2010-12-07

Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding

Keith P. Romano
University of Massachusetts Medical School

Et al.

Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding

Keith P. Romano[1], Akbar Ali[1], William E. Royer, and Celia A. Schiffer[2]

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605

Edited by John M. Coffin, Tufts University School of Medicine, Boston, MA, and approved September 14, 2010 (received for review May 13, 2010)

Hepatitis C virus infects an estimated 180 million people worldwide, prompting enormous efforts to develop inhibitors targeting the essential NS3/4A protease. Resistance against the most promising protease inhibitors, telaprevir, boceprevir, and ITMN-191, has emerged in clinical trials. In this study, crystal structures of the NS3/4A protease domain reveal that viral substrates bind to the protease active site in a conserved manner defining a consensus volume, or substrate envelope. Mutations that confer the most severe resistance in the clinic occur where the inhibitors protrude from the substrate envelope, as these changes selectively weaken inhibitor binding without compromising the binding of substrates. These findings suggest a general model for predicting the susceptibility of protease inhibitors to resistance: drugs designed to fit within the substrate envelope will be less susceptible to resistance, as mutations affecting inhibitor binding would simultaneously interfere with the recognition of viral substrates.

drug design | hepatitis C | substrate envelope

Drug resistance is a major obstacle in the treatment of quickly evolving diseases. Hepatitis C virus (HCV) is a genetically diverse Hepacivirus of the Flaviviridae family infecting an estimated 180 million people worldwide (1). The viral RNA genome is translated as a single polyprotein and subsequently processed by host-cell and viral proteases into structural (C, E1, E2, and p7) and nonstructural (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins (2). The viral RNA-dependent RNA polymerase, NS5B, is inherently inaccurate and misincorporation of bases accounts for a very high mutation rate (3). While some mutations are neutral, others will alter the viability of the virus and propagate with varying efficiencies in each patient. Thus HCV infected individuals will develop a heterogeneous population of virus variants known as quasispecies (4). As patients begin treatment, the selective pressures of antiviral drugs will favor drug resistant variants (5). Therefore, an inhibitor must not only recognize one protein variant, but an ensemble of related enzymes. A detailed understanding of the atomic mechanisms of resistance is essential to effectively combat drug resistance against HCV antivirals.

The essential HCV NS3/4A protease is an attractive therapeutic target responsible for cleaving at least four sites along the viral polyprotein. These sites share little sequence homology except for an acid at position P6, Cys or Thr at P1, and Ser or Ala at P1′ (Table S1). The first cleavage event at the 3-4A junction occurs *in cis* as a unimolecular process, while the remaining substrates are processed bimolecularly *in trans*. The NS3/4A protease also cleaves the human cellular targets TRIF and MAVS, which confounds the innate immune response to viral infection (6–8). Early drug design efforts were hampered by the relatively shallow, featureless architecture of the protease active site. The eventual observation of N-terminal product inhibition served as a stepping stone for the discovery of more potent peptidomimetic inhibitors (9, 10). Over the past decade, pharmaceutical companies have further developed these lead compounds. Many structure-activity-relationship (SAR) studies have been performed to evaluate the effect of different functional moieties on protease inhibition at positions P4-P1′ (11–17). Crystal structures have been determined of the NS3/4A protease domain bound to a variety of inhibitors as well as of several drug resistant protease variants, such as R155K and V36M (18, 19). These data elucidate the molecular interactions of NS3/4A with inhibitors and the effect of specific drug resistance mutations on binding. These efforts, conducted in parallel by several pharmaceutical companies, led to the discovery of many protease inhibitors. Proof-of-concept for the successful clinical activity of this drug class was first demonstrated by the macrocyclic inhibitor BILN-2061 (Boehringer Ingelheim) (20, 21), which was later dropped from clinical trials in 2006 due to cardiotoxicity (22). Many other NS3/4A protease inhibitors are currently in development, and telaprevir (Vertex), boceprevir (Schering-Plough), and ITMN-191 (Intermune) lead the way in advanced phases of human clinical trials (Fig. 1*A*).

Despite these successes, the rapid acquisition of drug resistance has limited the efficacy of the most potent NS3/4A protease inhibitors in both replicon studies and human clinical trials (Fig. 1*B* and Table 1). In this study, we show that mutations conferring the most severe resistance occur where the protease extensively contacts the inhibitors but not the natural viral substrates. Four crystal structures of the NS3/4A protease domain in complex with the N-terminal products of viral substrates reveal a conserved mode of substrate binding, with the consensus volume defining the substrate envelope. The protease inhibitors ITMN-191 (3M5L), TMC435 (3KEE) (23), and boceprevir (2OC8) (24) protrude extensively from the substrate envelope in regions that correlate with known sites of resistance mutations. Most notably, the P2 moieties of all three drugs protrude to contact A156 and R155, which mutate to confer high-level resistance against nearly all drugs reported in the literature (25–30). These findings suggest that drug resistance results from a change in molecular recognition and imply that drugs designed to fit within the substrate envelope will be less susceptible to resistance, as mutations altering inhibitor binding will simultaneously interfere with the binding of substrates.

Results

Synthesis of ITMN-191. We synthesized the macrocyclic inhibitor ITMN-191 using a convergent reaction sequence described in *SI Text*. Briefly, the P2 and P1-P1′ fragments were preassembled

Author contributions: K.P.R., A.A., and C.A.S. designed research; K.P.R. and A.A. performed research; A.A. and W.E.R. contributed new reagents/analytic tools; K.P.R., W.E.R., and C.A.S. analyzed data; and K.P.R., A.A., W.E.R., and C.A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 3M5L, 3M5M, 3M5N, and 3M5O).

[1]K.P.R. and A.A. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: Celia.Schiffer@umassmed.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1006370107/-/DCSupplemental.

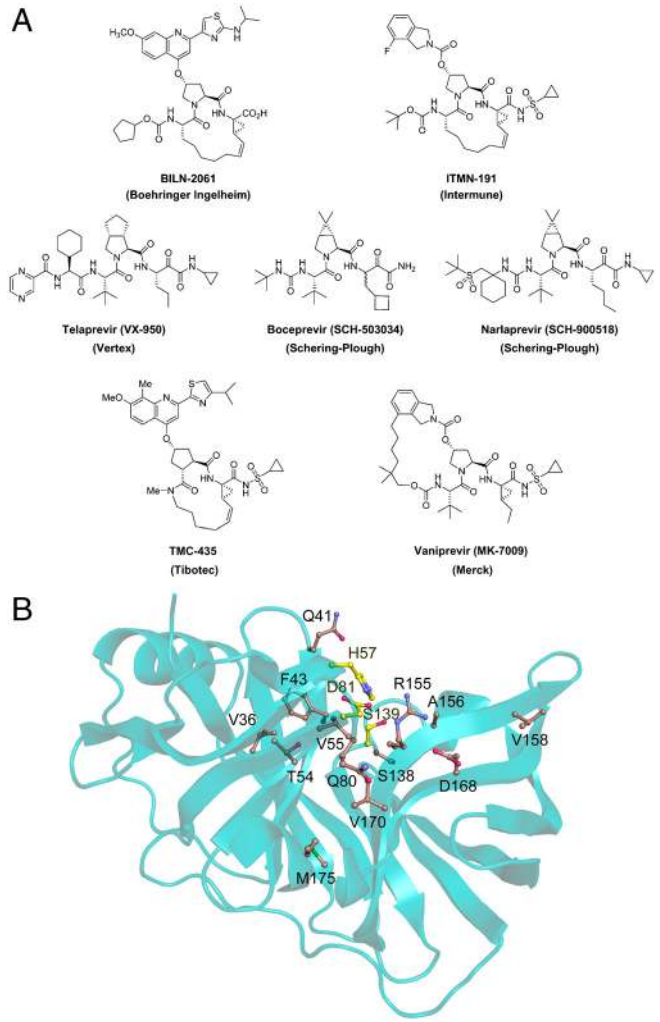

Fig. 1. NS3/4A protease inhibitors and reported sites of drug resistance. (*A*) The leading protease inhibitors in development mimic the N-terminal side of the viral substrates. (*B*) The majority of reported drug resistance mutations cluster around the protease active site with the catalytic triad depicted in yellow.

and the macrocyclic drug compound was generated by a four-step reaction sequence, including P2-P3 amide coupling, ester hydrolysis, coupling with the P1-P1′ fragment, and ring-closing metathesis. The P2-P3 fragment was assembled by coupling the commercially available Boc-protected amino acid (*S*)-2-(*tert*-butoxycarbonylamino)non-8-enoic acid (Acme Biosciences, Inc) with the preassembled P2 fragment, (3*R*, 5*S*)-5-(methoxycarbonyl)pyrrolidin-3-yl 4-fluoroisoindoline-2-carboxylate (31), using O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU)/diisopropylethylamine (DIPEA). Hydrolysis of the P2-P3 methyl ester with $LiOH.H_2O$ in a mixture of THF-MeOH-H_2O followed by coupling of the resulting acid under HATU/DIPEA conditions with the preassembled P1-P1′ fragment, (1*R*, 2*S*)-1-amino-N-(cyclopropylsulfonyl)-2-vinylcyclopropanecarboxamide (32), provided the bis-olefin precursor for ring-closing metathesis. Cyclization of the bis-olefin intermediate was accomplished using a highly efficient ring-closing metathesis catalyst Zhan 1B and provided the protease inhibitor ITMN-191.

Structure Determination of Inhibitor and Substrate Complexes. Although NS3/4A cleaves the viral polyprotein of over 3,000 residues at four specific sites in vivo, we focused on the local interactions of the protease domain with short peptide sequences corresponding to the immediate cleavage sites. All structural studies were carried out with the highly soluble, single-chain construct of the NS3/4A protease domain described previously (33), which contains a fragment of the essential cofactor NS4A covalently linked at the N terminus by a flexible linker. A similar protease construct was shown to retain comparable catalytic activity to the authentic protein complex (34). Crystallization trials were initially carried out using the inactive (S139A) protease variant in complex with substrate peptides spanning P7-P5′. The 4A4B substrate complex revealed cleavage of the scissile bond and no ordered regions for the C-terminal fragment of the substrate. Similar observations were previously described for two other serine proteases where catalytic activity was observed, presumably facilitated by water, despite Ala substitutions of the catalytic Ser (35, 36). Thus all subsequent crystallization trials with the NS3/4A protease were performed using N-terminal cleavage products of the viral substrates spanning P7-P1.

Table 1. Drug resistance mutations reported in replicon studies and clinical trials*

Residue	Mutation	Drug
V36	A, M, L, G	Boceprevir, telaprevir
Q41	R	Boceprevir, ITMN-191
F43	S, C, V, I	Boceprevir, telaprevir, ITMN-191, TMC435
V55	A	Boceprevir
T54	A, S	Boceprevir, telaprevir
Q80	K, R, H, G, L	TMC435
S138	T	ITMN-191, TMC435†
R155	K, T, I, M, G, L, S, Q	Boceprevir, telaprevir, ITMN-191, BILN-2061, TMC435
A156	V, T, S, I, G	Boceprevir, telaprevir, ITMN-191, BILN-2061, TMC435
V158	I	Boceprevir
D168	A, V, E, G, N, T, Y, H, I	ITMN-191, BILN-2061, TMC435
V170	A	Boceprevir, telaprevir
M175	L	Boceprevir

*References (18, 25, 26, 28, 30–37).
†TMC435 displays reduced activity against S138T, but the mutation was not observed in selection experiments.

NS3/4A crystal structures in complex with ITMN-191 and peptide products 4A4B, 4B5A, and 5A5B were determined and refined at 1.25 Å, 1.70 Å, 1.90 Å, and 1.60 Å resolution, respectively (Table S2). The complexes crystallized in the space groups $P2_12_12_1$ and $P2_1$ with one, two, or four molecules in the asymmetric unit. The average B factors range from 16.8–29.7 $Å^2$ and there are no outliers in the Ramachandran plots. These structures represent the highest resolution crystal structures of NS3/4A protease reported to date.

Overall Structure Analysis. The NS3/4A protease domain adopts a tertiary fold characteristic of serine proteases of the chymotrypsin family (37, 38). A total of nine protease molecules were modeled in the four crystal structures solved in this study with an overall rms deviation (rmsd) of 0.28 Å. The rmsds reveal the five most variable regions of the protease to be (Fig. S1): (*i*) the linker connecting cofactor 4A at the N terminus, (*ii*) the loop containing residues 65–70, (*iii*) the zinc-binding site containing residues 95–105, (*iv*) the 3_{10} helix region spanning residues 128–136, and (*v*) the active site antiparallel β-sheet containing residues 156–168. These structural differences likely indicate inherent flexibility in the protease and do not appear to correlate with ligand type or active site occupancy.

Analysis of Product Complexes. Product complexes 4A4B, 4B5A, and 5A5B were further analyzed with the C terminus of the full-length NS3/4A structure (1CU1), which contains the N-terminal cleavage product of viral substrate 3-4A (39). All four products bind to the protease active site in a conserved manner (Fig. 2), forming an antiparallel β-sheet with residues 154–160

and burying 500–600 Å^2 of solvent accessible surface area as calculated by PISA (40). The peptide backbone torsions are very similar, being most conserved at position P1 and deviating slightly toward position P4. Eight hydrogen bonds between backbone amide and carbonyl groups are completely conserved, involving protease residues S159 (C159 in product 3-4A), A157, R155, S139A, S138, and G137. S159 (C159 in product 3-4A), and A157 each contribute two hydrogen bonds with the P5 and P3 peptide residues, respectively. All P1 terminal carboxyl groups sit in the oxyanion hole, hydrogen bonding with the Nϵ atom of H57 and the amide nitrogens of residues 137–139. Although only product 4B5A contains a proline at P2, the other substrate sequences still adopt constrained P2 φ torsion angles. Thus products bind similarly despite their high sequence diversity.

The P1 and P6 residues are most conserved among the substrate sequences, as are most of their interactions with the protease. The P1 side chains interact with the aromatic ring of F154. In all structures but product complex 4B5A, K165 forms salt-bridges with the P6 acids, while residues R123, D168, R155, and the catalytic D81 form an ionic network along one surface of the bound products (Fig. 2). In complex 4B5A, R123 interacts directly with the P6 acid, while D168 reorients and no longer contacts R155. Other molecular interactions in the product complexes are more diverse. Notably, K136 interacts differently with the cleavage products, forming: (*i*) a hydrogen bond with the P2 backbone carbonyl oxygen of 3-4A; (*ii*) a salt-bridge with the P3 glutamate of product 4A4B; and (*iii*) nonspecific van der Waals interactions with the P2 and P3 side chains of products 4B5A and 5A5B. Also, in product complex 4A4B, an intramolecular hydrogen bond forms between the P3 and P5 glutamate residues, while the unique P4 acid of product 3-4A forms salt-bridges with the guanidinium groups of R123 and R155. Thus distinct patterns of side chain interactions underlie the set of conserved features involved in NS3/4A cleavage product binding.

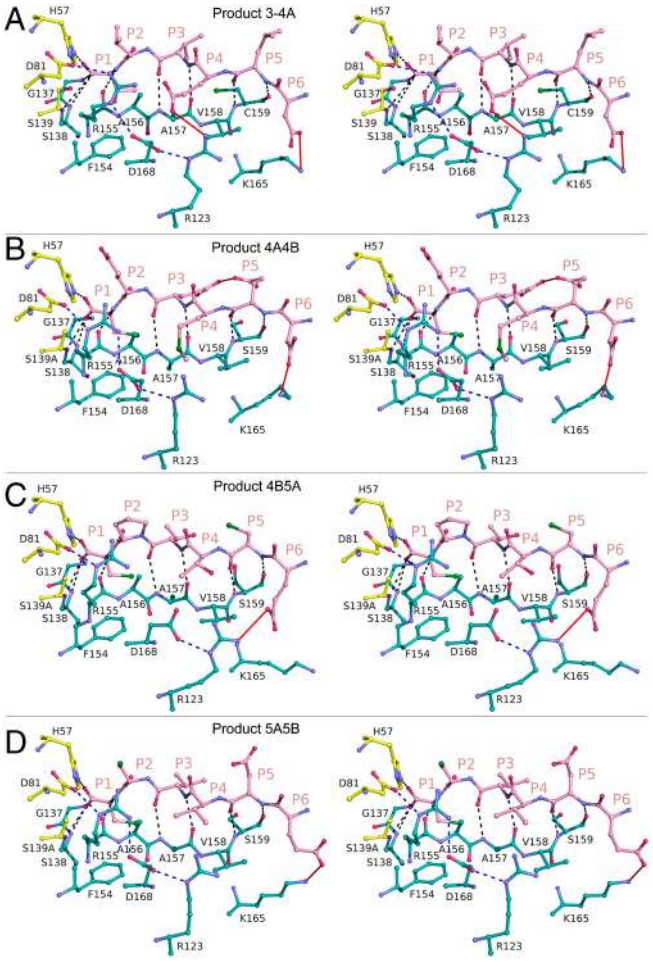

Fig. 2. Stereo view of N-terminal cleavage product binding to NS3/4A protease. N-terminal protease cleavage products (*A*) 3-4A, (*B*) 4A4B, (*C*) 4B5A, and (*D*) 5A5B are depicted as they bind to the protease active site. All conserved interactions are indicated by black dashes, while red lines depict interactions that are not present in all structures. The electrostatic network involving residues R123, D168, R155, and D81 is indicated by blue dashes.

The Substrate Envelope. To further analyze the structural similarities of the four NS3/4A product complexes, the active sites were superposed on the Cα atoms of residues 137–139 and 154–160, revealing that both the active site residues and substrate products spanning P6-P1 align closely with an average Cα rmsd of 0.24 Å and 0.35 Å, respectively. The consensus van der Waals volume shared by any three of the four cleavage products was then calculated to generate the NS3/4A substrate envelope (Fig. 3*A*). This shape could not be predicted by the primary sequences alone and highlights the conserved mode of viral substrate recognition despite their high sequence diversity.

Analysis of Inhibitor Complexes. ITMN-191, TMC435, and boceprevir are all peptidomimetic NS3/4A protease inhibitors. Active site superpositions of these drug complexes reveal that the inhibitors interact with many of the same protease residues as the cleavage products. Despite the P3-P1 cyclization of ITMN-191 and TMC435, the functional groups are positioned similarly in all three inhibitor complexes. The P1 cysteine surrogates interact with the aromatic ring of F154, while the P2 and P3 moieties overlap closely. Although TMC435 does not contain a P4 substituent, the P4 *tert*-butyl groups of ITMN-191 and boceprevir also align closely. In addition, the P1 and P3 backbone atoms of all inhibitors hydrogen bond with the carbonyl oxygens of R155 and A157, respectively. These observations verify the peptidomimetic nature of these drugs and support their observed mechanism as competitive active site inhibitors.

The largest variation between these three protease inhibitors occurs at P2 where the aromatic rings of ITMN-191 and TMC435 stack against the guanidinium group of R155 (Fig. 3). This molecular interaction alters the electrostatic network involving R123, D168, R155, and D81. R155 rotates nearly 180° around Cδ relative to its conformation observed in product complexes, losing its hydrogen bond with D81 but maintaining interaction with D168. Mutations at R155 or D168 would disrupt the electrostatic network and destabilize this packing thereby lowering the affinity of these macrocyclic drugs. This observation provides a structural rationale for the drug resistance mutations R155K, as previously proposed (19), and D168A/V, which both confer a selective advantage in vitro in the presence of ITMN-191 or TMC435 (26, 30). In addition, the TMC435 complex reveals that R155 is stabilized by a hydrogen bond with Q80, which also mutates to confer resistance to TMC435 (30). Thus many of the primary drug resistance mutations can be explained by the disruption of atomic interactions involving the P2 functional groups of the drugs.

Insights into Drug Resistance. To determine the locations where the inhibitors protrude from the substrate envelope, the inhibitor and product complexes were also superposed using residues 137–139 and 154–160. The van der Waals volumes of inhibitor protrusion from the substrate envelope (V_{out}) (41, 42) were calculated for each drug and compared with published EC_{50} fold-change data for drug resistance variants (30). The magnitudes of the EC_{50} fold-change data determined for each NS3/4A mutant generally trend with the V_{out} values for the three drugs. The P2 moieties of boceprevir, ITMN-191 and TMC435 protrude most extensively from the substrate envelope with V_{out} values of 105, 294, and

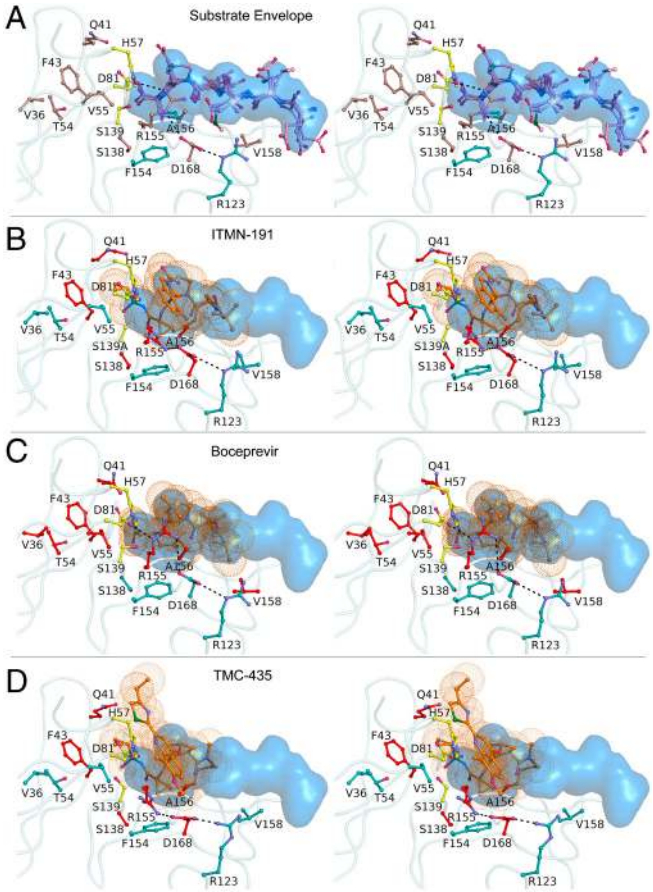

Fig. 3. Stereo view of the NS3/4A substrate envelope and protease inhibitors. (*A*) After active site superpositions, the overlapping van der Waals volume shared by any three of the four cleavage products defines the substrate envelope, depicted in blue. NS3/4A protease residues which mutate to confer drug resistance are shown in brown. (*B*) ITMN-191, (*C*) boceprevir and (*D*) TMC435 protrude from the substrate envelope at several locations, which correlate with known sites of drug-resistant mutations to each inhibitor, shown in red.

496 Å^3, respectively (Table 2). However, the precise level of drug resistance observed is also determined by the particular change in molecular interaction occurring for a given mutation. For example, A156 and R155 pack with the P2 moieties of these three inhibitors where they protrude beyond the substrate envelope. Mutations of A156 to bulkier side chains would result in a steric clash with the P2 drug moieties. Indeed, the rigid dimethylcyclopropane group of boceprevir protrudes from the substrate envelope at the P2 subsite, and A156V or A156T confer 65 and 75 fold-changes in EC_{50}, respectively (Table 2). Similarly, molecular changes at R155 and D168 would result in a substantial loss of interactions with P2. The most extensive protrusions of ITMN-191 and TMC435 at P2 trend with their greatest fold-change in potency of nearly 450 and 600, respectively, from mutations in this subsite. Thus the extent by which an inhibitor protrudes from the substrate envelope in a given subsite is indicative of its vulnerability to resistance.

Further structural analyses with the substrate envelope provide insights into other NS3/4A drug resistance mutations. The P1′ sulfonamide groups of ITMN-191 and TMC435, as well as the P1′ ketoamide of boceprevir, protrude from the substrate envelope near residues Q41 and F43, which both mutate to confer low-level resistance to these drugs (25, 30, 43). The keto group of boceprevir also projects outside the substrate envelope near T54 and V55. T54A/S confers low-level resistance to boceprevir, while V55A was recently identified in patient isolates after treatment with boceprevir (44). The analogous carbonyl groups of ITMN-191 and TMC435, however, are orientated in the opposite direction and protrude toward S138. In fact, in vitro studies reveal reduced activity for ITMN-191 and TMC435 against S138T variants, while boceprevir remains fully active (30, 43). The bulky P4 *tert*-butyl group of boceprevir extends outside the substrate envelope contacting V158; the V158I variant has lower affinity for this drug, likely due to a steric clash (45). This variant may also impact the affinity of ITMN-191, as its P4 *tert*-butyl also protrudes at the same location. These findings demonstrate that in regions outside the P2 subsite, positions where ITMN-191, TMC435, and boceprevir protrude from the substrate envelope also correlate with many other known sites of drug resistance mutations.

Table 2. EC_{50} fold-change (FC) data * for several NS3/4A drug resistant variants tabulated with V_{out}, the van der Waals volume of protrusion from the substrate envelope, at each subsite of the enzyme

		Boceprevir		ITMN-191		TMC435	
Subsite	Resistance mutation	V_{out} (Å^3)	EC_{50} FC	V_{out} (Å^3)	EC_{50} FC	V_{out} (Å^3)	EC_{50} FC
Total		292		500		649	
P1		76		67		64	
P2		105		294		496	
	Q80R		0.5		3.5		6.9
	Q80K		0.8		2.3		7.7
	R155K		4.7		447		30
	A156V		75		63		177
	A156T		65		41		44
	D168A		0.7		153		594
	D168E		0.8		75		40
P3		34		70		67	
P4		76		69		0	
	V158I		3.3†		ND		ND

*Antiviral activity was reported previously by Lenz et al., 2010 (30).
†Fold-change in EC_{50} reported in replicon assay by Qiu et al., 2009 (45).

Discussion

The emergence of drug resistance is a major obstacle in modern medicine that limits the long-term usefulness of the most promising therapeutics. By considering how HCV NS3/4A protease inhibitors bind relative to natural viral substrates, we discovered that primary sites of resistance occur in regions of the protease where drugs protrude from the substrate envelope. In particular, R155 and A156, which mutate to confer severe resistance against ITMN-191, TMC435, and boceprevir, interact closely with the P2 drug moieties where they protrude most extensively from the substrate envelope. Molecular changes at these residues confer resistance by selectively weakening inhibitor binding without compromising the binding of viral substrates. We further speculate that these mutations will not considerably affect the binding of the host cellular substrates TRIF and MAVS, which likely fit well within the substrate envelope as they share many features with the viral substrates. However, TRIF contains a track of eight proline residues instead of an acidic residue at position P6, which may modulate its binding. Further structural studies are warranted to better ascertain the molecular details of how these cellular substrates are recognized by the NS3/4A protease.

Although this study focuses on ITMN-191, TMC435, and boceprevir, other NS3/4A protease inhibitors in clinical trials, including telaprevir, narlaprevir, and vaniprevir (Fig. 1*A*), contain similar functional groups that likely protrude from the substrate envelope. Most notably, all these drug candidates contain bulky P2 moieties and are therefore susceptible to cross-resistance against mutations at R155 and A156. R155 and A156 mutations have been shown to confer telaprevir resistance in treated patients (46). Cross-resistance studies have also shown that nar-

laprevir displays similar fold losses in activity against most of the known drug resistance mutations for telaprevir and boceprevir (47). Ultimately, to slow the emergence of multidrug resistant viral strains, inhibitors should be confined within the substrate envelope, particularly at the P2 position. To compensate for the loss of binding affinity that will likely accompany these changes, additional interactions could potentially be optimized spanning the S4-S6 subsites of the protease and the catalytic triad.

Our findings further suggest a general model for using the substrate envelope to predict patterns of drug resistance in other quickly evolving diseases. For drug resistance to occur, mutations must selectively weaken a target's affinity for an inhibitor without significantly altering its natural biological function. Mutations occurring outside the substrate envelope are better able to achieve this effect, as these molecular changes can selectively alter inhibitor binding without compromising the binding of natural substrates. Whenever the interaction of a drug target with its biological substrates can be structurally characterized, we predict that drugs designed to fit within the substrate envelope will be less susceptible to resistance. Structure-based design strategies can utilize this model as an added constraint to develop inhibitors that fit within the substrate envelope. In fact, previous work in our laboratory provides proof-of-concept for the successful incorporation of the substrate envelope in the design of unique HIV protease inhibitors, which maintain high affinities against a panel of multidrug resistant variants of HIV-1 protease (42, 48–53). As a general paradigm, design efforts incorporating the substrate envelope would facilitate a more rationale evaluation of drug candidates and lead to the development of more robust inhibitors that are less susceptible to resistance.

Materials and Methods

Protein Crystallization. The NS3/4A protease construct was expressed and purified as reported previously (33, 54), detailed in *SI Text*. Purified protein was concentrated to ~3 mg/mL and loaded on a HiLoad Superdex75 16/60 column equilibrated with 25 mM 2-(N-morpholino)ethanesulfonic acid (MES) at pH 6.5, 500 mM NaCl, 10% glycerol, 30 μM zinc chloride, and 2 mM DTT. The protease fractions were pooled and concentrated to 20–25 mg/mL with an Amicon Ultra-15 device (10 kD; Millipore). The concentrated samples were then incubated at 4 °C for 1 h with 2–20 M excess of viral substrate 4A4B, peptide products 4B5A or 5A5B, or ITMN-191. Information about the synthesis of viral peptides and ITMN-191 is provided in *SI Text*. Diffraction-quality crystals were obtained overnight for all ligands by mixing equal volume of concentrated protein solution with precipitant solution (20–26% PEG-3350, 0.1 M sodium MES buffer at pH 6.5, and 4% ammonium sulfate) in 24-well VDX hanging drop trays.

Data Collection and Structure Solution. Crystals were flash-frozen in liquid nitrogen and mounted under constant cryostream. X-ray diffraction data were collected at Advanced Photon Source BioCARS 14-IDB, 14-BMC, and LS-CAT 21-ID-F. Diffraction intensities of the complexes were indexed, integrated, and scaled using the programs HKL2000 (55) and XDS (56). 5% of the data was used to calculate R-free (57). All structure solutions were generated using isomorphous molecular replacement with PHASER (58) or AMORE. The NS3/4A protease domain (PDB code 2A4G) (59) was used for molecular replacement in solving the product 4A4B structure, and this structure was subsequently used for solving the other complexes. In all cases, initial refinement was carried out in the absence of modeled ligand, which was subsequently built in. Phases were improved using ARP/wARP (60). Iterative rounds of translation, libration, and screw (TLS) and restrained refinement with CCP4 (61) and graphical model building with COOT (62) until convergence was achieved. The final structures were evaluated with MolProbity (63) prior to deposition in the protein data bank.

Structural Analysis. Double-difference plots (64) were used to determine the structurally invariant regions of the protease, consisting of residues 32–36, 42–47, 50–54, 84–86, and 140–143. Structures were superposed with PyMOL (65) using the Cα atoms of these residues for all protease molecules from the solved structures (nine total). The B chain of product complex 4A4B was used as the reference structure in all alignments. Fit of individual inhibitors into the substrate envelope was quantified by mapping the substrate envelope and the van der Waals volume of each inhibitor on a three-dimensional grid with spacing of 0.5 Å. V_{out} for each drug moiety was computed by counting the grid cells, which were occupied by any inhibitor atom of that site but not the substrate envelope, and multiplied by the grid cell size, 0.125 $Å^3$ (41, 42).

Substrate Envelope and Inhibitor Analyses. NS3/4A substrate envelope was computed using product complexes 4A4B (B chain), 4B5A (D chain), and 5A5B (A chain). In structures with multiple protease molecules in the asymmetric unit, the one containing the most ordered peptide product was used for the alignment. The protease domain of the full-length NS3/4A structure (A chain; PDB code 1CU1) (39), including the C-terminal six amino acids, was included as a product complex 3-4A. All active site alignments were performed in PyMOL using Cα atoms of protease residues 137–139 and 154–160. After superposition, Gaussian object maps were generated in PyMOL for each cleavage product. Four consensus Gaussian maps were then calculated, representing the intersecting volume of a group of three object maps. Finally, the summation of these four consensus maps was generated to construct the substrate envelope, depicting the van der Waals volume shared by any three of the four products. The previously determined boceprevir complex (PDB code 2OC8) (24) and TMC435 complex (PDB code 3KEE) (23) were used in this study (66).

ACKNOWLEDGMENTS. We thank H. Klei for helpful discussions. We also thank Z. Wawrzak, M. Bolbat, and K. Brister of the LS-CAT beamline at Argonne National Laboratory for data collection of the ITMN-191 complex; M. Nalam and R. Bandaranayake for assistance with structural refinement; A. Ozen for providing V_{out} calculations; and S. Shandilya and Y. Cai for computational support. The National Institute of Health (NIH) Grants R01-GM65347 and R01-AI085051 supported this work. Use of Advanced Photon Source (APS) was supported by the Department of Energy (DOE), Basic Energy Sciences, Office of Science, under Contract No. DE-AC02-06CH11357. Use of the BioCARS Sector 14 was supported by NIH-NCRR RR007707. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology TriCorridor under Grant 085P1000817.

1. World Health Organization. Barnes E, ed. (2010) Vaccine research: hepatitis C. *Hepatitis C Virus: Disease Burden*. Available at http://www.who.int/vaccine_research/diseases/viral_cancers/en/index2.html. March 15, 2010.
2. Major ME, Feinstone SM (1997) The molecular virology of hepatitis C. *Hepatology* 25:1527–1538.
3. Qureshi SA (2007) Hepatitis C virus—biology, host evasion strategies, and promising new therapies on the horizon. *Med Res Rev* 27:353–373.
4. Martell M, et al. (1992) Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J Virol* 66:3225–3229.
5. Paolucci S, et al. (2001) Analysis of HIV drug-resistant quasispecies in plasma, peripheral blood mononuclear cells and viral isolates from treatment-naive and HAART patients. *J Med Virol* 65:207–217.
6. Chen Z, et al. (2007) GB virus B disrupts RIG-I signaling by NS3/4A-mediated cleavage of the adaptor protein MAVS. *J Virol* 81:964–976.
7. Li XD, Sun L, Seth RB, Pineda G, Chen ZJ (2005) Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proc Natl Acad Sci USA* 102:17717–17722.
8. Li K, et al. (2005) Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. *Proc Natl Acad Sci USA* 102:2992–2997.
9. Llinas-Brunet M, et al. (1998) Peptide-based inhibitors of the hepatitis C virus serine protease. *Bioorg Med Chem Lett* 8:1713–1718.
10. Steinkuhler C, et al. (1998) Product inhibition of the hepatitis C virus NS3 protease. *Biochemistry* 37:8899–8905.
11. Arasappan A, et al. (2006) P2-P4 macrocyclic inhibitors of hepatitis C virus NS3-4A serine protease. *Bioorg Med Chem Lett* 16:3960–3965.
12. Bogen S, et al. (2008) Hepatitis C virus NS3-4A serine protease inhibitors: SAR of new P1 derivatives of SCH 503034. *Bioorg Med Chem Lett* 18:4219–4223.
13. Malancona S, et al. (2004) SAR and pharmacokinetic studies on phenethylamide inhibitors of the hepatitis C virus NS3/NS4A serine protease. *Bioorg Med Chem Lett* 14:4575–4579.
14. Nilsson M, et al. (2010) Synthesis and SAR of potent inhibitors of the hepatitis C virus NS3/4A protease: exploration of P2 quinazoline substituents. *Bioorg Med Chem Lett* 20:4004–4011.
15. Venkatraman S, et al. (2009) Discovery and structure-activity relationship of P1-P3 ketoamide derived macrocyclic inhibitors of hepatitis C virus NS3 protease. *J Med Chem* 52:336–346.
16. Raboisson P, et al. (2008) Structure-activity relationship study on a novel series of cyclopentane-containing macrocyclic inhibitors of the hepatitis C virus NS3/4A protease leading to the discovery of TMC435350. *Bioorg Med Chem Lett* 18:4853–4858.
17. Perni RB, et al. (2004) Inhibitors of hepatitis C virus NS3.4A protease 2. Warhead SAR and optimization. *Bioorg Med Chem Lett* 14:1441–1446.

18. Zhou Y, et al. (2008) Phenotypic characterization of resistant Val36 variants of hepatitis C virus NS3-4A serine protease. *Antimicrob Agents Chemother* 52:110–120.
19. Zhou Y, et al. (2007) Phenotypic and structural analyses of hepatitis C virus NS3 protease Arg155 variants: sensitivity to telaprevir (VX-950) and interferon alpha. *J Biol Chem* 282:22619–22628.
20. Lamarre D, et al. (2003) An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature* 426:186–189.
21. Hinrichsen H, et al. (2004) Short-term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients. *Gastroenterology* 127:1347–1355.
22. Vanwolleghem T, et al. (2007) Ultra-rapid cardiotoxicity of the hepatitis C virus protease inhibitor BILN 2061 in the urokinase-type plasminogen activator mouse. *Gastroenterology* 133:1144–1155.
23. Cummings MD, et al. (2010) Induced-fit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target. *Angew Chem Int Ed Engl* 49:1652–1655.
24. Prongay AJ, et al. (2007) Discovery of the HCV NS3/4A protease inhibitor (1R,5S)-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) II. Key steps in structure-based optimization. *J Med Chem* 50:2310–2318.
25. Tong X, et al. (2008) Characterization of resistance mutations against HCV ketoamide protease inhibitors. *Antiviral Res* 77:177–185.
26. He Y, et al. (2008) Relative replication capacity and selective advantage profiles of protease inhibitor-resistant hepatitis C virus (HCV) NS3 protease mutants in the HCV genotype 1b replicon system. *Antimicrob Agents Chemother* 52:1101–1110.
27. Sarrazin C, et al. (2007) SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders. *Gastroenterology* 132:1270–1278.
28. Kieffer TL, et al. (2007) Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. *Hepatology* 46:631–639.
29. Yi M, Ma Y, Yates J, Lemon SM (2009) Transcomplementation of an NS2 defect in a late step in hepatitis C virus (HCV) particle assembly and maturation. *PLoS Pathog* 5:e1000403.
30. Lenz O, et al. (2010) In vitro resistance profile of the HCV NS3/4A protease inhibitor TMC435. *Antimicrob Agents Chemother* 54:1878–1887.
31. Arasappan A, Njoroge FG, Girijavallabhan VM (2005) *PCT Int.Appl. WO 2005/113581* Patent Application.
32. Wang XA, et al. (2006) *US Patent* 6995174.
33. Wittekind M, Weinheirner S, Zhang Y, Goldfarb V (2002) *US Patent* 6333186.
34. Taremi SS, et al. (1998) Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis C virus protease. *Protein Sci* 7:2143–2149.
35. Carter P, Wells JA (1988) Dissecting the catalytic triad of a serine protease. *Nature* 332:564–568.
36. Krishnan R, Sadler JE, Tulinsky A (2000) Structure of the Ser195Ala mutant of human alpha—thrombin complexed with fibrinopeptide A(7—16): evidence for residual catalytic activity. *Acta Crystallogr D* 56:406–410.
37. Kim JL, et al. (1996) Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* 87:343–355.
38. Bode W, Huber R (1978) Crystal structure analysis and refinement of two variants of trigonal trypsinogen: trigonal trypsin and PEG (polyethylene glycol) trypsinogen and their comparison with orthorhombic trypsin and trigonal trypsinogen. *FEBS Lett* 90:265–269.
39. Yao N, Reichert P, Taremi SS, Prosise WW, Weber PC (1999) Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* 7:1353–1363.
40. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
41. Nalam MN, et al. (2010) Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. *J Virol* 84:5368–5378.
42. Chellappan S, Kairys V, Fernandes MX, Schiffer C, Gilson MK (2007) Evaluation of the substrate envelope hypothesis for inhibitors of HIV-1 protease. *Proteins* 68:561–567.
43. Kuntzen T, et al. (2008) Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naive patients. *Hepatology* 48:1769–1778.
44. Susser S, et al. (2009) Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. *Hepatology* 50:1709–1718.
45. Qiu P, et al. (2009) Identification of HCV protease inhibitor resistance mutations by selection pressure-based method. *Nucleic Acids Res* 37:e74.
46. Sarrazin C, et al. (2007) Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology* 132:1767–1777.
47. Tong X, et al. (2010) Preclinical characterization of the antiviral activity of SCH 900518 (Narlaprevir), a novel mechanism-based inhibitor of hepatitis C virus NS3 protease. *Antimicrob Agents Chemother* 54:2365–2370.
48. Chellappan S, et al. (2007) Design of mutation-resistant HIV protease inhibitors with the substrate envelope hypothesis. *Chem Biol Drug Des* 69:298–313.
49. Ali A, et al. (2009) Substrate envelope based design of new HIV-1 protease inhibitors active against drug-resistant HIV-1. *238th ACS National Meeting, Washington, DC, United States* (American Chemical Society, Washington, DC), MEDI-102.
50. King NM, Prabu-Jeyabalan M, Nalivaika EA, Schiffer CA (2004) Combating susceptibility to drug resistance: lessons from HIV-1 protease. *Chem Biol* 11:1333–1338.
51. Prabu-Jeyabalan M, et al. (2006) Substrate envelope and drug resistance: crystal structure of RO1 in complex with wild-type human immunodeficiency virus type 1 protease. *Antimicrob Agents Chemother* 50:1518–1521.
52. Altman MD, Nalivaika EA, Prabu-Jeyabalan M, Schiffer CA, Tidor B (2008) Computational design and experimental study of tighter binding peptides to an inactivated mutant of HIV-1 protease. *Proteins* 70:678–694.
53. Altman MD, et al. (2008) HIV-1 protease inhibitors from inverse design in the substrate envelope exhibit subnanomolar binding to drug-resistant variants. *J Am Chem Soc* 130:6099–6113.
54. Gallinari P, et al. (1998) Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. *J Virol* 72:6758–6769.
55. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology,* 276 pp:307–326.
56. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 26:795–800.
57. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475.
58. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
59. Arasappan A, et al. (2005) Hepatitis C virus NS3-4A serine protease inhibitors: SAR of P′ 2 moiety with improved potency. *Bioorg Med Chem Lett* 15:4180–4184.
60. Morris RJ, Perrakis A, Lamzin VS (2002) ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr D* 58:968–975.
61. Collaborative Computational Project Number 4 (1994) The CCP4 Suite: programs for protein crystallography. *Acta Cryst,* D50 pp:760–763.
62. Emsley P, Cowtan K (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
63. Davis IW, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–383.
64. Prabu-Jeyabalan M, Nalivaika EA, Romano K, Schiffer CA (2006) Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. *J Virol* 80:3607–3616.
65. DeLano WL (2008) *The PyMOL Molecular Graphics System* (DeLano Scientific LLC, San Carlos, CA).
66. Berman HM, et al. (2000) The protein data bank. *Nucleic Acids Res* 28:235–242.