

## Challenges in Modern Drug Discovery: A Case Study of Boceprevir, an HCV Protease Inhibitor for the Treatment of Hepatitis C Virus Infection

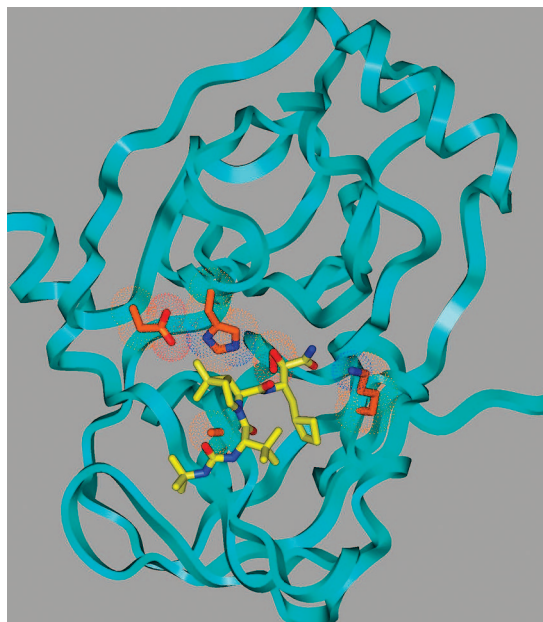
F. GEORGE NJOROGE,\* KEVIN X. CHEN, NENG-YANG SHIH,  
AND JOHN J. PIWINSKI

Schering-Plough Research Institute, 2015 Galloping Hill Road,  
Kenilworth, New Jersey 07033

RECEIVED ON MAY 4, 2007

### CON SPECTUS

**M**ore than 170 million people worldwide are affected by the hepatitis C virus (HCV). The disease has been described as a “silent epidemic” and “a serious global health crisis”. HCV infection is a leading cause of chronic liver disease such as cirrhosis, carcinoma, or liver failure. The current pegylated interferon and ribavirin combination therapy is effective in only 50% of patients. Its moderate efficacy and apparent side effects underscore the need for safer and more effective treatments. The non-structural NS3 protease of the virus plays a vital role in the replication of the HCV virus. The development of small molecule inhibitors of NS3 protease as antiviral agents has been intensively pursued as a viable strategy to eradicate HCV infection. However, it is a daunting task. The protease has a shallow and solvent-exposed substrate binding region, and the inhibitor binding energy is mainly derived from weak lipophilic and electrostatic interactions. Moreover, lack of a robust *in vitro* cell culture system and the absence of a convenient small animal model have hampered the assessment of both *in vitro* and *in vivo* efficacy of any antiviral compounds. Despite the tremendous challenges, with access to a recently developed cell-based replicon system, major progress has been made toward a more effective small molecule HCV drug. In our HCV program, facing no leads from our screening effort, a structure-based design approach was carried out. An  $\alpha$ -ketoamide-type electrophile was designed to trap the serine hydroxyl of the protease. Early ketoamide inhibitors mimicked the structures of the peptide substrates. With the aid of X-ray structures, we successfully truncated the undecapeptide lead that had a molecular weight of 1265 Da stepwise to a tripeptide with a molecular weight of 500 Da. In an attempt to depeptidize the inhibitors, various strategies such as hydrazine urea replacement of amide bonds and P2 to P4 and P1 to P3 macrocyclizations were examined. Further optimization of the tripeptide inhibitors led to the identification of the best moieties for each site: primary ketoamide at P', cyclobutylalanine at P1, *gem*-dimethylcyclopropylproline at P2, *tert*-leucine at P3, and *tert*-butyl urea as capping agent. The combination of these led to the discovery of compound **8** (SCH 503034, boceprevir), our clinical candidate. It is a potent inhibitor in both enzyme assay ( $K_i^* = 14$  nM) and cell-based replicon assay ( $EC_{90} = 0.35$   $\mu$ M). It is highly selective (2200 $\times$ ) against human neutrophil elastase (HNE). Boceprevir is well tolerated in humans and demonstrated antiviral activity in phase I clinical trials. It is currently in phase II trials. This Account details the complexity and challenges encountered in the drug discovery process.

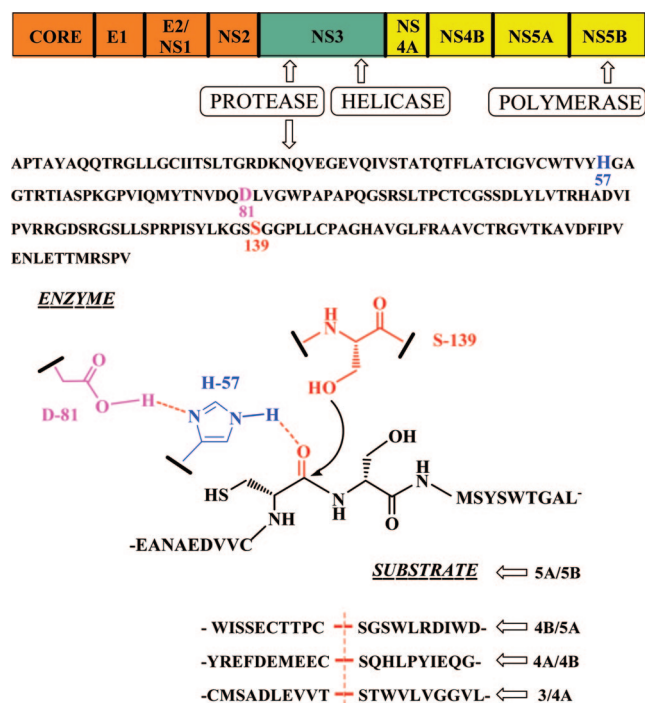


## Introduction

Hepatitis C virus (HCV) infection is one of the most significant health problems affecting humans. An estimated 170 million individuals (3%) worldwide and more than 4 million Americans (1.3%) are infected with HCV.<sup>1</sup> In roughly 80% of cases, the virus leads to a chronic form of hepatitis, a condition that is incurable in many patients. Without therapeutic intervention, it can lead to morbidity or mortality in 10–20 years through either cirrhosis and hepatic failure or hepatocellular carcinoma.<sup>2</sup> It is anticipated that a significant percentage of those currently infected will develop cirrhosis and other associated hepatic sequelae. HCV infection is the most common cause of liver transplantation.<sup>2</sup>

Current medical treatment options are limited. The only drugs available are subcutaneous interferon- $\alpha$  (IFN- $\alpha$ ) or pegylated-IFN- $\alpha$ , alone or in combination with oral ribavirin.<sup>3</sup> IFN- $\alpha$  is a protein that stimulates the immune system, while ribavirin is a nucleoside analog that works in concert with IFN- $\alpha$  to control the infection. The efficacy of this combination therapy against the predominant genotype 1 affecting North America, Europe, and Japan is moderate, with only about 50% of the patients showing sustained virological response. Some patients also experience significant side effects related to the treatment. With few alternatives available, more effective agents with fewer side effects are clearly needed.<sup>3</sup>

HCV, the etiologic agent of non-A, non-B hepatitis, was identified in 1989<sup>4</sup> and is a member of the Flaviviridae. HCV is an enveloped, positive-strand RNA virus of approximately 9.6 kilobases. Upon entering a suitable host cell, the HCV genome serves as a template for cap-independent translation through an internal ribosome entry site (IRES).<sup>5</sup> The resulting polyprotein undergoes both co- and post-translational proteolytic maturation by host and virally encoded proteases. The virally encoded protease responsible for processing the non-structural (NS) portion of the polyprotein is located in the N-terminal third of the NS3 protein<sup>5</sup> (Figure 1). Following autoproteolysis of the NS3–NS4A junction, the protease cleaves the polyprotein at the NS4A–NS4B, NS4B–NS5A, and NS5A–NS5B junctions to release the downstream NS proteins. The mature proteins subsequently self-assemble on the endoplasmic reticulum to generate the replicative complex or replisome. The replisome, using the viral genome as a template, generates negative-strand viral RNA intermediates, which are then used as templates to synthesize new positive-strand (genomic) RNAs. These are either translated to yield more polyprotein or, later in the infection cycle, encapsulated to



**FIGURE 1.** Schematic representation of the HCV genome and NS3 protease polyprotein. NS3 is a bifunctional protein with protease and helicase activities. The green box represents the NS3 protease domain, and thin arrows depict cleavages mediated by the NS3/NS4A protease complex.

generate progeny virions. Inhibition of the maturational activities of the NS3 protease would therefore suppress replisome formation, RNA replication, and ultimately new virions.

The X-ray structures of NS3 protease, both as an isolated domain and in the full-length NS3 protein, have been determined.<sup>6</sup> The structural data has provided detailed insights to facilitate rational inhibitor design. The NS3 protease is in many ways a typical  $\beta$ -barrel serine protease, with a canonical Asp-His-Ser catalytic triad similar to the well-studied digestive enzymes trypsin and chymotrypsin. By contrast, the NS3 protease uses an extended polydentate binding cleft, with several recognition subsites to ensure specificity. It forms a heterodimeric complex with the NS4A protein, an essential cofactor that activates the protease and assists in anchoring the heterodimer to the endoplasmic reticulum. On the other hand, the RNA-dependent RNA polymerase (RdRp) contained within the NS5B protein is the catalytic component of the HCV RNA replication machinery.<sup>5,7</sup> This enzyme synthesizes RNA using the RNA template. This biochemical activity is not present in mammalian cells, offering the opportunity to identify very selective inhibitors of the viral enzyme.

Much of the recent effort has been directed toward developing drugs that inhibit viral replication. Several promising small-molecule inhibitors of the NS3/4A protease and the

NS5B polymerase are in clinical development.<sup>8</sup> Early testing has demonstrated strong antiviral activity both *in vitro* and in patients. Inhibitors toward other potential targets such as IRES and NS5A are in earlier stages of preclinical investigation. Although efforts are ongoing to develop a vaccine,<sup>9</sup> the unusually rapid genetic drift of HCV makes this a daunting task.

## Challenges in Discovering HCV Protease Inhibitors

Although HCV was characterized more than a decade ago, the lack of a robust *in vitro* cell culture system capable of supporting its replication has complicated traditional approaches to developing or evaluating antiviral compounds. Likewise, the absence of a convenient small animal model has hampered the assessment of *in vivo* efficacy. Most of our knowledge of HCV has been derived from surrogate experimental systems that approximate infection and often preclude definitive interpretation. Only since the development of the HCV autonomous subgenomic replicon system,<sup>10</sup> a severe combined immunodeficiency disease (SCID) mouse with chimeric human liver model,<sup>11</sup> and the chronically infected chimpanzee model<sup>10</sup> has the preclinical evaluation of potential anti-HCV agents become possible.

There are currently two animal models for preclinical evaluation of anti-HCV therapies; both of which suffer from limitations that make them less than ideal for preliminary studies. HCV infects only humans and chimpanzees. The chronically infected chimpanzee model,<sup>12</sup> the "gold standard" for HCV studies, is challenging and expensive because one out of three chimpanzees spontaneously resolve their HCV infection. An immune-deficient SCID mouse–human liver xenograft system<sup>11</sup> was developed by researchers at the University of Alberta. In this model, the livers of neonate SCID beige mice are colonized with infused human hepatocytes, which rescues them from a fatal transgene. Infection of these human liver grafts by several genotypes of HCV and the therapeutic effects of INF- $\alpha$  have been reported. Unfortunately, the animals are fragile and scale up of the colony has been slower than expected, thus limiting access to the system.

Owing to the fact that the NS3–NS4A protease is playing a critical role in HCV viral replication, it has been viewed as an ideal target for the creation of new HCV therapy.<sup>8,13</sup> However, developing HCV protease inhibitors as drugs was no trivial task. At the onset of our work, there were no viable lead structures from which to develop potential drug candidates. Our screening effort of four million compounds did not generate any meaningful leads to initiate a drug discovery effort.

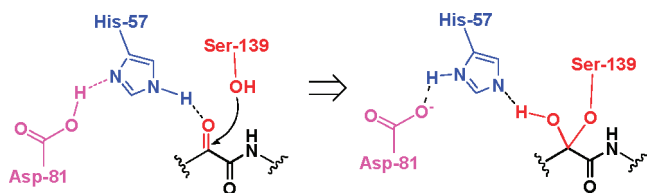
Thus, early inhibitors were designed based on the substrate–enzyme active site interactions.<sup>14</sup> However, the HCV protease requires an extensive peptide substrate, with which it establishes multiple weak interactions distributed along an extended surface. The requirement of large substrates was a major concern in the development of orally bioavailable small molecule drugs. It was feared that the enzyme may only be inhibited by molecules large enough to mimic the natural substrate. Indeed, early leads were long peptidic compounds that occupied much of the substrate-binding site to take advantage of multiple hydrogen bonding and hydrophobic interactions. The major challenge was to modify these large molecules to the less peptidic and lower molecular weight drug candidates with desirable pharmacokinetic (PK) profiles, while retaining or improving potency in the enzymatic and cellular assays. The resolution of the three-dimensional structure<sup>6</sup> of the enzyme damped the enthusiasm of medicinal chemists further, because the substrate-binding cleft of NS3–NS4A protease seemed flat and featureless, lacking the cavities, holes, and flaps, or so-called binding pockets, that had been exploited as anchor points to design potent and selective inhibitors of other proteases.

Moreover, there was another major challenge for any successful anti-HCV therapy: drug-resistant viruses emerge rapidly under the selective pressure. The fast turnover rate and the intrinsic low fidelity of the HCV replication machinery endowed the virus with the ability to fully explore its genome space and quickly come up with mutations that rendered it resistant to antiviral drugs.

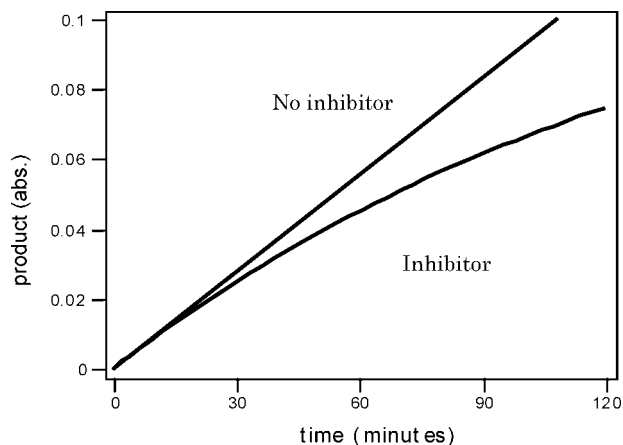
As a result of all these difficulties, no HCV protease inhibitors have been approved yet for the treatment of HCV disease in spite of the fact that the virus was fully characterized in 1989. However, efforts within the pharmaceutical industry have resulted in several candidates in clinical development,<sup>8</sup> including boceprevir.<sup>15</sup>

## Our Approach for Inhibitor Design

Since our screening efforts failed to generate potential leads to initiate a drug discovery effort, we embarked on a structure-based design approach. Early research for inhibitors capitalized on the observation that the enzyme is susceptible to marked inhibition by the N-terminal peptide products released from the substrates upon enzymatic cleavage. Learning from experiences from others in developing potent serine-protease inhibitors,<sup>16</sup> incorporation of a serine trap into the molecule was a promising approach. We envisioned that reaction of the active site serine (Ser139)<sup>6</sup> with conventional electrophiles such as aldehydes, ketones, trifluoromethyl ketones, and  $\alpha$ -ke-



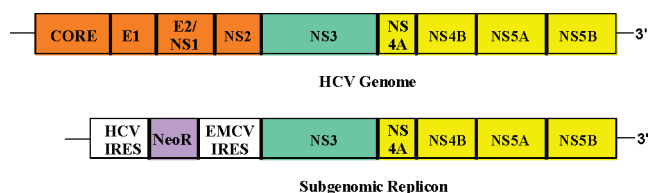
**FIGURE 2.** Nucleophilic attack of the  $\alpha$ -ketoamide by Ser139 led to a covalent tetrahedron intermediate stabilized by residues His57 and Asp81.



**FIGURE 3.** Progress curve of peptide substrate hydrolysis by the HCV protease domain showing the time-dependent inhibition.

toamides and subsequent trapping of the resulting transition-state analogues by the active site triad (Ser139, His57, and Asp81, Figure 2)<sup>6</sup> would provide effective inhibition. After considerable research, a few large peptidic molecules, mimicking the peptide substrate structures but containing an  $\alpha$ -ketoamide functionality, were discovered to be potent HCV protease inhibitors.

The  $\alpha$ -ketoamide moiety was essential for the inhibitory activity. When it is attacked by Ser139 of the NS3 protease, it forms a stable, covalent, and reversible complex with the enzyme. The time required for stable covalent adduct formation is on the order of minutes. To assess the potency of such slowly equilibrating or "slow-binding" inhibitors accurately, proteolytic reactions containing inhibitor are usually monitored until equilibrium is evident using progress curve analysis (Figure 3). In this continuous assay,<sup>17</sup> the extent of hydrolysis of chromogenic 4-phenylazophenyl (PAP) ester from the peptide fragment Ac-DTEDVVP(Nva)-O-4-PAP was spectrophotometrically determined. To underscore the slow-binding nature of these molecules and distinguish them from simple, instantaneous competitive inhibitors, the equilibrium binding constant was usually designated  $K_i^*$ ,<sup>18</sup> although for most purposes it could be considered equivalent to a traditional  $K_i$ .



**FIGURE 4.** Schematic representation of the RNA encoding the bicistronic subgenomic HCV replicon.

Besides an enzymatic assay, a cell-based assay was also essential for optimizing inhibitor potency. An HCV subgenomic replicon system was developed by Bartenschlager and colleagues in 1999.<sup>10</sup> The replicon cell-based assay has been used extensively to evaluate the functional potency and subsequent antiviral efficacy of HCV protease inhibitors. The HCV replicon is essentially a defective (i.e., noninfectious) viral genome in which the sequences encoding the structural proteins at the 5' end of the RNA have been replaced by a selectable marker, the neomycin resistance gene (NeoR) (Figure 4). The NeoR marker allows selection of cells harboring functional replicons following transfection and antibiotic treatment. Replicon constructs, including those developed to evaluate potential antiviral agents, use a bicistronic design where two independent IRES elements are present. The HCV IRES sequence drives expression of the neomycin resistance gene to allow selection of replicon bearing cells and a second IRES sequence from encephalomyocarditis virus (EMCV) initiates translation of the RNA segment encoding HCV nonstructural proteins from NS3 to NS5B.

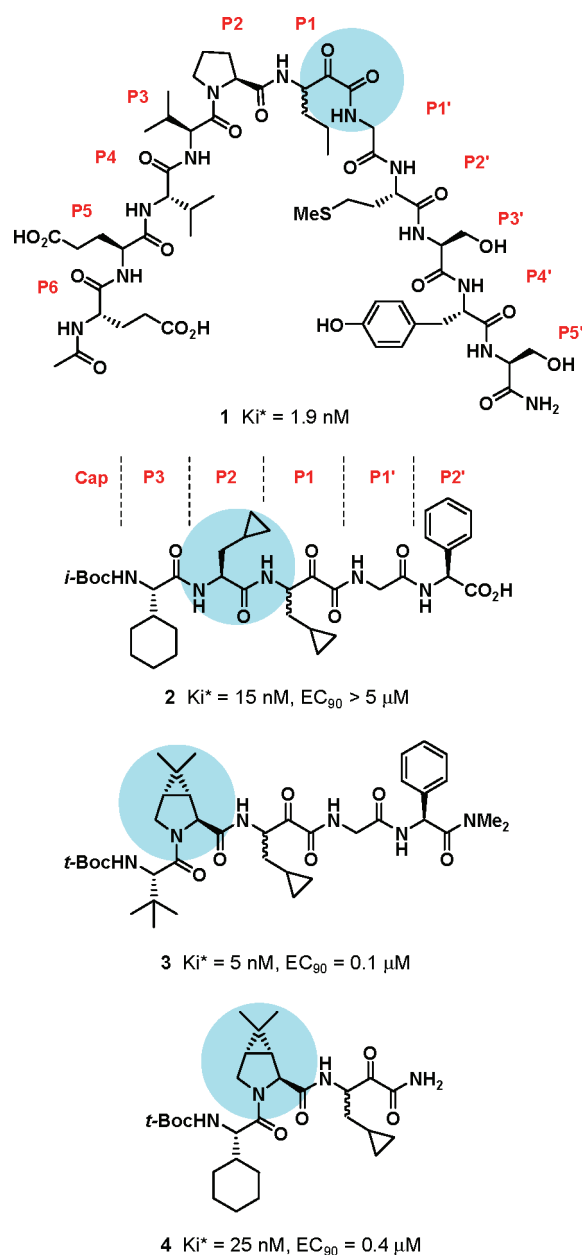
Unlike a true HCV infection, cells bearing HCV replicons, even full-length replicons expressing structural proteins, do not generate progeny virions. At the time, the replicon system was the only germane *in vitro* system for evaluating potential antiviral agents directed against the HCV nonstructural proteins and, consequently, provided an essential and stringent system for the evaluation of potent inhibitors of HCV protease. The  $EC_{50}$  and  $EC_{90}$  values for suppression of the bicistronic subgenomic replicon (genotype 1b) were obtained through a 72 h assay in HuH-7 cells. At 72 h, cells were lysed, and the replicon RNA level was determined using real-time polymerase chain reaction (PCR) analysis (Taqman) that targeted the NS5B portion of the viral genome. Changes in replicon RNA level were compared to an internal control, cellular glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA levels, in a single-tube multiplex reaction. Dose-response curves were generated and drug concentrations resulting in a 2-fold or a 10-fold reduction in replicon RNA were estimated using a grid search method to give  $EC_{50}$  and  $EC_{90}$  values.

One of the key elements contributing to the potency of an  $\alpha$ -ketoamide is the electrophilic ketoamide functionality, which forms a reversible covalent bond with serine 139 of the HCV protease. However, the ketoamide motif could also be susceptible to attack by a variety of other nucleophiles present on other proteins. Thus, to address the selectivity issue, inhibitory activity against the human neutrophil elastase (HNE) was measured as a gauge of selectivity (HNE/HCV) in the structure–activity–relationship (SAR) development. The active site of HNE resembles closely the active site structure of HCV NS3 protease. Although the clinical relevance of this selectivity parameter has not been demonstrated, it helped to serve as a guide in designing selective inhibitors versus similar proteases, thereby potentially minimizing side effects that could surface in the clinic.

Our early work led to the discovery of undeca-peptide  $\alpha$ -ketoamide inhibitors,<sup>15</sup> which had structures with 11 amino acid residues that spanned from P6 to P5'.<sup>19</sup> One of the earliest leads, compound **1** (Figure 5), exhibited excellent HCV NS3 protease inhibition ( $K_i^* = 1.9$  nM). It was a mixture of two diastereomers at the epimerizable P1  $\alpha$ -center. However, with a molecular weight of 1265 Da, it was not surprising that this compound did not exhibit a desirable PK profile. Nevertheless, compound **1** and related analogs served as a starting point for further SAR studies. Our research toward improving the overall profile of these initial leads was greatly aided by using X-ray structures of inhibitors bound to the HCV NS3 protease and continuous drug metabolism/pharmacokinetic evaluation.

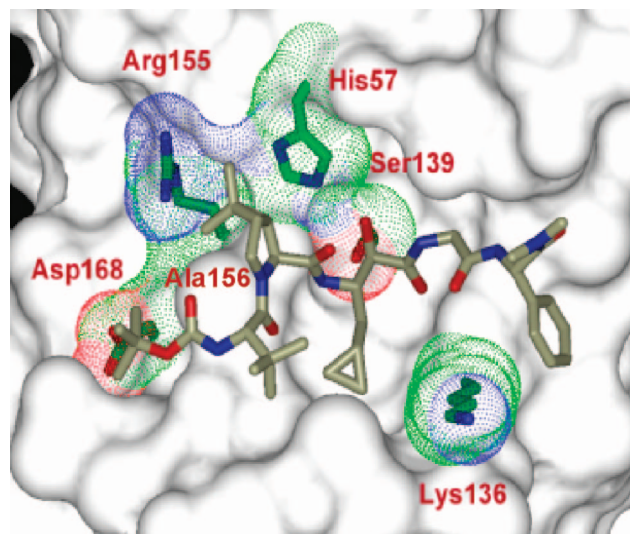
**Truncation Effort.** Starting from the lead compound **1**, an undeca-peptide, a series of stepwise truncations at both the prime and nonprime ends led to a pentapeptide inhibitor **2**.<sup>20</sup> The molecular weight was reduced to approximately half of that of compound **1**. A number of amino acids were also modified to optimize potency. Thus, P1 and P2 were replaced with cyclopropylalanines, P3 was replaced with cyclohexylglycine, and P2' was substituted with phenylglycine. The compound was capped as isobutyl carbamate at P3 and as carboxylic acid at P2'. Inhibitor **2** was quite potent with a  $K_i^*$  of 15 nM. However, it was inactive in the cell-based HCV replicon assay ( $EC_{90} > 5$   $\mu$ M), possibly due to the presence of the charged carboxylic acid preventing it from getting into the cells.

Further optimization along the entire backbone was carried out with the aid of X-ray crystal structures. The most important discovery occurred at P2 position. The *gem*-dimethylcyclopropylproline was discovered to be a superior leucine surrogate. The *tert*-butylglycine was found to be a good replacement of cyclohexylglycine at P3. The capping groups



**FIGURE 5.** From undeca-peptide to pentapeptide HCV NS3 protease inhibitors.

at both ends were also changed to *tert*-butyl carbamate for P3 and dimethylamide for P2', respectively. The latter would address the cell penetration issue of compound **2**. Indeed, as a result of all these optimizations, compound **3**<sup>21</sup> demonstrated a dramatic improvement over compound **2**, with good binding affinity ( $K_i^* = 5$  nM) and replicon activity ( $EC_{90} = 0.1$   $\mu$ M). Shown in the X-ray structure of **3** bound to the protease (Figure 6), the improvement in binding potency was a result of favorable interaction of the P2 *gem*-dimethylcyclopropylproline moiety with the methyl side chain of Ala156. The cyclopropylmethyl side chain of P1 and the phenyl group of P2' formed a "C"-shaped clamp around the side chain of

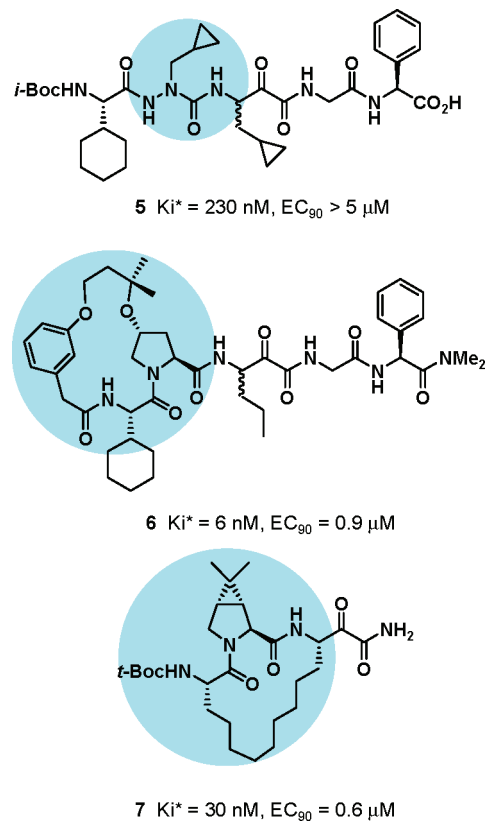


**FIGURE 6.** X-ray structure of inhibitor **3** bound to the HCV protease.

Lys136. However, the molecular weight (725 Da) of **3** was still larger than that of a typical drug candidate, which, preferably, should be less than 500 Da.<sup>22</sup> Thus, it was not unexpected that this translated to a poor PK profile in rats and monkeys. Generally, an oral area-under-curve (po AUC) of greater than 1.0  $\mu\text{M}\cdot\text{h}$  and bioavailability of at least 10% are desired. In rats, compound **3** had a po AUC of 0.35  $\mu\text{M}\cdot\text{h}$  at 10 mg/kg and a bioavailability of 4%, while in monkeys the AUC was 0.03  $\mu\text{M}\cdot\text{h}$  at 3 mg/kg and bioavailability was 1%. Clearly, further reduction in the molecular weight of the inhibitors was necessary to achieve desirable PK properties.

Due to the importance of the remaining residues at non-prime side, we focused our additional truncation efforts on the prime side. First, the P2' phenylglycine residue was removed with modest loss of potency. The P1' glycine was replaced by simple alkyl groups such as methyl, ethyl, or allyl. Unfortunately, all secondary and tertiary P1 ketoamides were significantly less potent than compound **3**. However, it was discovered that the nonsubstituted primary ketoamides had the best potency in both enzyme and replicon assays. The representative inhibitor after this round of SAR development was compound **4** (Figure 5). It had a  $K_i^*$  of 25 nM, and a replicon  $\text{EC}_{90}$  of 0.40  $\mu\text{M}$ . Although both values were higher than those of analog **3**, the molecular weight was significantly lower and fell within the preferred range for a developable candidate.

**Depeptization Studies.** It is difficult to acquire desirable pharmacokinetic properties from large peptides because they are susceptible to hydrolysis by various peptidases. Thus, substantial efforts were devoted to depeptizing the lead molecules. Shown in Figure 7 are several examples of that effort.

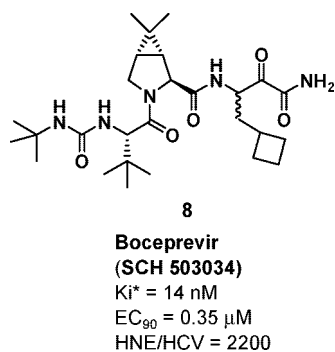


**FIGURE 7.** Efforts toward depeptization of HCV protease inhibitors.

First, replacement of P2 amino acid residue with an aza motif was executed. Thus, in compound **5**,<sup>18</sup> the cyclopropylalanyl P2 was replaced with a substituted hydrazine urea moiety. Unfortunately, that resulted in a significant loss in binding affinity ( $K_i^* = 230 \text{ nM}$ ) and a total loss of potency in replicon assay ( $\text{EC}_{90} > 5 \mu\text{M}$ ).

Efforts were also made to improve potency and PK profiles through macrocyclization of P2 and P3 residues, resulting in macrocyclic inhibitors of type **6**<sup>23</sup> (Figure 7). A phenylacetamide capping from P3 was linked to C-4 of P2 proline through a *tert*-alkyl ether linkage. The 16-membered macrocyclic ring formed a donut-shaped circle over the methyl group of Ala156 as evidenced by X-ray crystallography. The inhibitor had excellent  $K_i^*$  of 6 nM and was moderately active in the cell-based assay ( $\text{EC}_{90} = 0.90 \mu\text{M}$ ). However, it had an oral AUC of only 0.46  $\mu\text{M}\cdot\text{h}$  at 10 mg/kg in rats with a low bioavailability of 2.2%. Similarly sized compounds with modifications along the macrocyclic ring gave equally poor PK results. When truncated at the prime side, smaller macrocyclic inhibitors did not have the desired level of potency in replicon assay.

On the other hand, macrocyclization from P1 to P3 residues through a 16-membered ring provided novel inhibitors such as **7**.<sup>24</sup> With these analogs, we were able to incorpo-



**FIGURE 8.** HCV protease inhibitor **8**.

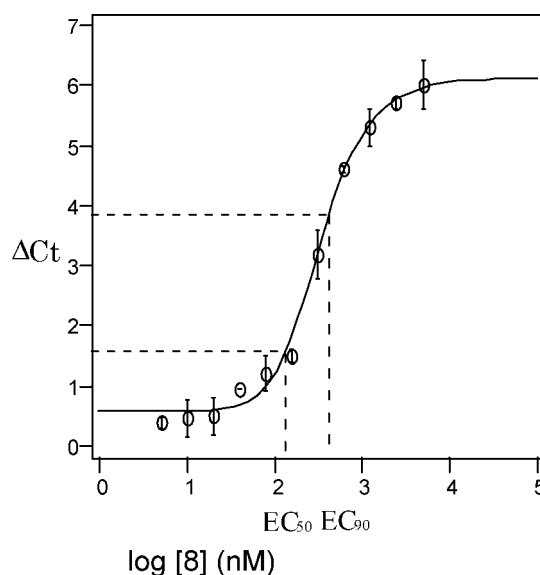
rate the dimethylcyclopropyl-proline at P2. Compound **7** was a single diastereomer with a nonepimerizable P1  $\alpha$ -center. It had a respectable  $K_i^*$  of 30 nM, and a good replicon  $EC_{90}$  of 0.60  $\mu\text{M}$ . However, as in the other cases described above, further exploration of this series of inhibitors did not provide an improved PK profile.

In summary, the depeptization exercises carried out neither resulted in significant improvement in potency nor provided measurable enhancement in PK properties in these HCV protease inhibitors.

### Discovery of Boceprevir (**8**)

Extensive SAR investigation at the P1, P3, and P3-capping positions from the lead primary ketoamide **4** was continued. The dimethylcyclopropylproline moiety at P2 was established to be optimal and was retained in all subsequent analogs. Systematic variation in chain length and ring size on the P1 side chain led to the discovery of cyclobutylalanine as the optimal choice. Rings with other sizes, such as cyclopentyl or cyclohexyl, were found to be too large for the S1 pocket. A large effort was also spent on exploring many P3 capping groups. Various substituents extending from the P3 amino group were examined; these included alkyls, aryls, amides, carbamates, ureas, sulfonamides, and sulfonyl ureas. It was discovered that the urea type of P3 capping gave the best overall profile. Ultimately, the combination of various optimized moieties led to the discovery of boceprevir (**8**, SCH 503034, Figure 8),<sup>15</sup> which was selected for drug safety studies and potential development in clinical trials.

The  $\alpha$ -center of the P1 residue of **8** was racemic. The two diastereomeric compounds could be separated by HPLC. However, when either pure isomer was treated with an organic or inorganic base (e.g., triethyl amine or lithium hydroxide), they underwent rapid isomerization. Fast equilibration was also demonstrated under the conditions of biological assays. This alleviated any need for separation of the two entities for pharmacological evaluations. The ratio of the two isomers varied



**FIGURE 9.** Effect of **8** on HuH-7 cells bearing the subgenomic HCV replicon. RNA level was measured relative to an internal control ( $\Delta\text{Ct}$ ). Increase in  $\Delta\text{Ct}$  indicated decreasing replicon RNA levels; each  $\Delta\text{Ct}$  reflects a 2-fold change in RNA level from baseline.

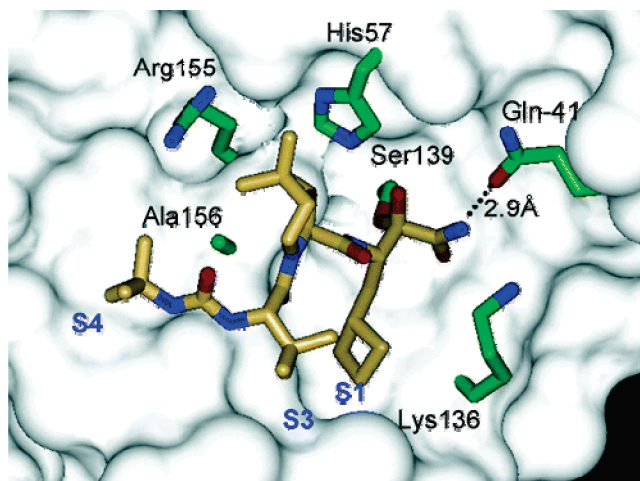
significantly depending on the experimental conditions. The isomer with an (*S*)-configuration at the P1  $\alpha$ -center was the major isomer in most cases.

Compound **8** had an optimal overall profile. In the HCV NS3 protease continuous assay, it had a potency of 14 nM ( $K_i^*$ ) averaged over a large number of runs. In the 72-h bicistronic subgenomic cell-based replicon assay in HuH-7 cells, the  $EC_{50}$  and  $EC_{90}$  values were determined to be 0.20  $\mu\text{M}$  and 0.35  $\mu\text{M}$ , respectively (Figure 9). Inhibitor **8** was also found to be a very weak inhibitor of HNE ( $K_i = 26 \mu\text{M}$ ) representing a selectivity of 2200. Additionally, the reactivity of **8** toward a panel of other serine proteases was measured, and **8** showed no cross-reactivity when tested up to 50  $\mu\text{M}$  with trypsin, chymotrypsin, thrombin, and factor Xa. The cross-reactivity against a broad panel of other general enzymes was also evaluated, and no major issues were identified. All these studies indicated that compound **8** was highly selective toward the HCV serine protease.

The pharmacokinetic profile of **8** was evaluated in several animal species. Following oral administration, it was moderately absorbed in rats, dogs, and monkeys. Absorption was relatively rapid in dogs but slower in mice, rats, and monkeys, as evidenced by mean absorption times (MAT) ranging from 0.5 to 1.4 h (Table 1). The AUC was good in dogs and rats, moderate in mouse, and low in monkeys. The absolute oral bioavailability was modest in mouse, rats, and dogs (26–34%) but low in monkeys (4%). There was no issue with CYP 2D6, 2C9, 2C19, and 3A4 inhibition, either co- or preincubated. Target organ analysis in rats revealed that **8** was

**TABLE 1.** Mean ( $n = 3$ ) Pharmacokinetic Parameters of **8** Following Oral Dosing

species	mouse	rat	dog	monkey
dose (mg/kg)	10	10	3	3
AUC (po, $\mu\text{M} \cdot \text{h}$ )	0.93	1.5	3.1	0.12
bioavailability (%)	34	26	30	4
MAT (h)	1.2	1.4	0.5	1.4
$C_{\text{max}}$ ( $\mu\text{M}$ )	2.3	0.66	2.3	0.09

**FIGURE 10.** X-ray structure of inhibitor **8** bound to the HCV protease.

highly concentrated in liver with a remarkable liver/plasma concentration ratio of approximately 30.

The X-ray structure of compound **8** bound to NS3 protease was solved and is shown in Figure 10. It was clear that the diastereomer with an (*S*)-configuration at the P1  $\alpha$ -center was the active inhibitor. The cyclobutylalanine moiety effectively occupied most of the space available at the S1 pocket. This group was largely responsible for the excellent selectivity observed with **8** versus human neutrophil elastase, which has a much smaller S1 pocket. The P2 dimethylcyclopropylproline residue adopted a bent conformation that allowed maximum overlap of the moiety with Ala156 of the enzyme. The *exo*-methyl on the cyclopropane ring had favorable interaction with imidazole of His57, and the *endo*-methyl had contact with Ala156 and Arg155. The side chain of P3 *tert*-butylglycine occupied the S3 pocket, providing good hydrophobic interaction with the enzyme. The *tert*-butyl group of the P3 urea capping group had good contact in the S<sub>4</sub> pocket, presumably also through a pure hydrophobic interaction. The ketoamide was reversibly trapped by Ser139 to form a covalent bond and at the same time donated a hydrogen bond to the protein backbone. In addition to van der Waals contacts, **8** formed a series of specific hydrogen bonds with the protein backbone, which involved P1-NH, P3-carbonyl, and

both urea NH's. Combination of a number of hydrophobic interactions and the array of hydrogen bonds contributed greatly to the binding potency and the selectivity of **8**.

In summary, we discovered boceprevir (**8**) as a novel, potent, highly selective, orally bioavailable HCV NS3 protease inhibitor. It has been advanced to clinical trials in human beings for the treatment of hepatitis C viral infections. It was well tolerated and demonstrated antiviral activity in phase I clinical trials and is currently in phase II.

## Summary and Outlook

The pursuit of a potent and orally bioavailable HCV NS3 protease inhibitor as a drug candidate for the treatment of hepatitis C has been a difficult task. The shallow and featureless nature of the enzyme's active site presented a significant challenge for the discovery of enzyme inhibitors. Without any viable leads, a structure-based drug design approach guided by X-ray crystal structures of the enzyme was pursued. Systematic truncations and depeptidizations on both prime and non-prime sites gave rise to smaller pentapeptides that were potent inhibitors, but did not possess the desirable pharmacokinetic properties. Modifications on the prime side resulted in the discovery of the primary  $\alpha$ -ketoamide moiety, which gave excellent potency. Further SAR optimization identified P1 cyclobutylalanine, P2 dimethylcyclopropylproline, P3 *tert*-butylglycine, and *tert*-butyl urea capping group as the best combination, which led to the discovery of boceprevir (**8**). It had an *in vitro* potency of 14 nM ( $K_i^*$ ) and cell-based replicon assay potency of 350 nM ( $\text{EC}_{90}$ ). Compound **8** demonstrated good oral bioavailabilities in rats and dogs and was found to be highly concentrated in the liver.

Across the pharmaceutical industry, several novel drug candidates have entered or will soon enter clinical evaluation to establish their clinical effectiveness for HCV patients. Aside from the safety and efficacy requirements common to all new drugs, the success of HCV-targeted agents will be heavily influenced by their ability to inhibit all viral variants and prevent the emergence of escape mutants. As is the case for HIV, combinations of several antiviral agents attacking different targets along the viral life cycle and, perhaps, the hosts themselves will certainly be required to control infection and prevent the emergence of drug-resistant viral variants.

*The authors wish to thank the many contributors whose efforts resulted in the successful outcome of this project. Especially noteworthy is the successful collaborative efforts of scientists in the Medicinal Chemistry (SPRI and Corvas), Structural Chemistry, Virology, and Drug Metabolism departments.*



## BIOGRAPHICAL INFORMATION

**George Njoroge** obtained his B.S. in Chemistry at University of Nairobi, Kenya, in 1979 and Ph.D. in Organic Chemistry at Case Western Reserve University (CWRU) in 1985. He was a postdoctoral fellow in the Institute of Pathology at CWRU working with Professor Vincent Monnier. He is currently director of Chemical Research at Schering-Plough Research Institute (SPRI).

**Kevin Chen** obtained his B.S. in Chemistry at University of Science and Technology of China in 1987 and Ph.D. in Organic Chemistry in 1996 at University of Wisconsin—Madison with Professor Edwin Vedejs. He was a NIH Postdoctoral Fellow until 1998 with Professor Stephen Martin at University of Texas at Austin. He is currently a Senior Principal Scientist at SPRI.

**Neng-Yang Shih** is Executive Director of Chemical Research at SPRI. He obtained his B.S. in Chemistry from Tamkang University in Taiwan and his Ph.D. in Organic Chemistry from University of Illinois at Chicago in 1981. He was a Postdoctoral Fellow at Columbia University with Professor Thomas J. Katz and at Harvard University with Professor Elias J. Corey.

**John Piwinski** is Group Vice President of Chemical Research at SPRI. He received his B.S. in Chemistry and Biochemistry from SUNY, Stony Brook, in 1976 and his Ph.D. in Organic Chemistry in 1980 from Yale University working with Professor Frederick Ziegler.

## REFERENCES

- (a) Cohen, J. The Scientific Challenge of Hepatitis C. *Science* **1999**, *285*, 26–30. (b) Alter, M. J.; Kruszon-Moran, D.; Nainan, O. V.; McQuillan, G. M.; Gao, F.; Moyer, L. A.; Kaslow, R. A.; Margolis, H. S. The Prevalence of Hepatitis C Virus Infection in the United States, 1988 through 1994. *N. Engl. J. Med.* **1999**, *341*, 556–562. (c) Cuthbert, J. A. Hepatitis C: Progress and Problems. *Clin. Microbiol. Rev.* **1994**, *7*, 505–532.
- (a) Hoofnagle, J. H. Course and Outcome of Hepatitis C. *Hepatology* **2002**, *36*, S21–S29. (b) Alter, H. J.; Seeff, L. B. Recovery, Persistence, and Sequelae in Hepatitis C Virus Infection: A Perspective on Long-Term Outcome. *Semin. Liver Dis.* **2000**, *20*, 17–35. (c) Brown, R. S., Jr.; Gaglio, P. J. Scope of Worldwide Hepatitis C Problem. *Liver Transplant.* **2003**, *9*, S10–S13.
- (a) Feld, J. J.; Hoofnagle, J. H. Mechanism of Action of Interferon and Ribavirin in Treatment of Hepatitis C. *Nature* **2005**, *436*, 967–972. (b) Neumann, A. U.; Lam, N. P.; Dahari, H.; Gretch, D. R.; Wiley, T. E.; Layden, T. J.; Perelson, A. S. Hepatitis C Virus Dynamics *In Vivo* and the Antiviral Efficacy of Interferon- $\alpha$  Therapy. *Science* **1998**, *282*, 103–107. (c) Fried, M. W.; Shiffman, M. L.; Reddy, D. K. R.; Smith, C.; Marinos, G.; Goncalves, F. L.; Häussinger, D.; Diago, M.; Carosi, G.; Dhumeaux, D.; Craxi, A.; Lin, A.; Hoffman, J.; Yu, J. PEGinterferon  $\alpha$ -2a plus Ribavirin for Chronic Hepatitis C Virus Infection. *New Engl. J. Med.* **2002**, *347*, 975–982.
- Choo, Q. L.; Kuo, G.; Weiner, A. J.; Overby, L. R.; Bradley, D. W.; Houghton, M. Isolation of a cDNA Clone Derived from a Blood-Borne Non-A, Non-B Viral Hepatitis Genome. *Science* **1989**, *244*, 359–362.
- (a) Lindenbach, B. D.; Rice, C. M. Unravelling Hepatitis C Virus Replication from Genome to Function. *Nature* **2005**, *436*, 933–938. (b) Kolykhalov, A. A.; Mihalik, K.; Feinstone, S. M.; Rice, C. M. Hepatitis C Virus-Encoded Enzymatic Activities and Conserved RNA Elements in the 3' Nontranslated Region Are Essential for Virus Replication *In Vivo*. *J. Virol.* **2000**, *74*, 2046–2051.
- (a) Yan, Y.; Li, Y.; Munshi, S.; Sardana, V.; Cole, J. L.; Sardana, M.; Steinkuehler, C.; Tomei, L.; De Francesco, R.; Kuo, L. C.; Chen, Z. Complex of NS3 Protease and NS4A Peptide of BK Strain Hepatitis C Virus: A 2.2 Å Resolution Structure in a Hexagonal Crystal Form. *Protein Sci.* **1998**, *7*, 837–847. (b) Love, R. A.; Parge, H. E.; Wichersham, J. A.; Hostomsky, Z.; Habuka, N.; Moomaw, E. W.; Adachi, T.; Hostomsky, Z. The Crystal Structure of Hepatitis C Virus NS3 Proteinase Reveals a Trypsin-Like Fold and a Structural Zinc Binding Site. *Cell* **1996**, *87*, 331–342.
- (a) Behrens, S. E.; Tomei, L.; De Francesco, R. Identification and Properties of the RNA-Dependent RNA Polymerase of Hepatitis C Virus. *EMBO J.* **1996**, *15*, 12–22. (b) Lesburg, C. A.; Cable, M. B.; Ferrari, E.; Hong, Z.; Mannarino, A. F.; Weber, P. C. Crystal Structure of the RNA-Dependent RNA Polymerase from Hepatitis C Virus Reveals a Fully Encircled Active Site. *Nat. Struct. Biol.* **1999**, *6*, 937–943.
- (a) Lamarre, D.; Anderson, P. C.; Bailey, M.; Beaulieu, P.; Bolger, G.; Bonneau, P.; Bös, M.; Cameron, D. R.; Cartier, M.; Cordingley, M. G.; Faucher, A.-M.; Goudreau, N.; Kawai, S. H.; Kukolj, G.; Lagacé, L.; LaPlante, S. R.; Narjes, H.; Poupart, M.-A.; Rancourt, J.; Sentjens, R. E.; George, T. S.; Simoneau, B.; Steinmann, G.; Thibeault, D.; Tsantrizos, Y. S.; Weldon, S. M.; Yong, C.-L.; Llinàs-Brunet, M. An NS3 Protease Inhibitor with Antiviral Effects in Humans Infected with Hepatitis C Virus. *Nature* **2003**, *426*, 186–189. (b) Perni, R. B.; Almqvist, S. J.; Byrn, R. A.; Chandorkar, G.; Chaturvedi, P. R.; Courtney, L. F.; Decker, C. J.; Dinehart, K.; Gates, C. A.; Harbeson, S. L.; Heiser, A.; Kalkeri, G.; Kolaczowski, E.; Lin, K.; Luong, Y.-P.; Rao, B. G.; Taylor, W. P.; Thomson, J. A.; Tung, R. D.; Wei, Y.; Kwong, A. D.; Lin, C. Preclinical Profile of VX-950, a Potent, Selective, And Orally, Bioavailable Inhibitor of Hepatitis C Virus NS3–NS4A Serine Protease. *Antimicrob. Agents Chemother.* **2006**, *50*, 899–909. (c) Saha, A. K.; Young, C.; Del Vecchio, A. M.; Bailey, T. A.; Reinhardt, J. A.; Kulkarni, B. A.; Falt, T. H.; Feng, H.; Rippin, S. R.; Blackledge, C. W.; Rys, D. J.; Lessen, T. A.; Swestock, J.; Deng, Y.; Nitz, T. J.; Chanduru, S.; Chopra, R.; Collett, M.; Pevear, D.; Howe, A. Y. M.; O'Connell, J.; Mansour, T.; Burns, C. J. Discovery of Hcv - 796: A Potent and Orally Bioavailable Hepatitis-C Polymerase Inhibitor under Clinical Development. Abstracts of Papers, 233rd ACS National Meeting, Chicago, IL, USA, March 25–29, 2007.
- Houghton, M.; Abrignani, S. Prospects for a Vaccine against the Hepatitis C Virus. *Nature* **2005**, *436*, 961–966.
- Lohmann, V.; Korner, F.; Koch, J.-O.; Herian, U.; Theilmann, L.; Bartenschlager, R. Replication of Subgenomic Hepatitis C Virus RNAs in a Hepatoma Cell Line. *Science* **1999**, *285*, 110–113.
- Mercer, D. F.; Schiller, D. E.; Elliott, J. F.; Douglas, D. N.; Hao, C.; Rinfret, A.; Addison, W. R.; Fischer, K. P.; Churchill, T. A.; Lakey, J. R. T.; Tyrrell, D. L. J.; Kneteman, N. M. Hepatitis C Virus Replication in Mice with Chimeric Human Livers. *Nat. Med.* **2001**, *7*, 927–933.
- Walker, C. M. Comparative Features of Hepatitis C Virus Infection in Humans and Chimpanzees. *Springer Semin. Immunopathol.* **1997**, *19*, 85–98.
- For a review, see: De Francesco, R.; Migliaccio, G. Challenges and Successes in Developing New Therapies for Hepatitis C. *Nature* **2005**, *436*, 953–960.
- (a) Llinàs-Brunet, M.; Bailey, M. D.; Fazal, G.; Goulet, S.; Halmos, T.; LePlante, S.; Maurice, R.; Poirier, M.; Poupart, M.-A.; Thibeault, D.; Wernic, D.; Lamarre, D. Peptide-Based Inhibitors of the Hepatitis C Virus Serine Protease. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1713–1718. (b) Steinkuehler, C.; Biasiol, G.; Brunetti, M.; Urbani, A.; Koch, U.; Cortese, R.; Pessi, A.; De Francesco, R. Product Inhibition of the Hepatitis C Virus NS3 Protease. *Biochemistry* **1998**, *37*, 8899–8905.
- (a) Venkatraman, S.; Bogen, S. L.; Arasappan, A.; Bennett, F.; Chen, K.; Jao, J.; Liu, Y.-T.; Lovey, R.; Hendrata, S.; Huang, Y.; Pan, W.; Parekh, T.; Pinto, P.; Popov, V.; Pike, R.; Ruan, S.; Santhanam, B.; Vibulbhan, B.; Wu, W.; Yang, W.; Kong, J.; Liang, X.; Wong, J.; Liu, R.; Butkiewicz, N.; Chase, R.; Hart, A.; Agrawal, S.; Ingravallo, P.; Pichardo, J.; Kong, R.; Baroudy, B.; Malcolm, B.; Guo, Z.; Prongay, A.; Madison, V.; Broske, L.; Cui, X.; Cheng, K.-C.; Hsieh, Y.; Brisson, J.-M.; Prelusky, D.; Korfmacher, W.; White, R.; Bogdanowich-Knipp, S.; Pavlovsky, A.; Bradley, P.; Saksena, A.; K.; Ganguly, A.; Piwinski, J.; Girijavallabhan, V.; Njoroge, F. G. Discovery of (1*R*,5*S*)-*N*-[3-Amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(*S*)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(*S*)-carboxamide (SCH 503034), a Selective, Potent, Orally Bioavailable Hepatitis C Virus NS3 Protease Inhibitor: A Potential Therapeutic Agent for the Treatment of Hepatitis C Infection. *J. Med. Chem.* **2006**, *49*, 6074–6086. (b) Malcolm, B. A.; Liu, R.; Lahser, F.; Agrawal, S.; Belanger, B.; Butkiewicz, N.; Chase, R.; Gheyas, F.; Hart, A.; Hesk, D.; Ingravallo, P.; Jiang, C.; Kong, R.; Lu, J.; Pichardo, J.; Prongay, A.; Skelton, A.; Tong, X.; Venkatraman, S.; Xia, E.; Girijavallabhan, V.; Njoroge, F. G. SCH 503034, a Mechanism-Based Inhibitor of Hepatitis C Virus NS3 Protease, Suppresses Polyprotein Maturation and Enhances the Antiviral Activity of Alpha Interferon in Replicon Cells. *Antimicrob. Agents Chemother.* **2006**, *50*, 1013–1020.
- For reviews about protease inhibitors, see: (a) Turk, B. Targeting Proteases: Successes, Failures and Future Prospects. *Nat. Rev. Drug Discovery* **2006**, *5*, 785–799. (b) Tyndall, J. D.; Nall, T.; Fairlie, D. P. Proteases Universally Recognize Beta Strands in Their Active Sites. *Chem. Rev.* **2005**, *105*, 973–999. (c) Powers, J. C.; Asgian, J. L.; Ekici, D.; James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002**, *102*, 4639–4643. (d) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease Inhibitors: Current Status and Future Prospects. *J. Med. Chem.* **2000**, *43*, 305–341.
- Zhang, R.; Beyer, B. M.; Durkin, J.; Ingram, R.; Njoroge, F. G.; Windsor, W. T.; Malcolm, B. A. A Continuous Spectrophotometric Assay for the Hepatitis C Virus Serine Protease. *Anal. Biochem.* **1999**, *270*, 268–275. The substrate Ac-DTEVVP(Nva)-O-PAP was used in the present study.

- 18 For a definition of  $K_i^*$  and discussions, see: Morrison, J. F.; Walsh, C. T. In *Advances in Enzymology and Related Areas of Molecular Biology*, Meister, A., Ed.; John Wiley & Sons, Inc., Hoboken, NJ, 1988; Vol 61, pp 201–301.
- 19 Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
- 20 Bogen, S.; Ruan, R.; Saksena, A. K.; Njoroge, F. G.; Girijavallabhan, V.; Agrawal, S.; Liu, R.; Pichardo, J.; Baroudy, B.; Prongay, A. Depeptidization Efforts on  $P_3$ - $P_2'$   $\alpha$ -Ketoamide Inhibitors of HCV NS3–NS4A Serine Protease: Effect on HCV Replicon Activity. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1621–1627.
- 21 Bogen, S. L.; Arasappan, A.; Bennett, F.; Chen, K.; Jao, E.; Liu, Y.-T.; Lovey, R. G.; Venkatraman, S.; Pan, W.; Parekh, T.; Pike, R. E.; Ruan, S.; Liu, R.; Baroudy, B.; Agrawal, S.; Ingravallo, P.; Pichardo, J.; Prongay, A.; Brisson, J.-M.; Hsieh, T. Y.; Cheng, K.-C.; Kemp, S. J.; Levy, O. E.; Lim-Wilby, M.; Tamura, S. Y.; Saksena, A. K.; Girijavallabhan, V.; Njoroge, F. G. Discovery of SCH 446211(SCH6): A New Ketoamide Inhibitor of the HCV NS3 Serine Protease and HCV Subgenomic RNA Replication NS3 Serine. *J. Med. Chem.* **2006**, *49*, 2750–2757.
- 22 Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- 23 Chen, K. X.; Njoroge, F. G.; Arasappan, A.; Venkatraman, S.; Vibulbhan, B.; Yang, W.; Parekh, T. N.; Pichardo, J.; Prongay, A.; Cheng, K.-C.; Butkiewicz, N.; Yao, N.; Madison, V.; Girijavallabhan, V. Novel Potent Hepatitis C Virus NS3 Serine Protease Inhibitors Derived from Proline-Based Macrocycles. *J. Med. Chem.* **2006**, *49*, 995–1005.
- 24 Velazquez, F.; Venkatraman, S.; Wu, W.; Blackman, M.; Prongay, A.; Girijavallabhan, V.; Shih, N.-Y.; Njoroge, F. G. Application of Ring-Closing Metathesis for the Synthesis of Macrocyclic Peptidomimetics as Inhibitors of HCV NS3 Protease. *Org. Lett.* **2007**, *9*, 3061–3064.