

The Design and Enzyme-Bound Crystal Structure of Indoline Based Peptidomimetic Inhibitors of Hepatitis C Virus NS3 Protease

Jesus M. Ontoria,[†] Stefania Di Marco,[‡]
Immacolata Conte,[†] M. Emilia Di Francesco,[†]
Cristina Gardelli,[†] Uwe Koch,[†] Victor G. Matassa,[§]
Marco Poma,[†] Christian Steinkühler,[‡]
Cinzia Volpari,[‡] and Steven Harper^{*†}

IRBM-MRL Rome, Via Pontina km 30,600,
00040 Roma, Italy

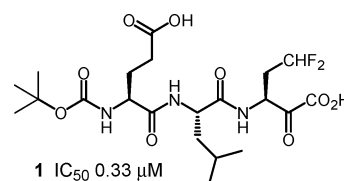
Received July 16, 2004

Abstract: The design of a series of peptidomimetic inhibitors of the hepatitis C virus NS3 protease is described. These inhibitors feature an indoline-2-carboxamide as a novel heterocyclic replacement for the P3 amino acid residue and N-terminal capping group of tripeptide based inhibitors. The crystal structure of the ternary NS3/NS4A/inhibitor complex for the most active molecule in this series highlights its suitability as an N-terminal capping group of a dipeptide inhibitor of the NS3 protease.

Despite improved therapeutic regimens for hepatitis C virus¹ (HCV) the disease remains the leading cause of liver transplantation in the United States² and affects over 170 million individuals worldwide.³ Current treatments based on interferon- α dosed in combination with ribavirin are often unsuccessful due to reemergence of virus following cessation of therapy, creating a pressing need for new antivirals to tackle HCV infection.⁴

The NS3 gene product encoded by the (+)-stranded RNA genome of HCV contains a trypsin-like serine protease enzyme that plays a key role in the viral life cycle.⁵ After liberation of its N-terminus from the initially translated HCV poly-protein, NS3 is responsible for four crucial poly-protein cleavages that release the viral nonstructural proteins NS3–NS5B. The catalytic activity of NS3 is augmented by formation of a complex with the viral NS4A protein, and heterodimeric NS3–NS4A is believed to be the physiologically relevant form of the enzyme.^{6,7} There is now compelling evidence to suggest that inhibition of NS3 is a viable strategy for the development of antiviral agents,⁸ validating the search for novel, potent, and drug-like inhibitors of this enzyme.

Medicinal chemistry efforts toward inhibitors of NS3 have been complicated by the unusual structural features of the enzyme. Crystal structures^{7,9} reveal that while the overall fold of the enzyme and the orientation of its catalytic machinery are in register with those of mammalian serine proteases, the viral enzyme differs markedly in structure at its substrate-binding site. Substrates bind in an extended conformation in a shallow and solvent exposed channel, the binding



energy being generated by hydrophobic contacts and electrostatic interactions along the length of this surface.¹⁰ Short synthetic substrates for the enzyme have not been found, and only peptides containing 10 or more amino acids (spanning the P6–P4' region) are processed.^{11,12} Remarkably, the N-terminal hexapeptide cleavage products of these substrates inhibit NS3¹³ and have served as a starting point for drug discovery efforts. Sequence optimization¹⁴ and the design of a non-nucleophilic mimetic for the canonical P1 cysteine residue¹⁵ led to a series of potent tripeptide “serine trap” inhibitors¹⁶ (e.g. **1**) that became the lead structures in our medicinal chemistry program. Here we report our efforts aimed at reducing the peptidic character of **1** through the design and optimization of a novel replacement for the N-terminal amino acid and capping group. We additionally disclose the enzyme bound crystal structure of one compound from this series that validates a C2-alkylated indoline ring as a peptidomimetic replacement for the P3 residue of tripeptide NS3 protease inhibitors.

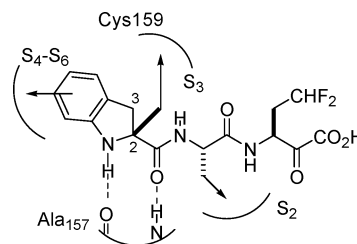


Figure 1. Design of indoline based peptidomimetic replacements for the N-terminal amino acid of **1**.

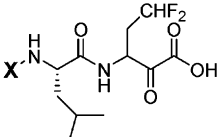
To replace the N-terminal capping group and P3 residue of **1**, we sought heterocyclic scaffolds that could (i) retain the double hydrogen bonding interaction with Ala157 observed in the enzyme bound crystal structure¹⁷ of peptide based inhibitors; (ii) orient a P2–P1 amino acid sequence as an active site “anchor”; (iii) direct pendant functionality to engage the P3/S3 side chain interaction; and (iv) function as a bridge to auxiliary binding regions, such as those accessed by the N-terminal capping group in the peptide series. Using the enzyme bound crystal structure of tripeptide keto-acid inhibitors as a starting point, a range of aromatic and nonaromatic nitrogen heterocycles were evaluated. Aromatic heterocycles appeared unattractive in view of the difficulty of accessing the S3 site from a substituent on an sp^2 hybridized ring carbon atom, leading to the selection of an indoline-2-carboxylic acid (Figure 1) as a promising scaffold. When the indoline engages the Ala157 NH and CO bonds in a twin hydrogen-bonding interaction, the substituent on its sp^3 hybridized C2 atom projects perpendicular to the plane of the indoline ring and (in one diastereoisomer) can access the S3

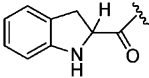
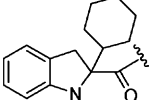
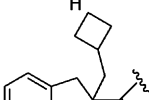
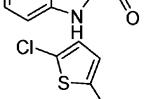
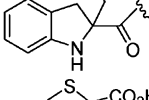
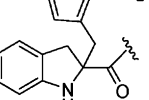
* To whom correspondence should be addressed. Phone: +39 06 91093444. Fax: +39 06 91093654. E-mail: Steven_Harper@Merck.com.

[†] Department of Medicinal Chemistry.

[‡] Department of Biochemistry.

[§] Present address: Graffinity Pharmaceuticals AG, Im Neuenheimer Feld 518, Heidelberg, Germany 69120.

Table 1. Structure and Biological Activity of Indole Based Ketoacid Inhibitors of HCV NS3 Protease


Entry	X	Cpd.	^a Isomeric Ratio	^b IC ₅₀
1		7a	1:1:0:0	^c n/a
		7b	0:0:1:0	50
		7c	0:0:0:1	n/a
2		7d	1:1:1:0.2	69
		7e	0:0:0:1	n/a
3		^d 7f	1:1	34
		^d 7g	1:1	n/a
4		7h	1	3.7
5		7i	1	0.7
		7j	1	27
		7k	1	38
		7l	1	73
6		7m	1	10
		7n	1	27
		7o	1	n/a
		7p	1	n/a

^a The use of (*R,S*)-difluoroAbu in this work results in the inhibitors being obtained as four diastereoisomers differing in stereochemistry at the P3 (indoline) and at P1 residues. The isomer ratio (as determined by ¹H NMR) of the biological testing sample is denoted. ^b IC₅₀ values (μM) are determined after 30 min preincubation and are the mean of 2 independent measurements. Standard deviations are within 35% of the mean. ^c n/a = not active (below 50% inhibition at 100 μM). ^d 1:1 mixture of diastereoisomers at the P1 difluoroAbu residue. **7f** and **7g** differ in stereochemistry at the P3 (indoline) residue.

surface. In this orientation the aromatic ring of the indoline lies somewhat askew to the plane of the enzyme with the C4 and C5 atoms exposed to solvent.

The C6 position of the indoline was envisioned as a site that could project substituents toward the enzyme surface in search of interactions such as those which enhance potency in the context of the N-terminal capping group of tripeptide inhibitors. The design retained the α-ketoacid functionality that proved an exceptional serine trap in the peptide series.¹⁸

The activities¹⁹ of indolines **7a–p** against the NS3 protease domain in the presence of a synthetic NS4A cofactor peptide are summarized in Table 1. In entry 5, relevant biological activity is shown by a single diastereoisomer of the inhibitor, with **7j–l** showing around a 40-fold or more loss in affinity with respect to the active component **7i**. The stereochemistry of **7i** was unambiguously assigned from the crystal structure of an NS3/NS4A/**7i** complex (Figure 2).²⁰ Both the difluoroAbu side chain at P1 and the thienylmethyl substituent at P3 lie on the opposite face of the inhibitor with

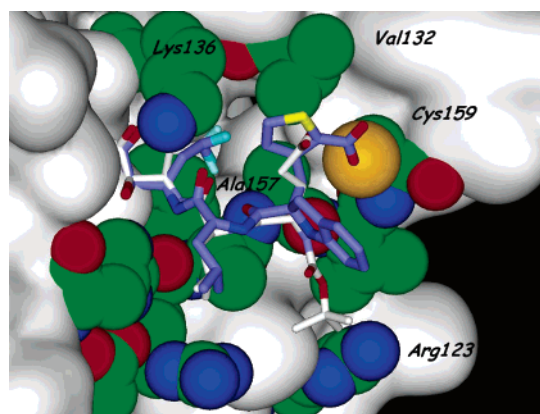


Figure 2. The crystal²⁰ structure of **7i** (purple) with the NS3/NS4 complex at 2.3 Å resolution. Superimposed is the crystal structure with the tripeptide inhibitor Boc-Glu-Leu-difluoro-Abu-CO₂H (white).¹⁷ Both inhibitors are shown in stick representation with heteroatoms colored as follows: F = cyan; O = red; S = yellow; N = blue. Atoms on the enzyme are shown in space-filled representation, with C = green; N = blue; O = red; S = yellow; solvent accessible protein surface = white.

respect to the side chain of the P2 (*L*)-leucine residue. The *S* absolute stereochemistry at each of the three chiral centers in **7i** is in line with the requirement for (*L*)-amino acids at P1–P3 in the peptide series and suggests that, for the C2-alkylated indolines in entries 2–3, the active component in the diastereoisomer mixtures (**7d** and **7f**) will likely also have *S* configuration at P1 and the same relative stereochemistry at P3, (i.e. *S* for **7d** and *R* for **7f**). Analogous stereochemistry is proposed for **7h** (entry 4), the only ketoacid isomer obtained in sufficient purity for biological testing from **6d**. In light of the good potency shown by **7h** and the strong drop in activity seen for **7i** with respect to **7j–l**, we speculate that **7h** likely has the *all-S* stereochemistry that led to activity in entry 5. Surprisingly, in entry 6 two active isomers for the C2 unsubstituted indoline were found. In view of the usual preference for *S* absolute configuration at P1, we presume that both *R* and *S* stereochemistry is tolerated at P3.

The C2 unsubstituted indoline **7b** showed weak inhibition of the NS3 protease. Surprisingly, only modest gains in potency were achieved by introducing alkyl substituents at the C2 position of the indoline. Both α-branched alkyl chains (entry 2) and simple β-branched alkyl groups (entry 3) gained only around 2-fold in potency: the IC₅₀'s for the postulated single active isomers for both **7d** and **7f** would be in the 15–20 μM range. More significant gains in potency were achieved by introduction of benzylic substituents at C2, and we were attracted to the thiophene ring as a means of engaging Cys159, a prominent residue in a convex lipophilic patch that constitutes the S3 site of the enzyme. Donor acceptor interactions²¹ with the electron deficient sulfur atom of the thiophene itself, or substituents such as halogen atoms, were envisaged. Chlorothiophene **7h** gained an order of magnitude in activity over the unsubstituted indoline and illustrates that low micromolar potency can be attained with no acidic functionality in the inhibitor other than the keto-acid moiety at the C-terminus. The combination of thienyl and carboxylic acid functionality at P3 (**7i**) did bring a further 5-fold gain in activity, however, and resulted in the most potent inhibitors in this study. Improvements

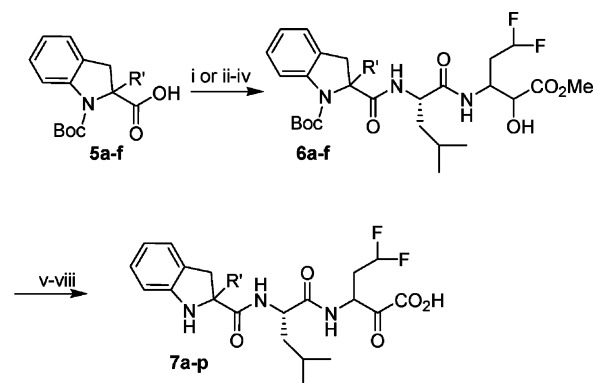
in activity could also be achieved through introduction of functionality on the aromatic ring of the indoline, the ether **7m** showing a 5-fold gain over its unsubstituted counterpart **7b**.

The enzyme bound crystal structure of a representative molecule (**7i**, Figure 2) from the current series of peptidomimetics retains many of the features observed previously for the tripeptide keto-acid inhibitors¹⁷ from which they were designed. The inhibitor combines with the E2 strand of the enzyme to form an antiparallel β -sheet, with the expected covalent bond between the keto group and the active site serine having been formed through nucleophilic attack on the opposite face to that usually observed for serine proteases. As a consequence the oxyanion hole is occupied by the P1 carboxylic acid rather than the hemiketal oxygen atom (which is instead stabilized through hydrogen bonding with the catalytic histidine residue). The difluoromethyl group binds snugly in the S1 specificity pocket and is in contact with the P3 substituent through contact of a fluorine atom with the thiophene C4 position (3.34 Å). This interaction, which may be considered a hydrophobic collapse that serves to preorganize the inhibitor for binding in its extended conformation, results in efficient filling of the lipophilic enzyme surface at S1–S3. The indoline ring engages in the expected P3/S3 hydrogen bonding that was inherent in our design, and presents the thienylmethyl group in an orientation suitable for interaction with Val132. There is an additional close P3/S3 contact between the thiophene carboxylic acid with Cys159. The orientation of the Lys136 side chain, a residue that plays a crucial role in stabilizing the transition states of peptide ketoacids, is also worthy of note. Here we find this chain in hydrogen bonding distance to both the P1 carboxylic acid (2.94 Å) and the P2 carbonyl group (3.15 Å), an orientation which places the aliphatic part of the side chain above the difluoroAbu P1 residue and encloses the S1 specificity pocket.

The importance of inhibitor–enzyme interactions at the S3 site of the HCV NS3 protease is now well appreciated. The S3 region is an open, lipophilic patch that stretches to S1 and in which Cys159 and Val132 are prominent. Optimal interaction at S3 in the peptide series¹⁶ was achieved with P3 amino acids that contain an α -branched aliphatic side chain such as cycloalkyl glycine, an approach which failed in the current setting. The restricted conformational space available to an α -branched substituent when attached to the indoline (**7d**) apparently precludes a favorable interaction at the S3 site. In contrast, the potency of inhibitors such as **7h** over the unsubstituted indoline **7a** is likely due in large part to an increase in the lipophilic contact area at S3. We have shown in previous work that the S3 region is characterized by a positive electrostatic potential created by a number of basic residues, including Lys136.¹⁰ Amphiphilic substituents at the indoline C2 position were therefore attractive, adding an apparently complementary electrostatic component to the lipophilic contact surface of the inhibitor. This substitution indeed proved effective, leading to sub-micromolar inhibitors of the enzyme (**7i**).

Compounds **7h** and **7i**, which may be considered capped dipeptides, show similar inhibition of NS3 as their related tripeptide counterparts containing, respec-

Scheme 1. Synthesis of Indoline Based Peptidomimetic Inhibitors of HCV NS3 Protease^a



^a Reagents: (i) HATU, DIEA, H-Leu-(±)-difluoroAbu-CH(OH)-CO₂Me; (ii) HATU, DIEA, H-Leu-OBn; (iii) Pd/C (10%), H₂ (1 atm); (iv) EDC, HOBT, DIEA, (±)-H-difluoroAbu-CH(OH)CO₂Me; (v) Dess–Martin periodinane, *tert*-BuOH; (vi) TFA, CH₂Cl₂, H₂O; (vii) NaOH, MeOH; (viii) reversed-phase HPLC.

tively, aliphatic (Cbz-Ile-Leu-difluoroAbu-COCO₂H, IC₅₀ = 1.7 μ M) or acidic (Boc-Glu-Leu-difluoroAbu-COCO₂H, IC₅₀ = 0.33 μ M) functionality at P3. The achievement of this level of potency through optimization of the P3/S3 interaction alone is notable given the significant contribution to binding that derives from the N-terminal capping group of the P3 amino acid in a peptide sequence.¹⁶ Figure 2 illustrates that the indoline ring does not superimpose well with the peptide capping group, and while the introduction of oxyacetic acid functionality on the aromatic ring of the indoline (**7m**) did bring a modest improvement of activity (potentially through interaction with Arg123 in the S6 region of the enzyme), there likely remains significant scope for further optimization. In contrast to our expectation, more direct access to the protein surface may in fact be achievable through substitution of the indoline C7 position. Optimization of both the aromatic ring of the indoline and its C2 substituent in tandem may also bring further gains in activity, and the combination of this work with P1/P2 modifications that lead to non-covalent inhibitor series²² offers an attractive strategy toward the development of more drug-like inhibitors of the HCV NS3 protease.

Indoline based inhibitors of the NS3 protease (Table 1) were accessed using the synthetic strategy illustrated in Scheme 1. The required 2-substituted indoline-2-carboxylic acids **5b–e** were obtained from the corresponding esters **3b–e** available via alkylation of **2a** or **2b** using KHMDS as base (Table 2).

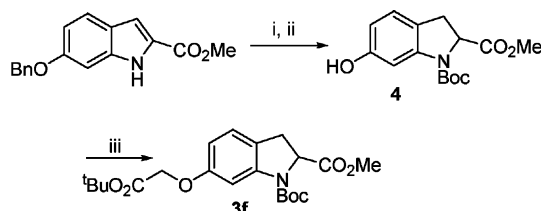
The ester of the 6-substituted indoline **3i** was prepared from methyl 6-benzyloxyindole-2-carboxylate as outlined in Scheme 2. Thus, Boc protection of the indole nitrogen was followed by concomitant hydrogenolysis of the benzyl ether and reduction of the indole ring to afford **4**, which was alkylated with *tert*-butyl bromoacetate.

The acids **5a–f** were elaborated to the keto-acid precursors **6** either by direct HATU mediated peptide coupling with H-Leu-(±)-difluoroAbu-CH(OH)CO₂Me¹⁶ or through stepwise elongation using the corresponding single amino acids. The latter route was used for **5b**, allowing hydrogenation of the cyclohexene ring in tandem with hydrogenolytic cleavage of the leucine

Table 2. Preparation of 2-Alkylindoline-2-carboxylic Acids

Entry	Substrate	R ₁ -X	Product	^a Yield
1	2a		3b	70
2	2a		3c	52
3	2a		3d	80
4	2b		3e	85

^a Unoptimized percentage yield of isolated material.

Scheme 2. Synthesis of 6-Substituted-indoline-2-carboxylic Ester 3f^a

^a Reagents: (i) Boc₂O, DMAP, MeCN; (ii) Pd/C (10%), H_{2(g)} (60 psi); (iii) Cs₂CO₃, *tert*-BuO₂CCH₂Br.

benzyl ester. The diastereomeric mixture of alcohols **6a–f** was oxidized under mild conditions using Dess–Martin periodinane, the reaction being followed by electrospray mass spectrometry to ensure complete reaction of each diastereoisomer. After removal of the protecting groups with trifluoroacetic acid followed by methanolic NaOH, the isomeric mixture of indolines **7** was purified by reversed-phase high performance liquid chromatography. In all but one case (see above) the expected four diastereoisomers of the products were obtained, though separation to isomeric purity was not always effected.

Using the enzyme bound crystal structures of capped tripeptide keto–acid inhibitors of the HCV NS3 protease as a starting point, we have designed 2-alkylindoline replacements for the P3 amino acid and N-terminal capping group of a tripeptide inhibitor sequence. The similarity of both the potency and binding mode of molecules from this series to those of the peptides from which they were designed validates the indoline scaffold as a novel peptidomimetic for use at the N-terminus of a peptide sequence.

Acknowledgment. We thank Maurizio Sollazzo for scientific discussion; Mirko Brunetti, Gabriella Biasiol, Sergio Serafini, and Mauro Cerretani for determination of biological activities; Silvia Pesci for NMR spectroscopy; and Fabio Talamo for HRMS data. This work was supported in part by a grant from the MIUR.

Supporting Information Available: Synthetic procedures and spectral data for compounds **3b–f** and **7a–p**. Experimental protocols for the collection and refinement of X-ray structure data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Beaulieu, P. L.; Llinas-Brunet, M. Therapies for Hepatitis C Infection: Targeting the Non-structural Proteins of HCV. *Curr. Med. Chem.: Anti-Infect. Agents* **2002**, *1*, 163–176.
- Monto, A.; Wright, T. L. The Epidemiology and Prevention of Hepatocellular Carcinoma. *Semin. Oncol.* **2001**, *28*, 441.
- Cohen, J. Hepatitis C. *Science* **1999**, *285*, 26.
- Dymock, B. W. Emerging Therapies for Hepatitis C Virus Infection. *Emerging Drugs* **2001**, *6*, 13.
- Steinkuhler, C.; Koch, U.; Narjes, F.; Matassa, V. G. Hepatitis C Virus Serine Protease Inhibitors: Current Progress and Future Challenges. *Curr. Med. Chem.* **2001**, *8*, 919.
- Bianchi, E.; Urbani, A.; Biasiol, G.; Brunetti, M.; Pessi, A.; De Francesco, R.; Steinkuhler, C. Complex Formation Between the Hepatitis C Virus Serine Protease and a Synthetic NS4A Cofactor Peptide. *Biochemistry* **1997**, *36*, 7890.
- Kim, J. L.; Morgenstern, K. A.; Lin, C.; Fox, T.; Dwyer, M. D.; Landro, J. A.; Chambers, S. P.; Markland, W.; Lepre, C. Crystal Structure of the Hepatitis C Virus NS3 Protease Domain Complexed with a Synthetic NS4A Cofactor Peptide. *Cell* **1996**, *87*, 343.
- Hinrichsen, H.; Benhamou, Y.; Wedemeyer, H.; Reiser, M.; Sentjens, R. E.; Calleja, J. L.; Forns, X.; Erhardt, A.; Crönlein, J.; Chaves, R.; Yong, C.-L.; Nehmiz, G.; Steinmann, G. G. Short-term Antiviral Efficacy of BILN 2061, a HCV Serine Protease Inhibitor, in Hepatitis C Genotype 1 Patients. *Gastroenterology* **2004**, *127*, 1347.
- Yao, N.; Reichert, P.; Taremi, S. S.; Prosser, W. W.; Weber, P. C. Molecular Views of Viral Polyprotein Processing Revealed by the Crystal Structure of the Hepatitis C Virus Bifunctional Protease-Helicase. *Structure* **1999**, *7*, 1353.
- Koch, U.; Biasiol, G.; Brunetti, M.; Fattori, D.; Pallaoro, M.; Steinkuhler, C. Role of Charged Residues in the Catalytic Mechanism of Hepatitis C Virus NS3 Protease: Electrostatic Precollision Guidance and Transition-State Stabilization. *Biochemistry* **2001**, *40*, 631.
- Kakiuchi, N.; Komoda, Y.; Hijikata, M.; Shimotohno, K. Cleavage Activity of Hepatitis C Virus Serine Proteinase. *J. Biochem.* **1997**, *122*, 749.
- Steinkuhler, C.; Urbani, A.; Tomei, L.; Biasiol, G.; Sardana, M.; Bianchi, E.; Pessi, A.; De Francesco, R. Activity of Purified Hepatitis C Virus Protease NS3 on Peptide Substrates. *J. Virol.* **1996**, *70*, 6694.
- Steinkuhler, 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.
- Ingallinella, P.; Altamura, S.; Bianchi, E.; Taliani, M.; Ingenito, R.; Cortese, R.; De Francesco, R.; Steinkuehler, C.; Pessi, A. Potent Peptide Inhibitors of Human Hepatitis C Virus NS3 Protease Are Obtained by Optimizing the Cleavage Products. *Biochemistry* **1998**, *37*, 8906.
- Narjes, F.; Koehler, K. F.; Koch, U.; Gerlach, B.; Colarusso, S.; Steinkuhler, C.; Brunetti, M.; Altamura, S.; De Francesco, R.; Matassa, V. G. A Designed P1 Cysteine Mimetic for Covalent and Non-covalent Inhibitors of HCV NS3 Protease. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 701.
- Colarusso, S.; Gerlach, B.; Koch, U.; Muraglia, E.; Conte, I.; Stansfield, I.; Matassa, V. G.; Narjes, F. Evolution, Synthesis and SAR of Tripeptide α -Ketoacid Inhibitors of the Hepatitis C Virus NS3/NS4A Serine Protease. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 705.
- Di Marco, S.; Rizzi, M.; Volpari, C.; Walsh, M. A.; Narjes, F.; Colarusso, S.; De Francesco, R.; Matassa, V.; Sollazzo, M. Inhibition of the Hepatitis C Virus NS3/4A Protease: The Crystal Structures of Two Protease-Inhibitor Complexes. *J. Biol. Chem.* **2000**, *275*, 7152.
- Narjes, F.; Brunetti, M.; Colarusso, S.; Gerlach, B.; Koch, U.; Biasiol, G.; Fattori, D.; De Francesco, R.; Matassa, V. G.; Steinkuhler, C. α -Ketoacids Are Potent Slow Binding Inhibitors of the Hepatitis C Virus NS3 Protease. *Biochemistry* **2000**, *39*, 1849.
- Urbani, A.; Bianchi, E.; Narjes, F.; Tramontano, A.; De Francesco, R.; Steinkuhler, C.; Pessi, A. Substrate Specificity of the Hepatitis C Virus Serine Protease NS3. *J. Biol. Chem.* **1997**, *272*, 9204.
- The coordinates have been deposited in the PDB. Coordinate ID code: 1w3c. Structure factor code: r1w3csf.
- Bent, H. A. Structural Chemistry of Donor Acceptor Interactions. *Chem. Rev.* **1968**, *68*, 525.
- Colarusso, S.; Koch, U.; Gerlach, B.; Steinkuehler, C.; De Francesco, R.; Altamura, S.; Matassa, V. G.; Narjes, F. Phenethyl Amides as Novel Noncovalent Inhibitors of Hepatitis C Virus NS3/4A Protease: Discovery, Initial SAR, and Molecular Modeling. *J. Med. Chem.* **2003**, *46*, 345.

JM049435D