, M., GRAKOUI, A., COSSET, F. L. & MELIKYAN, G. B. 2011. Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *J Biol Chem,* 286**,** 30361-76.

SHEPARD, C. W., FINELLI, L. & ALTER, M. J. 2005. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis,* 5**,** 558-67.

SHIMAKAMI, T., WELSCH, C., YAMANE, D., MCGIVERN, D. R., YI, M., ZEUZEM, S. & LEMON, S. M. 2011. Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. *Gastroenterology,* 140**,** 667-75.

SHIMAKAMI, T., YAMANE, D., JANGRA, R. K., KEMPF, B. J., SPANIEL, C., BARTON, D. J. & LEMON, S. M. 2012. Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc Natl Acad Sci U S A,* 109**,** 941-6.

SHIROTA, Y., LUO, H., QIN, W., KANEKO, S., YAMASHITA, T., KOBAYASHI, K. & MURAKAMI, S. 2002. Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. *J Biol Chem,* 277**,** 11149-55.

SIMMONDS, P. 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. *J Gen Virol,* 85**,** 3173-88.

SINGH, J., PETTER, R. C., BAILLIE, T. A. & WHITTY, A. 2011. The resurgence of covalent drugs. *Nat Rev Drug Discov,* 10**,** 307-17.

SLIWOSKI, G., KOTHIWALE, S., MEILER, J. & LOWE, E. W., JR. 2014. Computational methods in drug discovery. *Pharmacol Rev,* 66**,** 334-95.

STAPLEFORD, K. A. & LINDENBACH, B. D. 2011. Hepatitis C virus NS2 coordinates virus particle assembly through physical interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. *J Virol,* 85**,** 1706-17.

SUMPTER, R., JR., LOO, Y. M., FOY, E., LI, K., YONEYAMA, M., FUJITA, T., LEMON, S. M. & GALE, M., JR. 2005. Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *J Virol,* 79**,** 2689-99.

SUZUKI, R., MATSUDA, M., WATASHI, K., AIZAKI, H., MATSUURA, Y., WAKITA, T. & SUZUKI, T. 2013. Signal peptidase complex subunit 1 participates in the assembly of hepatitis C virus through an interaction with E2 and NS2. *PLoS Pathog,* 9**,** e1003589.

TAI, C. L., CHI, W. K., CHEN, D. S. & HWANG, L. H. 1996. The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *J Virol,* 70**,** 8477-84.

TANG, H. & GRISE, H. 2009. Cellular and molecular biology of HCV infection and hepatitis. *Clin Sci (Lond),* 117**,** 49-65.

TARGETT-ADAMS, P., GRAHAM, E. J., MIDDLETON, J., PALMER, A., SHAW, S. M., LAVENDER, H., BRAIN, P., TRAN, T. D., JONES, L. H., WAKENHUT, F., STAMMEN, B., PRYDE, D., PICKFORD, C. & WESTBY, M. 2011. Small molecules targeting hepatitis C virus-encoded NS5A cause subcellular redistribution of their target: insights into compound modes of action. *J Virol,* 85**,** 6353-68.

TARGETT-ADAMS, P. & MCLAUCHLAN, J. 2005. Development and characterization of a transient-replication assay for the genotype 2a hepatitis C virus subgenomic replicon. *J Gen Virol,* 86**,** 3075-80.

TEDBURY, P., WELBOURN, S., PAUSE, A., KING, B., GRIFFIN, S. & HARRIS, M. 2011. The subcellular localization of the hepatitis C virus non-structural protein NS2 is regulated by an ion channel-independent function of the p7 protein. *J Gen Virol,* 92**,** 819-30.

TEDBURY, P. R. & HARRIS, M. 2007. Characterisation of the role of zinc in the hepatitis C virus NS2/3 auto-cleavage and NS3 protease activities. *J Mol Biol,* 366**,** 1652-60.

TELLINGHUISEN, T. L., FOSS, K. L., TREADAWAY, J. C. & RICE, C. M. 2008. Identification of residues required for RNA replication in domains II and III of the hepatitis C virus NS5A protein. *J Virol,* 82**,** 1073-83.

TELLINGHUISEN, T. L., MARCOTRIGIANO, J., GORBALENYA, A. E. & RICE, C. M. 2004. The NS5A protein of hepatitis C virus is a zinc metalloprotein. *J Biol Chem,* 279**,** 48576-87.

TELLINGHUISEN, T. L., MARCOTRIGIANO, J. & RICE, C. M. 2005. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature,* 435**,** 374-9.

THIBEAULT, D., MAURICE, R., PILOTE, L., LAMARRE, D. & PAUSE, A. 2001. In vitro characterization of a purified NS2/3 protease variant of hepatitis C virus. *J Biol Chem,* 276**,** 46678-84.

TIMPE, J. M., STAMATAKI, Z., JENNINGS, A., HU, K., FARQUHAR, M. J., HARRIS, H. J., SCHWARZ, A., DESOMBERE, I., ROELS, G. L., BALFE, P. & MCKEATING, J. A. 2008. Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology,* 47**,** 17-24.

TONG, L., WENGLER, G. & ROSSMANN, M. G. 1993. Refined structure of Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures. *J Mol Biol,* 230**,** 228-47.

TOTROV, M. & ABAGYAN, R. 1997. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins,* Suppl 1**,** 215-20.

TSCHERNE, D. M., JONES, C. T., EVANS, M. J., LINDENBACH, B. D., MCKEATING, J. A. & RICE, C. M. 2006. Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J Virol,* 80**,** 1734-41.

VASSILAKI, N., FRIEBE, P., MEULEMAN, P., KALLIS, S., KAUL, A., PARANHOS-BACCALA, G., LEROUX-ROELS, G., MAVROMARA, P. & BARTENSCHLAGER, R. 2008. Role of the hepatitis C virus core+1 open reading frame and core cis-acting RNA elements in viral RNA translation and replication. *J Virol,* 82**,** 11503-15.

VIEYRES, G., DUBUISSON, J. & PIETSCHMANN, T. 2014. Incorporation of hepatitis C virus E1 and E2 glycoproteins: the keystones on a peculiar virion. *Viruses,* 6**,** 1149-87.

VIEYRES, G. & PIETSCHMANN, T. 2013. Entry and replication of recombinant hepatitis C viruses in cell culture. *Methods,* 59**,** 233-48.

VISHNUVARDHAN, D., KAKIUCHI, N., URVIL, P. T., SHIMOTOHNO, K., KUMAR, P. K. & NISHIKAWA, S. 1997. Expression of highly active recombinant NS3 protease domain of hepatitis C virus in E. coli. *FEBS Lett,* 402**,** 209-12.

VON DEM BUSSCHE, A., MACHIDA, R., LI, K., LOEVINSOHN, G., KHANDER, A., WANG, J., WAKITA, T., WANDS, J. R. & LI, J. 2010. Hepatitis C virus NS2 protein triggers endoplasmic reticulum stress and suppresses its own viral replication. *J Hepatol,* 53**,** 797-804.

VON SCHAEWEN, M. & PLOSS, A. 2014. Murine models of hepatitis C: what can we look forward to? *Antiviral Res,* 104**,** 15-22.

WATASHI, K., HIJIKATA, M., HOSAKA, M., YAMAJI, M. & SHIMOTOHNO, K. 2003. Cyclosporin A suppresses replication of hepatitis C virus genome in cultured hepatocytes. *Hepatology,* 38**,** 1282-8.

WAXMAN, L., WHITNEY, M., POLLOK, B. A., KUO, L. C. & DARKE, P. L. 2001. Host cell factor requirement for hepatitis C virus enzyme maturation. *Proc Natl Acad Sci U S A,* 98**,** 13931-5.

WELBOURN, S., GREEN, R., GAMACHE, I., DANDACHE, S., LOHMANN, V., BARTENSCHLAGER, R., MEEROVITCH, K. & PAUSE, A. 2005. Hepatitis C virus NS2/3 processing is required for NS3 stability and viral RNA replication. *J Biol Chem,* 280**,** 29604-11.

WELBOURN, S., JIRASKO, V., BRETON, V., REISS, S., PENIN, F., BARTENSCHLAGER, R. & PAUSE, A. 2009. Investigation of a role for lysine residues in non-structural proteins 2 and 2/3 of the hepatitis C virus for their degradation and virus assembly. *J Gen Virol,* 90**,** 1071-80.

WELBOURN, S. & PAUSE, A. 2006. HCV NS2/3 Protease. *In:* TAN, S. L. (ed.) *Hepatitis C Viruses: Genomes and Molecular Biology.* Norfolk UK: Horizon Bioscience.

WHITNEY, M., STACK, J. H., DARKE, P. L., ZHENG, W., TERZO, J., INGLESE, J., STRULOVICI, B., KUO, L. C. & POLLOCK, B. A. 2002. A collaborative screening program for the discovery of inhibitors of HCV NS2/3 cis-cleaving protease activity. *J Biomol Screen,* 7**,** 149-54.

WHO 2009. Global distribution of HCV genotypes.

WOZNIAK, A. L., GRIFFIN, S., ROWLANDS, D., HARRIS, M., YI, M., LEMON, S. M. & WEINMAN, S. A. 2010. Intracellular proton conductance of the hepatitis C virus p7 protein and its contribution to infectious virus production. *PLoS Pathog,* 6**,** e1001087.

WU, Z., YAO, N., LE, H. V. & WEBER, P. C. 1998. Mechanism of autoproteolysis at the NS2-NS3 junction of the hepatitis C virus polyprotein. *Trends Biochem Sci,* 23**,** 92-4.

WYLES, D. L., KAIHARA, K. A., KORBA, B. E., SCHOOLEY, R. T., BEADLE, J. R. & HOSTETLER, K. Y. 2009. The octadecyloxyethyl ester of (S)-9-[3-hydroxy-2-(phosphonomethoxy) propyl]adenine is a potent and selective inhibitor of hepatitis C virus replication in genotype 1A, 1B, and 2A replicons. *Antimicrob Agents Chemother,* 53**,** 2660-2.

WYLES, D. L., KAIHARA, K. A., VAIDA, F. & SCHOOLEY, R. T. 2007. Synergy of small molecular inhibitors of hepatitis C virus replication directed at multiple viral targets. *J Virol,* 81**,** 3005-8.

XU, Z., CHOI, J., YEN, T. S., LU, W., STROHECKER, A., GOVINDARAJAN, S., CHIEN, D., SELBY, M. J. & OU, J. 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J,* 20**,** 3840-8.

YAMAGA, A. K. & OU, J. H. 2002. Membrane topology of the hepatitis C virus NS2 protein. *J Biol Chem,* 277**,** 33228-34.

YAMASHITA, T., KANEKO, S., SHIROTA, Y., QIN, W., NOMURA, T., KOBAYASHI, K. & MURAKAMI, S. 1998. RNA-dependent RNA polymerase activity of the soluble recombinant hepatitis C virus NS5B protein truncated at the C-terminal region. *J Biol Chem,* 273**,** 15479-86.

YANG, F., ROBOTHAM, J. M., GRISE, H., FRAUSTO, S., MADAN, V., ZAYAS, M., BARTENSCHLAGER, R., ROBINSON, M., GREENSTEIN, A. E., NAG, A., LOGAN, T. M., BIENKIEWICZ, E. & TANG, H. 2010. A major determinant of cyclophilin dependence and cyclosporine susceptibility of hepatitis C virus identified by a genetic approach. *PLoS Pathog,* 6**,** e1001118.

YANG, F., ROBOTHAM, J. M., NELSON, H. B., IRSIGLER, A., KENWORTHY, R. & TANG, H. 2008. Cyclophilin A is an essential cofactor for hepatitis C virus infection and the principal mediator of cyclosporine resistance in vitro. *J Virol,* 82**,** 5269-78.

YI, M., MA, Y., YATES, J. & LEMON, S. M. 2009. Trans-complementation of an NS2 defect in a late step in hepatitis C virus (HCV) particle assembly and maturation. *PLoS Pathog,* 5**,** e1000403.

ZHONG, J., GASTAMINZA, P., CHENG, G., KAPADIA, S., KATO, T., BURTON, D. R., WIELAND, S. F., UPRICHARD, S. L., WAKITA, T. & CHISARI, F. V. 2005. Robust hepatitis C virus infection in vitro. *Proc Natl Acad Sci U S A,* 102**,** 9294-9.

ZHU, Y. Z., QIAN, X. J., ZHAO, P. & QI, Z. T. 2014. How hepatitis C virus invades hepatocytes: The mystery of viral entry. *World J Gastroenterol,* 20**,** 3457-3467.

ZSOLDOS, Z., REID, D., SIMON, A., SADJAD, B. S. & JOHNSON, A. P. 2006. eHiTS: an innovative approach to the docking and scoring function problems. *Curr Protein Pept Sci,* 7**,** 421-35.

Chapter 8 - Appendix

Synthetic procedures

1-[(1-Methylcyclohexyl)carbonyl]-spiro[indene-1,4'-piperidine]

(JS-160-1)
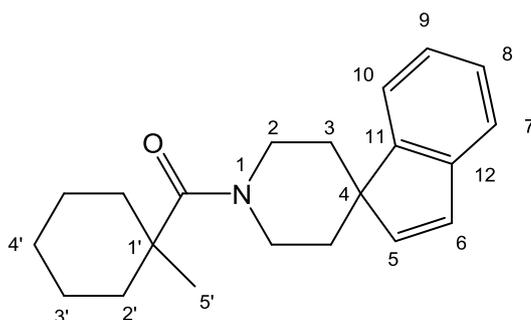
1-Hydroxybenzotriazole (457.1 mg, 3.383 mmol) and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (648.5 mg, 3.383 mmol) were added to a stirred solution of 1-methyl-1-cyclohexanecarboxylic acid (352.8 mg, 2.481 mmol) in DMF (12.5 ml), and the resulting solution was stirred at room temperature for 1 hour. 4-Spiroindene-piperidine hydrochloride (0.5 g, 2.255 mmol) and TEA (376 µl, 2.706 mmol) were then ad ded and stirred at room temperature for 16 hours. Following evaporation of the solvent the resulting residue was partitioned between EtOAc (20 ml) and water (2 x 20 ml) followed by brine (2 x 20ml) and the organic extracts were dried (MgSO$_4$) before the solvents were removed under reduced pressure. The resulting solid was purified via column chromatography to yield the <u>title compound</u> (294 mg, 0.950 mmol, 42%) as a colourless powder; **m.p** 116-118 °C ; **R$_f$** 0.88 (10% MeOH/DCM); **δ$_H$** (500 MHz, CDCl$_3$); 7.30 (m, 4H), 6.86 (dd, (J = 24.1, 5.7, 2H), 4.54 (d, J = 14.3, 2H), 3.20 (d, J = 13.0, 2H), 2.20-1.89 (m, 4H), 1.59 (m, 5H), 1.52-1.12 (m, 10H); **δ$_C$** (75 MHz, CDCl$_3$); 176.07 (CO), 151.37 (12-C), 142.79 (11-C), 140.47 (5-C), 130.41 (6-C), 127.13 (8-C), 125.46 (9-C), 121.66 and 121.55 (10-C and 7-C), 52.35 (4-C), 44.25 (2-C), 42.74 (1'-C), 37.22 (2'-C), 33.95 (3-C), 29.71 (4'-C), 25.99 (5'-C), 23.18 (3'-C); ***m/z* (ES)**; (found MH$^+$ 310.224; C$_{21}$H$_{27}$NO requires *MH* 310.21); **LC-MS** (RT= 2.15 min, *m/z* found MH$^+$ 310.2).

1'-[(Cyclohexyl)carbonyl]-spiro[indene-1,4'-piperidine]

(JS-160-2)
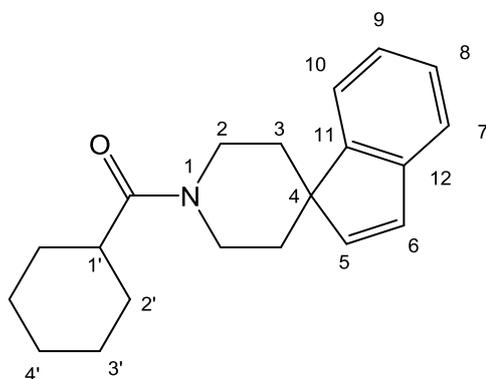
1-Hydroxybenzotriazole (182.8 mg, 1.353 mmol) and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (259.4 mg, 1.353 mmol) were added to a stirred solution of cyclohexane carboxylic acid (123 µl, 0.992 mmol) in DMF (5 ml). The solution was stirred at room temperature for 15 minutes. 4-Spiroindene-piperidine hydrochloride (200 mg, 0.902 mmol) and TEA (150 µl, 1.082 mmol) were then added and the resulting solution was stirred at room temperature for 16 hours. Following evaporation of the solvent the resulting residue was partitioned between EtOAc (20 ml) and water (2 x 20 ml) followed by brine (2 x 20ml) and the organic extracts were dried (MgSO$_4$) before the solvents were removed under reduced pressure. The resulting solid was purified via column chromatography to yield the <u>title compound</u> (148.8 mg, 0.504 mmol, 56%) as a colourless powder; **m.p** 134-135 °C ; **R$_f$** 0.80 (10% MeOH/DCM); **δ$_H$** (500 MHz, CDCl$_3$); 7.36-7.18 (m, 4H), 6.84 (dd, (J = 27.0, 5.7 , 2H), 4.70 (d, *J* = 13.7, 1H), 4.04 (d, *J* = 13.3, 1H), 3.40 (t, *J* = 12.3, 1H), 3.00 (t, *J* = 12.2, 1H), 2.56 (t, *J* = 11.5, 1H), 2.05-1.92 (m, 2H), 1.81 (m, 4H), 1.72 (s, 1H), 1.68-1.36 (m, 5H), 1.29 (m, 3H); **δ$_C$** (75 MHz, CDCl$_3$); 174.85 (CO), 151.30 (12-C), 142.76 (11-C), 140.12 (5-C), 130.59 (6-C), 127.17 (8-C), 125.49 (9-C), 121.68 and 121.58 (10-C and 7-C), 52.27 (4-C), 44.35 (2-C), 40.64 (1'-C), 34.58 (2'-C), 33.33 (3-C), 29.67 (4'-C), 25.94 (3'-C); ***m/z* (ES)**; (found MH$^+$ 296.192; C$_{20}$H$_{25}$NO requires *MH* 296.2); **LC-MS** (RT= 2.20 min, *m/z* found MH$^+$ 296.2).

1'-[(1-Methylcyclohexyl)carbonyl]-spiro[indeno[2,3-b]oxirene-1,4'-piperidine]
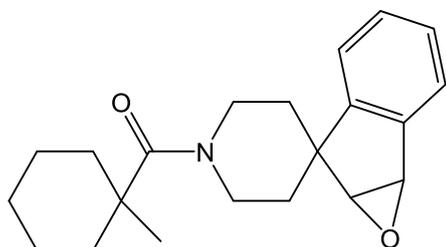
JS-160-1.2

Sodium bicarbonate (162.8 mg, 1.938 mmol) and meta-chloroperbenzoic acid (77%, 202.8 mg, 1.175 mmol) were added to 1-[(1-methylcyclohexyl)carbonyl]-2,3-dihydrospiro[indene]-1,4'-piperidine] (JS-160-1) (200 mg, 0.646 mmol) in DCM (7 ml) and the resulting mixture was stirred for 10 minutes at 0 °C. The solution was allowed to warm to room temperature and then stirred for 16 hours. Following evaporation of the solvent the resulting residue was partitioned between DCM (20 ml) and sodium sulphate solution (10% (w/v), 20 ml) and the solvents were removed under reduced pressure. The resulting brown residue (70.6 mg) proved insoluble in all tested solvents preventing spectroscopic analysis. Title compound was identified by mass spectrometry; **LC-MS** (RT= 2.05 min, *m/z* found MH^+ 326.2, $C_{21}H_{27}NO_2$ requires *MH* 326.21).

1'-[(1-Methylcyclohexyl)carbonyl]-2,3-dihydrospiro[indene-1,4'-piperidine]-2-ol, 3-amine

JS-160-1.3
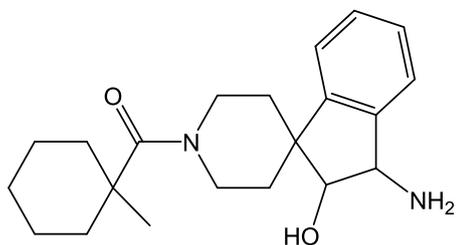
Lithium perchlorate (22.9 mg, 0.215 mmol) and aqueous ammonia (35%) (2 ml) were added to a solution of 1'-[(1-methylcyclohexyl)carbonyl]-spiro[indeno[2,3-b]oxirene-1,4'-piperidine] (JS-160-1.2) (70 mg, 0.215 mmol) in THF (2 ml) and the resulting solution was stirred at 125 °C at 220 psi for 45 minutes. Following evaporation of the solvent the resulting residue was partitioned between DCM (20 ml) and water (2 x 20 ml) followed by brine (2 x 20 ml) and the organic extracts were dried (MgSO$_4$) before the solvents were removed under reduced pressure. The resulting yellow-green powder proved insufficient for further purification or spectroscopic analysis (6 mg, 0.018 mmol, 8%). Title compound was identified by mass spectrometry; *m/z* **(ES)**; (found MH$^+$ 344.222; C$_{21}$H$_{30}$N$_2$O$_2$ requires *MH* 343.23); **LC-MS** (RT= 1.60 min, *m/z* found MH$^+$ 343.2).

(2R, 3R)-3-amino-1'-[(1-methylcyclohexyl) carbonyl]-2,3-dihydrospiro[indene-1.4'-piperidin]-2-ol
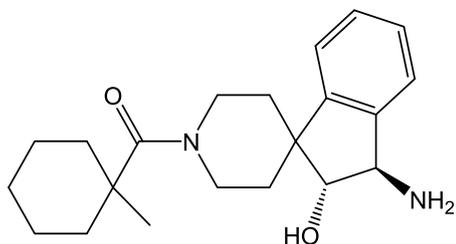
(JS-160)

Title compound was from Chembridge 'screening library' as a white film; *m/z* **(ES)**; (found MH$^+$ 343.239; $C_{21}H_{30}N_2O_2$ requires *MH* 343.23)

(2R, 3R)-3-amino-1'-(3-methylbutanoyl)-2,3-dihydrospiro[indene-1,4'-piperidin]-2-ol

(JS-160-3)
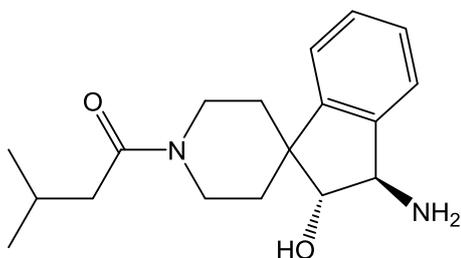
Title compound was from Chembridge 'screening library' as a white-yellow film; *m/z* **(ES)**; (found MH$^+$ 303.207; $C_{18}H_{26}N_2O_2$ requires *MH* 303.2)

(2R, 3R)-3-amino-1'-glycoloyl-2,3-dihydrospiro[indene-1,4'-piperidin]-2-ol

(JS-160-4)
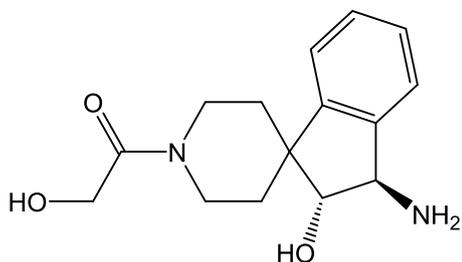
Title compound was from Chembridge 'screening library' as a white film; ***m/z* (ES)**; (found MH$^+$ 277.156; $C_{15}H_{20}N_2O_3$ requires *MH* 277.15)

(2R,3R)-3-amino-1'-(3-pyridinylcarbonyl)-2,3-dihydrospiro[indene-1,4'-piperidin]-2-ol

(JS-160-5)
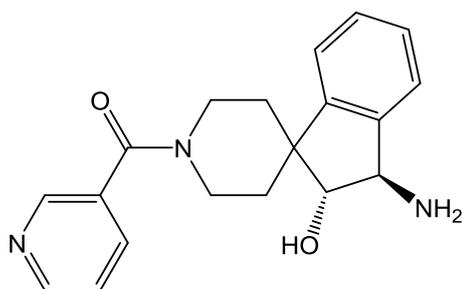
Title compound was from Chembridge 'screening library' as a white-yellow film; ***m/z* (ES)**; (found MH$^+$ 324.171; $C_{19}H_{21}N_3O_2$ requires *MH* 324.17)

(2R,3R)-3-amino-1'-(2-hydroxybenzoyl)-2,3-dihydropsiro[indene-1,4'-piperidin]-2-ol

(JS-160-6)
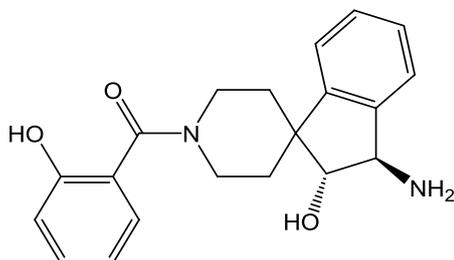
Title compound was from Chembridge 'screening library' as a white-yellow film; ***m/z* (ES)**; (found MH$^+$ 339.166; C$_{20}$H$_{22}$N$_2$O$_3$ requires *MH* 339.17)

1-{2-[(2R,3R)-3-amino-2-hydroxy-2,3-dihydro-1'H-spiro[indene-1,4'-piperidin]-1'-yl]-2-oxoethyl}-2-pyrrolidinone

(JS-160-7)
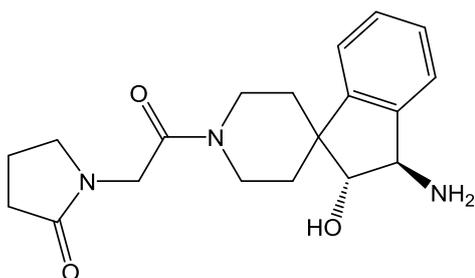
Title compound was from Chembridge 'screening library' as a yellow film; ***m/z* (ES)**; (found MH$^+$ 344.196; C$_{19}$H$_{25}$N$_3$O$_3$ requires *MH* 344.19)

223

1-methyl-1'-(tetrahydro-2H-pyran-4-yl)spiro[indole-3,4'-piperidin]-2(1H)-one

(JS-160-8)
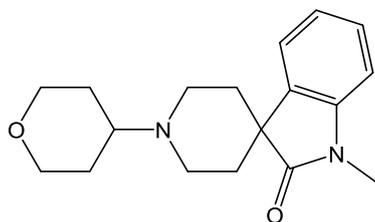
Title compound was from Chembridge 'screening library' as a white film; *m/z* **(ES)**; (found MH$^+$ 301.192; $C_{18}H_{24}N_2O_2$ requires *MH* 301.19)

1'-(2-hydroxyethyl)-1-methylspiro[indole-3,4'-piperidin]-2(1H)-one

(JS-160-9)
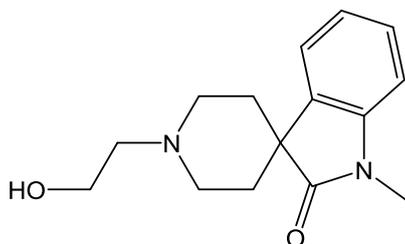
Title compound was from Chembridge 'screening library' as a yellow film; *m/z* **(ES)**; (found MH$^+$ 261.159; $C_{15}H_{20}N_2O_2$ requires *MH* 261.16)

Compound 1

(2S*,3R*)-N^2-(3-Chlorophenyl)oxirane-2,3-dicarboxamide
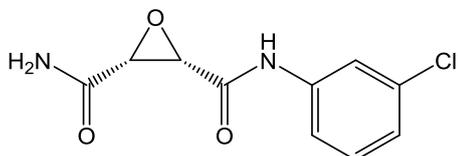
Title compound was synthesised in Leeds by Craig Avery (CA424, 6.26, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 263.019; $C_{10}H_9ClN_2O_3$ requires *MNa* 263.02).

Compound 2

(2S*,3R*)-N^2-Phenyloxirane-2,3-dicarboxamide
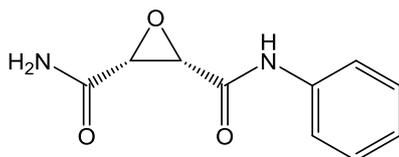
Title compound was synthesised in Leeds by Craig Avery (CA441, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 229.059; $C_{10}H_{10}N_2O_3$ requires *MNa* 229.06).

Compound 3

**(2S*,3R*)-N^2-(4-Bromophenyl)oxirane-2,3-dicarboxamide**
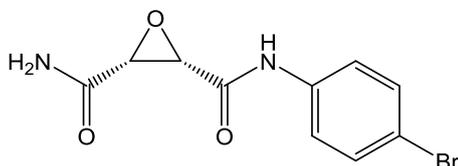
Title compound was synthesised in Leeds by Craig Avery (CA445, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 306.969; $C_{10}H_9BrN_2O_3$ requires *MNa* 306.97).

Compound 4

**(2S*,3R*)-N^2-(4-Bromo-3-chlorophenyl)oxirane-2,3-dicarboxamide**
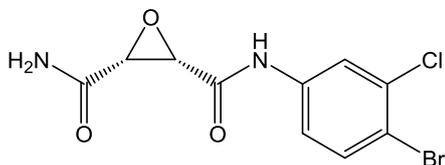
<u>Title compound</u> was synthesised in Leeds by Craig Avery (CA446, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 340.930; C$_{10}$H$_8$BrClN$_2$O$_3$ requires *MNa* 340.93).

Compound 5

(2S*,3R*)-N^2-(3,4-Dimethylphenyl)oxirane-2,3-dicarboxamide
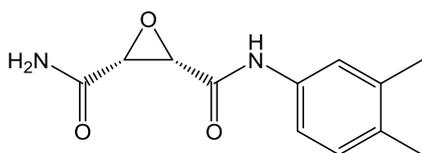

<u>Title compound</u> was synthesised in Leeds by Craig Avery (CA447, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 257.090; C$_{12}$H$_{14}$N$_2$O$_3$ requires *MNa* 257.09).

Compound 6

(2S*,3R*)-N^2-(3-Phenoxyphenyl)oxirane-2,3-dicarboxamide
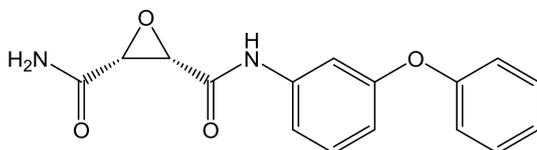
<u>Title compound</u> was synthesised in Leeds by Craig Avery (CA432, Craig Avery PhD Thesis, University of Leeds); *m/z* **(ES)**; (found MNa 321.085; C$_{16}$H$_{14}$N$_2$O$_4$ requires *MNa* 321.09).

Compound 7

(2S*,3R*)-N^2-(4-(4-Aminophenoxy)phenyl)oxirane-2,3-dicarboxamide
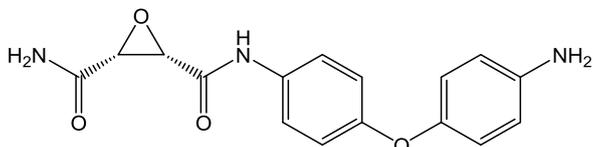
Title compound was synthesised in Leeds by Craig Avery (CA451, Craig Avery PhD Thesis, University of Leeds); ***m/z* (ES)**; (found MNa 336.096; $C_{16}H_{15}N_3O_4$ requires *MNa* 336.10).

Compound 8

(2S*,3R*)-N^2-(4-(4-Nitrophenoxy)phenyl)oxirane-2,3-dicarboxamide
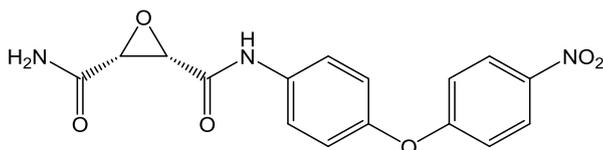
Title compound was synthesised in Leeds by Craig Avery (CA450, Craig Avery PhD Thesis, University of Leeds); ***m/z* (ES)**; (found MNa 366.070; $C_{16}H_{13}N_3O_6$ requires *MNa* 366.07).

Compound 9

**(2S*,3R*)-N^2-(3-Chloro-4-(2-ethoxyphenyl)phenyl)oxirane-2,3-dicarboxamide**
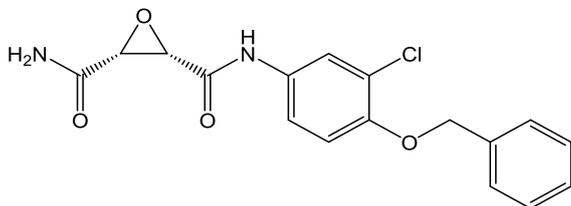
Title compound was synthesised in Leeds by Craig Avery (CA462, Craig Avery PhD Thesis, University of Leeds); ***m/z* (ES)**; (found MNa 369.062; $C_{17}H_{15}ClN_2O_4$ requires *MNa* 369.06).

Compound 10

(2S*,3R*)-N^2-(3-Chloro-4-(4-(*tert*-Butoxycarbonyl)piperidin-1-methoxy)phenyl)oxirane-2,3-dicarboxamide
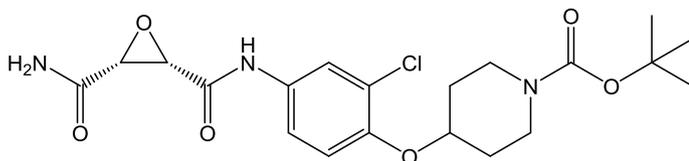
Title compound was synthesised in Leeds by Craig Avery (CA465, Craig Avery PhD Thesis, University of Leeds); ***m/z* (ES)**; (found MNa 462.140; C$_{20}$H$_{26}$ClN$_3$O$_6$ requires *MNa* 462.14).

Compound 11

(2S*,3R*)-N^2-(4-(2-(2-Nitrophenylsulphonamido)ethyl)phenyl)oxirane-2,3-dicarboxamide
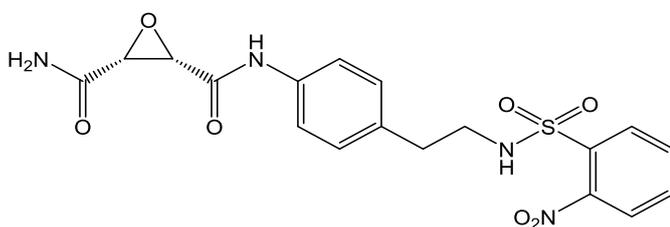
Title compound was synthesised in Leeds by Craig Avery (CA434, Craig Avery PhD Thesis, University of Leeds); ***m/z* (ES)**; (found MNa 457.079; C$_{18}$H$_{18}$ClN$_4$O$_7$S requires *MNa* 457.08).