Virology 405 (2010) 424–438

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Structural characterization of the Hepatitis C Virus NS3 protease from genotype 3a: The basis of the genotype 1b vs. 3a inhibitor potency shift

Mariana Gallo [a,1], Matthew James Bottomley [b,1], Matteo Pennestri [a], Tommaso Eliseo [a], Maurizio Paci [a], Uwe Koch [b], Renzo Bazzo [a,b], Vincenzo Summa [b], Andrea Carfì [b], Daniel O. Cicero [a,*]

[a] Department of Chemical Science and Technology University of Rome "Tor Vergata", Italy
[b] IRBM Istituto di Ricerche di Biologia Molecolare P. Angeletti, MRL-Rome, Pomezia Rome, Italy

ARTICLE INFO

Article history:
Received 2 April 2010
Returned to author for revision 30 April 2010
Accepted 26 May 2010
Available online 13 July 2010

Keywords:
HCV
NS3 protease
3a genotype
NMR
Macrocyclic inhibitors

ABSTRACT

The first structural characterization of the genotype 3a Hepatitis C Virus NS3 protease is reported, providing insight into the differential susceptibility of 1b and 3a proteases to certain inhibitors. Interaction of the 3a NS3 protease with a P2–P4 macrocyclic and a linear phenethylamide inhibitor was investigated. In addition, the effect of the NS4A cofactor binding on the conformation of the protease was analyzed. Complexation of NS3 with the phenethylamide inhibitor significantly stabilizes the protease but binding does not involve residues 168 and 123, two key amino acids underlying the different inhibition of genotype 1b vs. 3a proteases by P2–P4 macrocycles. Therefore, we studied the dynamic behavior of these two residues in the phenethylamide complex, serving as a model of the situation in the apo 3a protein, in order to explore the structural basis of the inhibition potency shift between the proteases of the genotypes 1b and 3a.

Introduction

Hepatitis C Virus (HCV) is a pathogen with 3% average infection prevalence, leading to ~170 million people infected worldwide (Wasley and Alter, 2000). A reduced transmission and improved therapeutic strategies for treating infected patients followed the discovery of the virus in 1989. However, current therapy is inadequate and development of resistance is foreseen as a major obstacle for emerging medications (De Francesco and Migliaccio, 2005). HCV displays an extensive genome heterogeneity, and to date six genotypes (1–6) and more than 50 subtypes (1a, 1b, 1c, and so forth) have been described (Bukh et al., 1995). Although genotype 1 accounts for ~70% of the infections in industrialized countries, genotypes 2 and 3 also have wide distributions, and in particular subtype 3a is common among intravenous drug users in the United States and Europe (Nolte, 2001; Pawlotsky, 2003).

The highest positive response rates to antiviral therapy for the treatment of chronic Hepatitis C have been achieved using the combination of peginterferon and ribavirin. The major determinant of therapeutic outcome is the HCV genotype (Di Bisceglie and Hoofnagle, 2002): patients infected with genotype 1 achieve response rates of 40% to 45%, compared with rates approaching 80% for those infected with genotypes 2 or 3. This fact and the high prevalence of genotype 1 have prompted many pharmaceutical companies to concentrate their efforts in the development of new antiviral agents using viral protein targets of genotype 1.

The HCV genome encodes a single polyprotein of about 3000 amino acid residues containing four structural proteins (Core–E1–E2–p7) and six non-structural proteins (NS2–NS3–NS4A–NS4B–NS5A–NS5B) (Reed and Rice, 2000). The NS3 protein shows both protease and helicase activities. The N-terminal third of NS3 (residues 1–180) constitutes a chymotrypsin-like serine protease domain (NS3p) which, together with the cofactor NS4A, is responsible for the cleavage of the non-structural proteins NS3–NS5B (Reed and Rice, 2000). In addition, NS3 is used by the virus to antagonize the host cell innate immune response by cleaving CARDIF and TRIF, two critical components by which the cells sense the invasion by viral pathogens and trigger the induction of the antiviral state (Gale and Froy, 2005). As a result of its central role in the viral life cycle, NS3p has been chosen as one of the major targets for the development of antiviral agents (De Francesco and Carfi, 2007).

In the past several years, significant progress has been made in developing inhibitors of the HCV NS3/NS4A protease. An important early observation was that the N-terminal cleavage products of HCV NS3p are inhibitors of the enzyme itself (Steinkühler et al., 1998). Three inhibitors, which have been designed with such lead

Abbreviations: HCV, Hepatitis C Virus; NS3p, non-structural protein 3 protease domain; NS3p1b, NS3p3a, and M7, NS3 protease domain of genotypes 1b, 3a and of genotype 3a containing seven mutations respectively; 4Ap, NS4A peptide from genotype 1b; HSQC, heteronuclear single quantum correlation; ASA, accessible surface area; NOE, nuclear Overhauser effect.

* Corresponding author. Department of Chemical Science and Technology, Università di Roma "Tor Vergata" Via della Ricerca Scientifica 1, 00133 Rome, Italy. Fax: +39 06 72594328.
E-mail address: cicero@scienze.uniroma2.it (D.O. Cicero).
[1] These authors contributed equally to this work.

0042-6822/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2010.05.035

compounds as starting points, have shown proof-of-concept in phase I clinical trials: BILN2061, VX-950 and SCH 503034 (De Francesco and Carfi, 2007), confirming that inhibition of NS3p is a viable strategy for HCV therapy.

BILN2061 is a P1–P3 macrocyclic tripeptide that was the first HCV protease inhibitor to enter a clinical trial. It presents an inhibition constant (K_i) for the genotype 1 enzymes in the sub-nanomolar range and an EC_{50} in the genotype 1 HCV replicon system in the low nanomolar range (Lamarre et al., 2003). In phase I clinical trials, administration of BILN2061 to genotype 1 patients resulted in a 2–3 log reduction of HCV RNA in the majority of patients after 2 days of dosing (Hinrichsen et al., 2004). However, the compound was less effective on genotype 2 and 3 patients; only about half of this group achieved more than a 1 log reduction in viral load (Reiser et al., 2003) compared with 100% of the genotype 1 patients at the same dosing level (Hinrichsen et al., 2004).

Mutagenesis studies have shown that the loss in inhibitory potency on genotypes 3 vs. 1 by some compounds is caused by naturally occurring amino acid substitutions in positions 168 or 123 of the NS3p (Tong et al., 2006; Thibeault et al., 2004; Beyer et al., 2001). These are the only two residues, involved in interactions with inhibitors, which present a notably different side chain in the two genotypes. Moreover, mutations at these positions are implicated in the development of resistance to inhibitors in the genotype 1 protease (Trozzi et al., 2003; Lu et al., 2004; Lin et al., 2004, 2005; Dahl et al., 2006). In the genotype 1 NS3 protease, residues 168 and 123 are part of a network of amino acids of alternating charges that starts with the catalytic H57 and spans residues D81, R155, D168, and R123. These last two residues are substituted by the uncharged Q168 and T123 in the 3a genotype, thus interrupting the network of favorable electrostatic interactions. In fact, the lack of a salt-bridge between Q168 and R155 in the genotype 3 protease was invoked (Thibeault et al., 2004) as the main factor contributing to the inhibitory potency loss *in vitro* of BILN2061. However, to date no high resolution structure exists for the NS3 protease of genotype 3, compared to the almost thirty structures available for the 1b protease domain, determined both by X-ray crystallography and NMR spectroscopy. With the purpose of investigating at atomic resolution the structural characteristics of the NS3 protease domain of genotype 3a (NS3p3a) and the basis for its resistance to inhibition, we have used NMR spectroscopy. Similarly to previous investigations using the P1–P3 inhibitor, BILN2061, we found that recently described potent P2–P4 macrocyclic inhibitors (Holloway et al., 2007; Liverton et al., 2008) show a large decrease in potency when tested on NS3p3a with respect to the 1b genotype NS3 protease (NS3p1b) (see Fig. 1, compounds **Ma** and **Mb**, for examples). In contrast, a phenethylamide inhibitor, (Fig. 1, compound **Pa**), shows no potency shift. The latter compound is a close analog of an inhibitor for which we recently characterized the interaction with NS3p1b (Gallo et al., 2009). Here we describe for the first time the overall structural characteristics of a soluble mutated form of NS3p3a. This form of the protease contains seven surface mutations that render soluble the protein and conserves the inhibition profile observed for the wild-type NS3p3a. We studied its interactions with the phenethylamide **Pa** and the macrocyclic **Ma** inhibitors, and with a peptide corresponding to the NS4A cofactor. In addition, we describe the dynamical properties of the side chains of residues 123 and 168 of NS3p3a and relate their structural characteristics to the inhibitory potency loss observed for P2–P4 macrocyclic compounds like **Ma** and **Mb**.

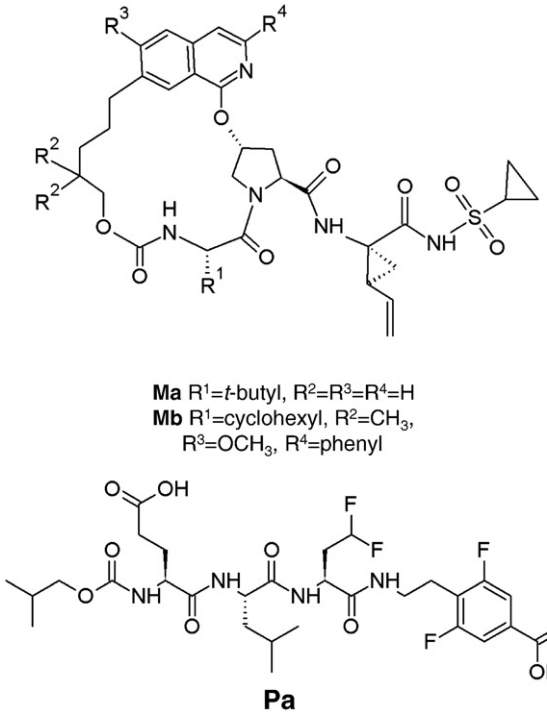

Fig. 1. Chemical structures of protease inhibitors used in this study.

Materials and methods

Expression and purification of recombinant NS3 protease domains

All NS3p expression constructs were PCR-cloned into the pETM-11 vector (EMBL Heidelberg, Germany). Point mutations were introduced using the Quikchange site-directed mutagenesis kit (Stratagene). Protein production was performed in *E. coli* BL21 (DE3) cells grown in LB medium supplemented with 0.05 mM $ZnCl_2$ and induced by 0.4 mM IPTG when the optical density was ~0.6 ($\lambda = 600$ nm). After incubation with shaking at 21 °C overnight, cells were harvested by centrifugation and resuspended in lysis buffer containing 40 mM Tris pH 7.5, 200 mM NaCl, 0.5% (v/v) NP-40, 25% (v/v) glycerol, 6 mM imidazole, 0.02 mM $ZnCl_2$, 1 mM DTT. Cells were lyzed using a microfluidizer and the NS3 proteins were purified from the soluble fraction by Ni-affinity chromatography (Ni-NTA resin, QIAGEN). While attached to the resin, the buffer was exchanged for 40 mM Tris pH 7.5, 200 mM NaCl, 0.1% (w/v) *n*-octyl-β-D-glucopyranoside, 10% (v/v) glycerol, 6 mM imidazole, 0.02 mM $ZnCl_2$, 1 mM DTT, followed by stepwise elution of the protein using the same buffer supplemented with 50–250 mM imidazole. The yield and purity of proteins were compared by reducing SDS-PAGE analysis followed by staining with Coomassie brilliant blue R-250. The NS3p proteins migrate at a position corresponding to their predicted monomeric molecular weight of ~23 kDa. The amount of insoluble protein produced was monitored by using buffer supplemented with 2% SDS to solubilize the pellet obtained by centrifugation of the lyzed cells.

For the most soluble genotype 3a protein, M7, uniformly ^{15}N- or $^{13}C/^{15}N$-labeled proteins were produced in *E. coli* BL21 (DE3) cells grown in M9 minimal medium supplemented with 0.05 mM $ZnCl_2$ and containing $(^{15}NH_4)_2SO_4$ with or without ^{13}C-labeled glucose. For these NMR samples, the lysis buffer was 40 mM Tris pH 7.5, 200 mM NaCl, 0.3% (w/v) CHAPS, 20% (v/v) glycerol, 1 mM DTT. During Ni-resin purification the buffer was exchanged for 40 mM Tris pH 7.5, 200 mM NaCl, 0.1% (w/v) CHAPS, 10% (v/v) glycerol, 1 mM DTT with elution by imidazole as above. Following elution from the Ni-resin the NS3 protein was incubated for 24 h at 4 °C with a solubility-enhanced TEV protease (van den Berg et al., 2006) (plasmid kindly provided by Dr. Helena Berglund) in order to remove the N-terminal 6His tag, with simultaneous dialysis to remove imidazole. The dialysis buffer was 5% (v/v) glycerol, 0.1% (w/v) CHAPS, 20 mM sodium phosphate pH 6.5, 100 mM NaCl and 1 mM DTT. The sample was subsequently re-passed

over a Ni-resin for removal of His-tagged species, including the TEV protease. Finally, the sample was purified by size-exclusion chromatography on a Superdex-75 (16/60) column equilibrated in dialysis buffer.

The purified M7 samples were concentrated to 0.5 mM and exchanged using NAP-5 columns (GE Healthcare) into the NMR buffer containing 4% deuterated glycerol, 0.3% deuterated *n*-octyl β-D-glucopyranoside, 20 mM sodium phosphate buffer pH 6.8, 100 mM NaCl, 0.01% sodium azide and 1 mM DTT at 25 °C.

The NS4A peptide (4Ap) used in NMR experiments corresponds to residues G21 to P39 of the NS4A protein of the 1b genotype, including two lysine residues added at the termini to increase the solubility of the peptide *KK*GSVVIVGRIVLSGKPAIIP*KK* (additional lysines are in italics).

Final samples for NMR experiments contained 0.6–0.8 mM apo-M7 or M7–**Pa** complex, 100 mM NaCl, 20 mM sodium phosphate pH 6.8, 1 mM DTT, 0.3% deuterated *n*-octyl β-D-glucopyranoside, 0.01% sodium azide, in a solution of 90% H_2O/10% D_2O or 100% D_2O.

Surface plasmon resonance (SPR)

All SPR experiments were performed using a Biacore 3000 instrument at 25 °C. Purified M7 protein in 10 mM sodium acetate buffer pH 4.5 was covalently immobilized by amine-coupling on a carboxymethylated dextran CM-5 sensor chip, yielding a surface density of ~1400 Response Units (RU). Experiments were performed at a flow rate of 40 μL/min in running buffer containing 40 mM sodium phosphate pH 7.0, 150 mM NaCl, 0.05% *n*-octyl-β-D-glucopyranoside, 0.1 mM DTT, 0.01 mM $ZnCl_2$. Binding of NS4A-like peptides was examined in multi-point titrations by injecting peptides at a range of concentrations: 0, 0.05, 0.1, 0.5, 1.0, 5, 10, 20 and 40 μM, all dissolved in running buffer. After each step, the sensor chip surface was regenerated with a 1 min injection of 0.6% (v/v) surfactant P-20. Data were analyzed using BIAevaluation software, with subtraction of reference sensorgrams (injection of buffer only) to yield curves representing specific binding. Steady-state analysis was used to obtain thermodynamic dissociation constants (K_D). Due to the hydrophobicity of the wild-type NS4A cofactor peptides, lysine (K) residues were added at both termini of the peptides used in the SPR studies in order to enhance their solubility. Thus, the NS4A peptide 1b (3 K) has the sequence *KKK*GSVVIVGRIVLSGKPAIIP*KKK*; and the NS4A peptide 3a(3 K) has the sequence *KKK*GCVVIVGHIELGGKPALVPDK*KK*, sequence differences between the two peptides are underlined, additional lysines are in italics.

Inhibition assays

Inhibitors **Ma**, **Mb** and **Pa** were chemically synthesized according to published protocols (Liverton et al., 2008; Malancona et al., 2004). The inhibition assays were performed by using the sensitive time-resolved fluorescence (TRF) assay capable of detecting very low NS3/NS4A concentrations developed by Mao et al. (2008). Briefly, NS3 was solubilized in 100 μL of buffer containing 50 mM Hepes pH 7.5, 150 mM NaCl, 15% glycerol, 0.15% Triton X-100, 10 mM DTT, and 0.1% PEG. The solutions containing the protease were preincubated with different concentrations of inhibitors in DMSO for 30 min and for 15 min with 5 μM 4Ap. Then the reactions were initiated by adding 100 nM of the TRF substrate (Ac-Cys(Eu)-Asp-Asp-Met-Glu-Glu-Abu-[COO^-]-X-Ser-Ala-Lys(ε-QSY7)-NH_2, where Eu and QSY7 are the fluorescent and quenching groups, respectively). The hydrolysis was quenched after 1 h at room temperature with 100 μL of 500 mM MES pH 5.5. Product fluorescence was detected with excitation at 340 nm and emission at 615 nm with a 400 μs delay in a Victor V2 fluorophoteometer (Perkin Elmer). The inhibition constants were derived using standard four-parameter fit to the data. Errors in the K_i constants measured are in the low subnanomolar range.

NMR spectroscopy

The NMR experiments were performed at 35 °C on Bruker Avance700, Avance 600, and Avance400 spectrometers equipped with triple resonance probes incorporating shielded z-axis gradient coils. The NMR data were processed on Silicon Graphics workstations using NMRPipe (Delaglio et al., 1995) and analyzed using NMRView software (Johnson, 2004).

The following standard set of triple resonance spectra were acquired in H_2O: HNCO, HNCA, HN(CO)CA, CBCA(CO)NH, HBHA(CO)NH and CBCANH and ^{15}N-edited TOCSY. For the sample dissolved in D_2O the following experiments were performed: (H)CCH-COSY, H(C)CH-COSY, (H)CCH-TOCSY, HACACO (Cicero et al., 2006). For glutamine and asparagine side chain resonance assignment the following experiments optimized for NH_2 groups were performed: HNCO, HNCA and HACACO. Side chain resonance assignment was achieved for almost all the T, Q, and N residues. The ^{15}N-{^{1}H} NOE experiment was performed at 70.94 MHz ^{15}N frequency using standard pulse schemes (Kay et al., 1989) optimized for NH_2 groups in an interleaved manner. The heteronuclear NOE values were determined by the ratio of peak volumes of spectra recorded with and without ^{1}H saturation, employing a net relaxation delay of 4 s for each scan in both experiments. Errors were estimated by evaluating the standard deviation of the NOE, σNOE:

$$\sigma\mathrm{NOE} = \left[{}^1\!/_2\left(\sigma I^2_{\mathrm{sat}}/I^2_{\mathrm{unsat}} + \sigma I^2_{\mathrm{unsat}} I^2_{\mathrm{sat}}/I^4_{\mathrm{unsat}}\right)\right]^{1/2}$$

where σI_{sat} and $\sigma I_{\mathrm{unsat}}$ are the standard deviations of the noise in the spectra. Errors were typically in the range 0.05–0.1.

Measurements of ^{15}N T_2 relaxation experiments were carried out at 35 °C in a 400 MHz spectrometer using standard pulse schemes (Kay et al., 1989) in an interleaved manner. Relaxation delays of 8, 24, 32, 40, 48, and 56 ms were employed. Data were fitted using the Rate Analysis routine of NMRView (Johnson, 2004). The uncertainties of peak intensities were evaluated as the standard deviation of the spectral noise measured in a region free of cross-peaks. Typically, errors were about 6%.

Results

A soluble form of the 3a genotype NS3 protease

Production of NS3p3a (amino acids 1–180) using the procedure previously described for obtaining NS3p1b, even after addition of a C-terminal solubilizing lysine tail (residues ASKKKK) (Steinkühler et al., 1998), yielded mostly protein in the insoluble fraction (Fig. 2A). Efforts to refold the protein from the urea-denatured state were unsuccessful. In view of these negative results, we engineered mutants of NS3p3a to favor its production as a soluble protein for structural studies. Fig. 2B shows the sequence alignment of the NS3 protease domains of genotypes 1 and 3. Residues close to the catalytic site, substrate binding region or involved in the interaction with the NS4A cofactor were left untouched. From the remaining residues we selected a set of seven NS3p3a surface exposed residues (boxed in Fig. 2B) to be mutated by introducing the corresponding residue of the more soluble 1b genotype. The sequence of the resulting mutant protein (hereafter termed M7) is shown at the bottom of Fig. 2B, and with respect to the NS3p3a wild-type presents the seven mutations at the N-terminal domain (V28Q, T47C, G49N, H69G, L72T, A98T, and A102S). These positions are distant from the catalytic site, as shown when mapped onto a model of the structure of M7 in Fig. 2C.

As predicted, the mutations resulted in the production of a greatly enhanced quantity of soluble protein that could be readily purified (see Materials and methods, Fig. 2A). This is particularly relevant for structural studies by NMR, for which a protein must reach

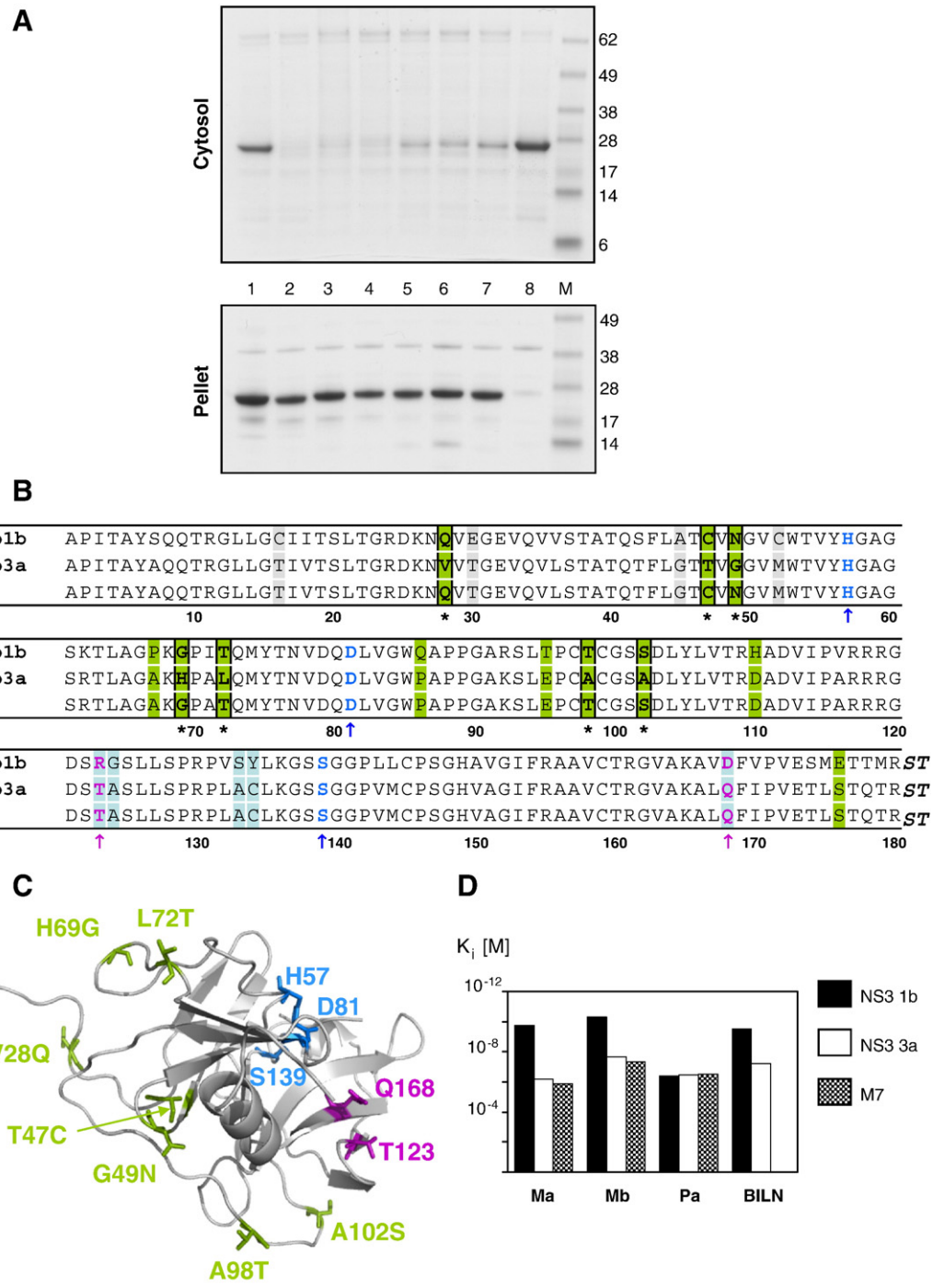

Fig. 2. Engineering a soluble form of the 3a genotype NS3 protease. (A) Direct comparison of the solubility of different NS3p proteins. SDS-PAGE analysis was used to compare the yields of each NS3p protein present in the soluble fraction (partially purified from the cytosol; upper panel) and in the pellet (lower panel). The SDS-PAGE gels were stained with Coomassie brilliant blue R-250 and are shown as grayscale images. The protein band at ~38 kDa observed for the pellet indicates that the expression level of a non-NS3 protein was similar in all cases. Key to lanes: (1) NS3p1b strain J, (2) NS3p3a wild-type, (3) NS3p3a— *ST* (*ST* indicates a solubizing tail added at the C-terminus of NS3p, —ASKKKK), (4) NS3p3a— *ST* V28Q, (5) NS3p3a— *ST* T47C and G49N, (6) NS3p3a— *ST* H69G and L72T, (7) NS3p3a— *ST* A98T and A102S, (8) M7, (9) Molecular Weight markers. (B) Multiple alignment of NS3p sequences: NS3p1b (used in our previous NMR studies), NS3p3a, and M7 mutant used here. Positions that present a notable side chain switch between genotypes 1 and 3 are grouped according to whether they are involved in the interaction with the NS4A cofactor (shaded light gray), if they are close to the catalytic or substrate binding region (shaded light blue), or if they do not belong to the other two categories (shaded green). Seven residues from this group chosen to be mutated in M7 are boxed and indicated with an asterisk at the bottom. The catalytic triad (H57, D81, and S139) is indicated in blue and highlighted by arrows. Residues at positions 123 and 168 responsible for the inhibitor potency shift between both genotypes are in violet and highlighted by arrows. (C) Model of the structure of M7 showing the positions that were mutated from NS3p3a wt to achieve the soluble form of the protease. Residues mutated are depicted in green, the catalytic triad in blue and residues 123 and 168 in violet. (D>) pK_i values for inhibition of NS3/NS4A protease of genotypes 1b, 3a and the mutant M7. Values are reported for inhibitors Ma,Mb and Pa, whose chemical structure is shown in Fig. 1, and BILN2061. Standard deviations in measurements of pK_i are too small to be appreciable in the log-scale graph (but were typically in the low subnanomolar range).

concentrations around 10^{-4} M. The solubilizing effect of the mutations appeared to be cumulative, since alternative 'intermediate' forms of NS3p3a each containing only one or two mutations showed only mildly increased solubility compared to the wild-type NS3p3a but were notably less soluble than the composite M7 protein. After purification, the M7 protein migrated as a single band in SDS-PAGE with the appropriate apparent molecular weight, displayed a monodisperse monomeric elution profile in analytical gel filtration studies and gave the expected molecular weight by electrospray mass spectroscopy (data not shown).

P2–P4 macrocyclic inhibitors **Ma** *and* **Mb** *show a large potency shift between 3a and 1b NS3 proteases*

In order to assess whether the M7 mutant retained the behavior of the wild-type form of the protease with respect to inhibitor potency shift, we tested the two NS3p3a proteins and the 1b genotype protease with a series of inhibitors. Recently, potent macrocyclic inhibitors of the NS3/NS4A protease were described, the design of which was based on the formation of a P2–P4 linker, using an isoquinoline moiety as the P2 group to facilitate the rapid synthesis of analogs (Liverton et al., 2008). Compounds **Ma** and **Mb** depicted in Fig. 1 are examples of these inhibitors, which show a K_i in the low picomolar range on the 1b genotype NS3/NS4A enzyme (Liverton et al., 2008). Both compounds were tested using the wild-type NS3/NS4A enzyme of the 3a genotype and M7.

A K_i decrease of between two and three orders of magnitude with respect to the 1b genotype enzyme was observed for both 3a enzymes in the case of **Ma** (K_i 0.18 nM and 659 nM for the 1b and 3a genotypes, respectively) and **Mb**, (K_i 0.05 nM and 20 nM) (Fig. 2D). This inhibitory potency decrease is comparable to that of the BILN2061 inhibitor, for which a 200-fold change in K_i was observed between 1b and 3a genotypes, when the inhibition assay was performed in the same experimental conditions (K_i 0.32 nM and 65 nM).

We also tested an inhibitor presenting a less bulky P2 side chain and lacking the P2–P4 cycle, like the phenethylamide inhibitor **Pa** (Fig. 1) (Malancona et al., 2004). Compound **Pa** does not show a significant potency difference between 1b and 3a proteases and, as observed for **Ma** and **Mb**, there is no significant difference between the wild-type and mutated 3a enzymes (K_i 0.40 μM, 0.35 μM and 0.31 μM for the 1b and 3a genotypes and M7, respectively) (Fig. 2D).

From these results we conclude that, as observed for wild-type NS3p3a, the M7 protein retains the same decreased susceptibility to inhibition by P2-exploiting compounds when compared to the genotype 1b protease, and that the interaction between M7 and **Pa** does not involve the key residues that cause this shift (123 and 168).

Binary complexes with **Pa** *and* **Ma**

M7 was produced uniformly ^{15}N- or ^{15}N/^{13}C-labeled to characterize its structure in solution by NMR spectroscopy. Fig. 3A shows the ^{1}H–^{15}N HSQC spectrum of M7. The quality of the spectrum is acceptable for NMR studies, but we observed signs of sample degradation at high protein concentrations, probably due to autoproteolysis.

A better quality NMR spectrum was obtained for the complex between M7 and the phenethylamide inhibitor **Pa** (M7–**Pa**) (Fig. 3B). Peaks in the HSQC spectrum appeared more disperse and more homogeneous than in spectrum without the inhibitor. We have recently reported that this type of inhibitor strongly stabilizes the N-terminal domain of NS3p1b, and in particular the active site conformation (Gallo et al., 2009). This fact is clearly reflected by the presence of the two peaks corresponding to H57 and G58, at the expected chemical shifts observed for other NS3p complexes with a stabilized active site (Gallo et al., 2009; Archer et al., 2002; McCoy et al., 2001). Using multinuclear NMR experiments that correlate the chemical shift of backbone nuclei, such as HNCO, HNCA, HNcoCA and CBCAcoNH, we were able to assign 86% of the M7–**Pa** complex ^{1}H, ^{13}C and ^{15}N backbone nuclei. By analyzing the chemical shifts of Hα, Cα, Cβ and C′ nuclei with the program TALOS (Cornilescu et al., 1999) we determined the secondary structure of the M7–**Pa** complex (Fig. 3C). The positions of the different elements of secondary structure correspond very closely with those observed in all the crystal structures for complexes of NS3p1b and with those of the complex between NS3p1b and an analog of **Pa** (Gallo et al., 2009). Regarding the N-domain, the secondary structure results indicate that the introduction of the seven mutations that rendered the protein more soluble did not change the structure of the protein, in line with the fact that the inhibitor potency profiles on the mutant protein are very similar to that on NS3p3a. In addition, the two key residues that are responsible for the loss in inhibitory potency in the 3a genotype, T123 and Q168, are located in the same relative position in the C2 and F2 strands, respectively. Importantly, this indicates that the loss of inhibition observed for compounds with large P2-substituents on the 3a vs. 1b protease is related to different chemical and/or conformational properties at the level of the side chains rather than a structural change in the backbone conformation between the 1b and 3a genotype proteases.

In the case of compound **Ma**, no structural data is available for complexes with NS3p of any genotype, so we decided to study in parallel its interaction with NS3p1b and M7. **Ma** interacts with both proteases under the NMR conditions. However, the **Ma** complexes demonstrate conformational flexibility in solution, and probably also aggregation effects, resulting in low T_2 values that precluded a detailed structural analysis (see Supplementary material, Fig. S1). Nevertheless, the inhibitor binding region was determined for both the 1b and 3a complexes (Fig. 4) by comparing the chemical shift of ^{1}H and ^{15}N nuclei between **Pa** and **Ma** complexes. As expected, regions showing chemical shift changes are located in the differential surface of interaction of these inhibitors. In both genotypes the interaction of compound **Ma** with NS3p occurs in the same region, indicating that the orientation of the inhibitor in the protease complexes is the same for both genotypes. In particular, the models show the different way the R155–D168 region is used by the two inhibitors, with a direct protein–ligand interaction with compound **Ma** but not **Pa**.

Fig. 5A shows selected regions of the ^{1}H–^{15}N HSQC spectra of **Pa** and **Ma** complexes with the two proteases. In particular, a peak corresponding to G58 was observed for the **Ma** complex with both proteases. As mentioned before, the appearance of this cross-peak, which is absent in the spectrum of the apo-protease, is an indication of the stabilization of the active site produced by inhibitor binding. Furthermore, residue 168 presents a different chemical shift in **Pa** and **Ma** complexes with NS3p1b and M7 proteases. This difference strongly suggests that this residue interacts only with **Ma** and not with **Pa**. Interestingly, residue 123 presents almost the same chemical shift in the two NS3p1b complexes, but a different chemical shift between the two M7 complexes.

Interaction with NS4A peptides

Interaction of NS3p with its cofactor was investigated using peptides spanning the central part of the NS4A protein belonging to genotypes 1b and 3a (Lin et al., 1995). We first tested the ability of a 23-mer peptide derived from the 3a genotype NS4A cofactor to interact with M7, but no new peaks were detected (data not shown). In contrast, the interaction of M7 with a peptide cofactor spanning residues G21–P39 of the genotype 1b NS4A protein (termed 4Ap) (Kim et al., 1996), provoked changes in the chemical shift of several amide HN nuclei detected in the HSQC spectrum (Fig. 6). More importantly, two peaks displaying the expected chemical shifts of H57 and G58 appeared in the M7–4Ap complex spectrum indicating a stabilization of the active site by cofactor binding very similar to that

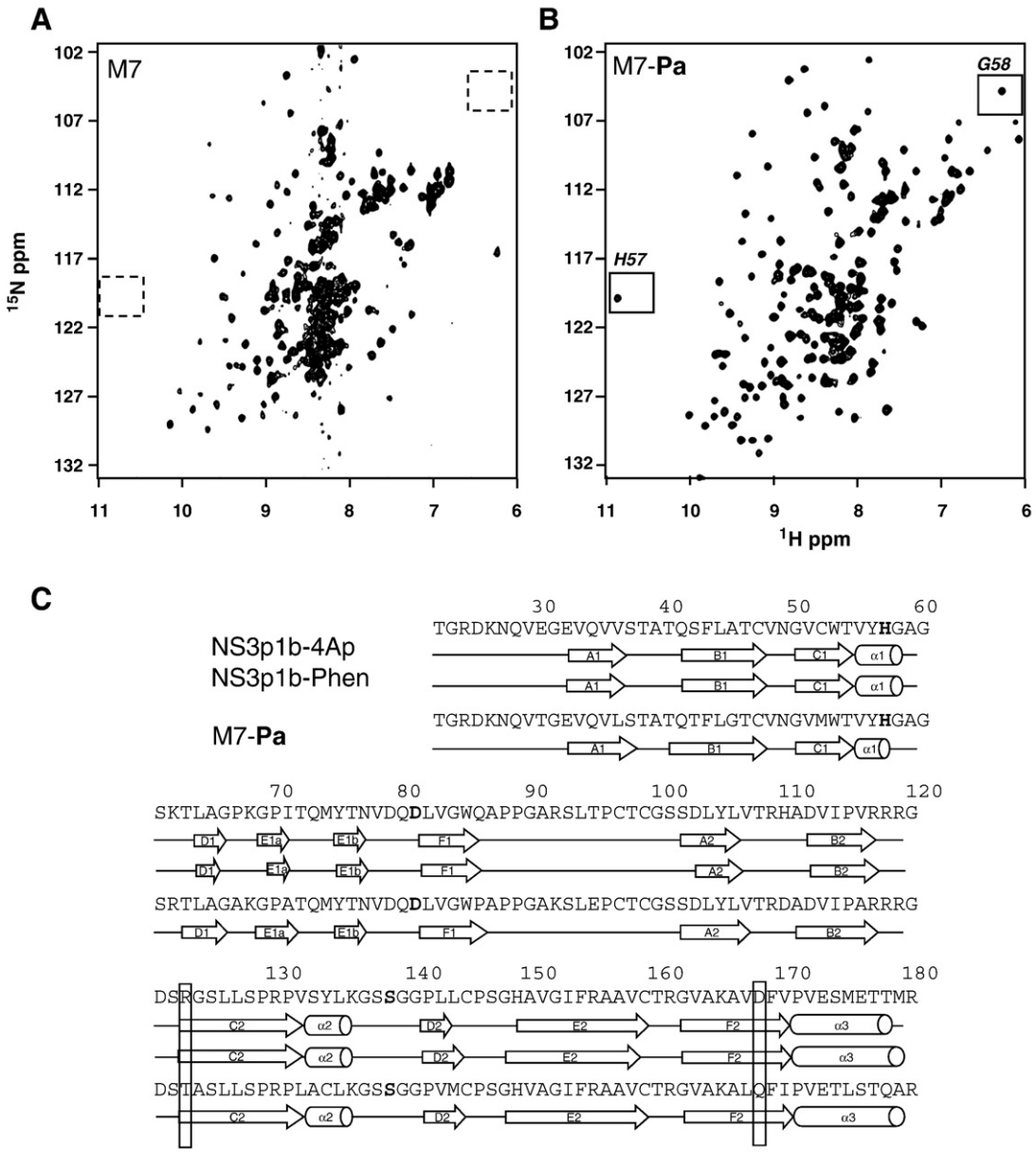

Fig. 3. Interaction of M7 protease with Pa. $^{1}H–^{15}N$ HSQC spectra on ^{15}N-labeled M7 alone (A), and after the addition of 1 equivalent of the phenethylamide inhibitor Pa (B). The presence of diagnostic signals belonging to the active site, H57 and G58, are boxed. (A) was carried out at 25 °C and in the presence of 4% deuterated glycerol, while (B) at 35 °C and in the absence of glycerol. (C) Secondary structure of NS3p complexes: NS3p1b–4Ap complex (PDB entry 1A1R, Kim et al., 1996), NS3p1b complexed with a phenetylamide inhibitor (Phen) analog of Pa (PDB entry 2K1Q, Gallo et al., 2009), and that observed for M7–Pa. Key residues for the inhibitor potency shift between genotypes 1b and 3a (positions 123 and 168) are boxed. NMR signals of residues 1–21 were not observed.

observed in the binary complex with **Pa**. The relative intensities of these two cross-peaks with respect to the average intensity of the N–H cross-peaks increased upon cofactor binding. This fact may indicate a degree of active site stabilization induced by 4Ap, reflecting either the disappearance of slow motions, and/or a reduced solvent exchange rate. Further studies will be conducted in order to clarify this point.

This improved binding to M7 of the peptide from genotype 1b compared to that from genotype 3a was also observed in surface plasmon resonance studies. In brief, using the M7 protein covalently immobilized on a carboxymethylated dextran-coated sensor chip, a steady-state affinity analysis revealed a modestly enhanced binding of the genotype 1b peptide ($K_D = 1.9 \pm 0.3$ μM) compared to the genotype 3a peptide ($K_D = 5.0 \pm 0.8$ μM). Nevertheless, the spectrum of the M7–4Ap complex at high protein concentrations was of poor quality, with heterogeneous line widths and an even faster degradation rate compared with apo-M7, precluding the study of this binary complex at high resolution.

Interaction of M7–**Pa** with 4Ap induced changes in the chemical shift of some peaks (Fig. 7). The $^{1}H–^{15}N$ HSQC spectrum was of an intermediate quality between those of M7 and M7–**Pa**. Using the assignment of M7–**Pa** we followed the chemical shift perturbation caused by the interaction with the 4Ap cofactor. The perturbation is very different in the N- and C-domains, as expected from the fact that, in all known crystals, the cofactor binds to NS3p forming an extra β-sheet in the N-domain. Fig. 7B shows regions of the $^{1}H–^{15}N$ HSQC spectra of the binary and ternary complexes. Most of the residues that show significant perturbations are close to the cofactor in the structures containing 4Ap (Fig. 7A). Conversely, the inhibitor binding site is nearly unperturbed by 4Ap binding, particularly in the S-region

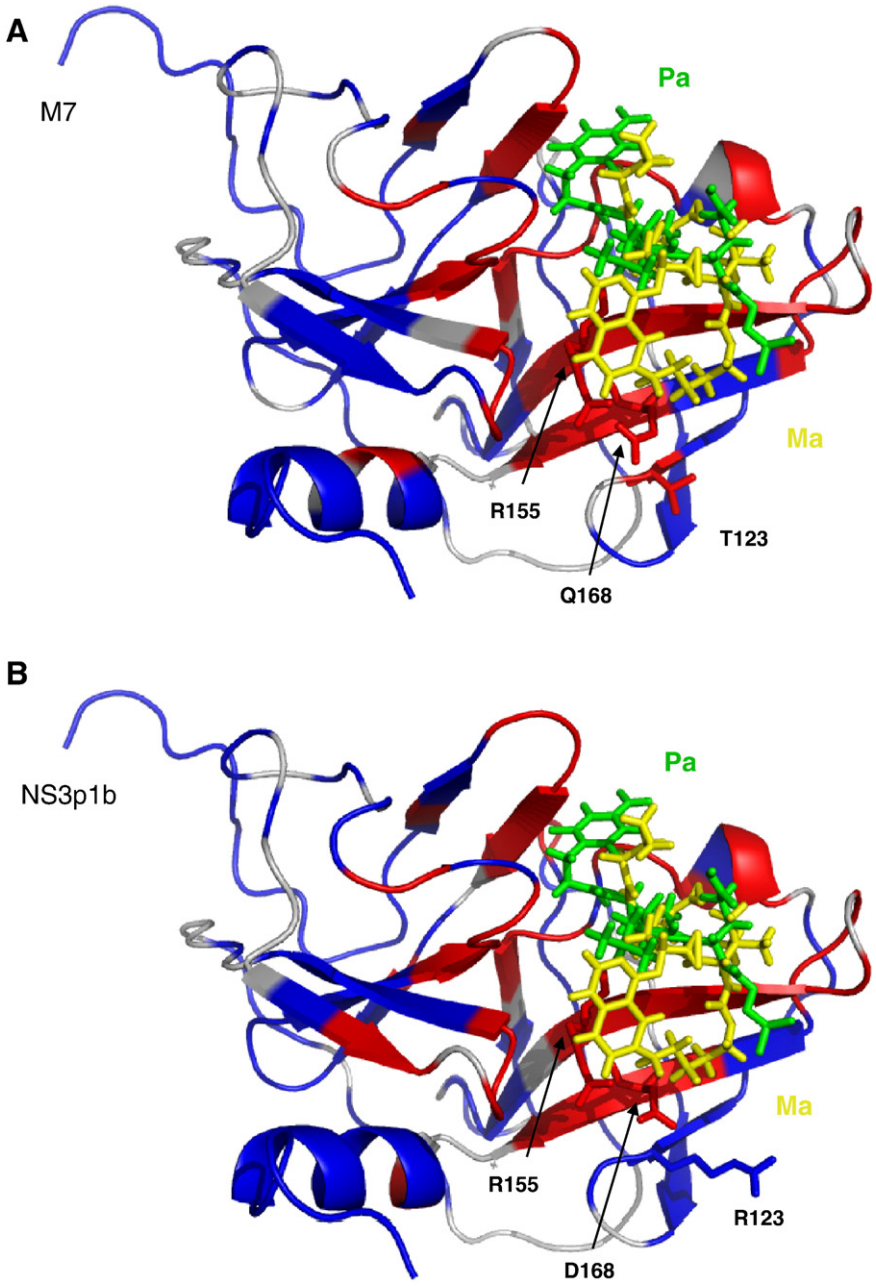

Fig. 4. Interaction of NS3p1b and M7 proteins with Ma. Chemical shift changes produced on M7 (A) and on NS3p1b (B) upon interaction with Ma, compared with the respective complexes with Pa, are mapped on the structure of the protease. Residues whose chemical shift remains invariable in both complexes are depicted in blue; residues for which the chemical shift is different in the Ma complex respect the Pa complex are depicted in red. Pa is designed in green and Ma in yellow. For the models, the structure of the complex of NS3p1b with a close analog of Pa (PDB entry 2K1Q) was used as a basis. Ma was docked in the latter structure preserving the conserved inhibitor orientation that was observed in all the available structures of NS3p1b–inhibitor complexes (Gallo et al., 2009). Residues of the charged network R123–D168–R155 are shown. The models show the different way the two types of inhibitor interact with this region.

and the catalytic site. Residues T123 and Q168 do not show any chemical shift perturbation by 4Ap binding indicating that their structural features do not depend on the presence of the cofactor.

The active site of M7 in the binary and ternary complexes

Residues H57 and G58, considered as markers of the stabilization of the active site, do not show any large change in chemical shift upon 4Ap binding to the M7–**Pa** complex (Fig. 7B). In the case of H57, the particular downfield chemical shift of the amide hydrogen (close to 11 ppm) was rationalized as a consequence of a hydrogen bond with the D81 side chain (McCoy et al., 2001). This situation is apparently not altered by the cofactor, and only a slight upfield shift is noticed. No perturbation in chemical shift was observed for the other two members of the catalytic triad (Fig. 7A), indicating that in the presence or absence of 4Ap the conformation of the active site induced by **Pa** is essentially identical.

A closer inspection of the active site is possible by monitoring the low field signal of the catalytic histidine δ1 proton (Fig. 8A). This peak was not observed for free M7 or for the M7–4Ap complex, indicating that, although the His-Asp dyad might be aligned by the cofactor, the fast exchange with the solvent does not allow its observation. In

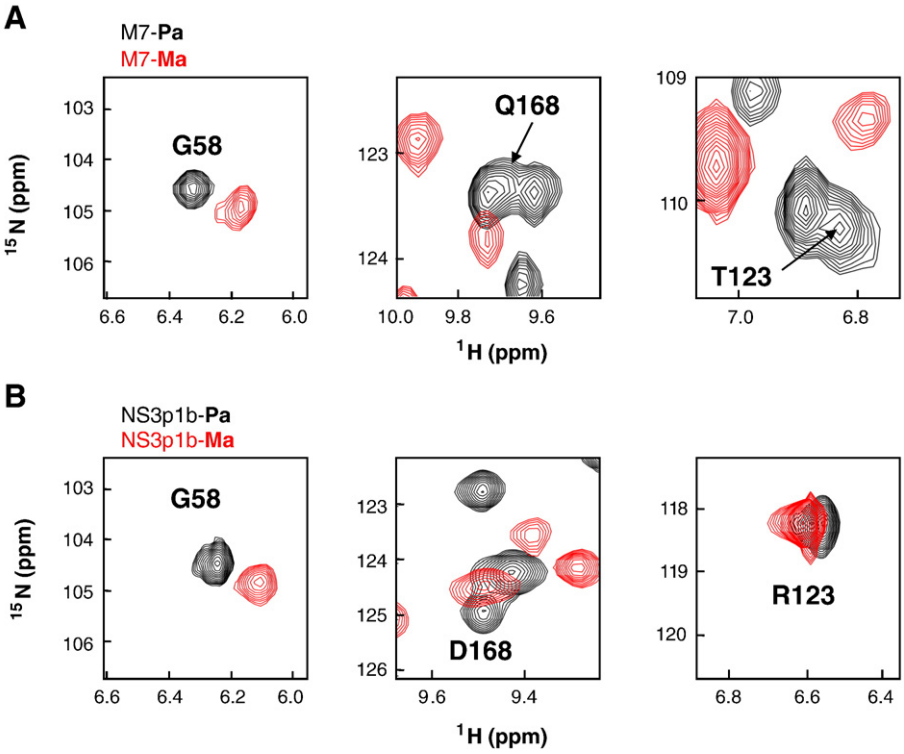

Fig. 5. Active site stabilization and interaction of positions 123 and 168 of NS3p1b and M7 with Ma. (A) Selected regions of the superposition of the ^{1}H–^{15}N HSQC spectra of the M7–Pa (black) and M7–Ma (red) complexes showing cross-peaks corresponding to G58, T123 and Q168 in the M7–Pa complex. (B) The same for NS3p1b–Pa (black) and NS3p1b–Ma (red) complexes.

contrast, the signal is readily observable for the M7–**Pa** complex, indicating that the inhibitor alone is able to align the His-Asp dyad and shield the region from solvent exchange by means of the P2 side chain, as already observed for other NS3p complexes (Gallo et al., 2009; Barbato et al., 2000). The signal remains sharp at temperatures up to 35 °C, thus providing evidence of very slow exchange kinetics with the solvent. Cofactor binding to form the ternary complex provokes a very small downfield shift of the signal. The tautomer of H57 in the absence of 4Ap corresponds unambiguously to the Nδ1 protonated form, as judged from the ^{1}H–^{15}N long-range HSQC spectrum (Pelton et al., 1993) (Fig. 8B). This tautomer, although less populated than the Nε2 protonated form of a solvent exposed histidine, is stabilized by the interaction between H57 and D81. The chemical shift of Nδ1 reflects its participation as a strong hydrogen bond donor, whereas Nε2 is not involved in any hydrogen bond interaction (Bachovchin, 1986). This situation remains almost unaffected upon 4Ap binding.

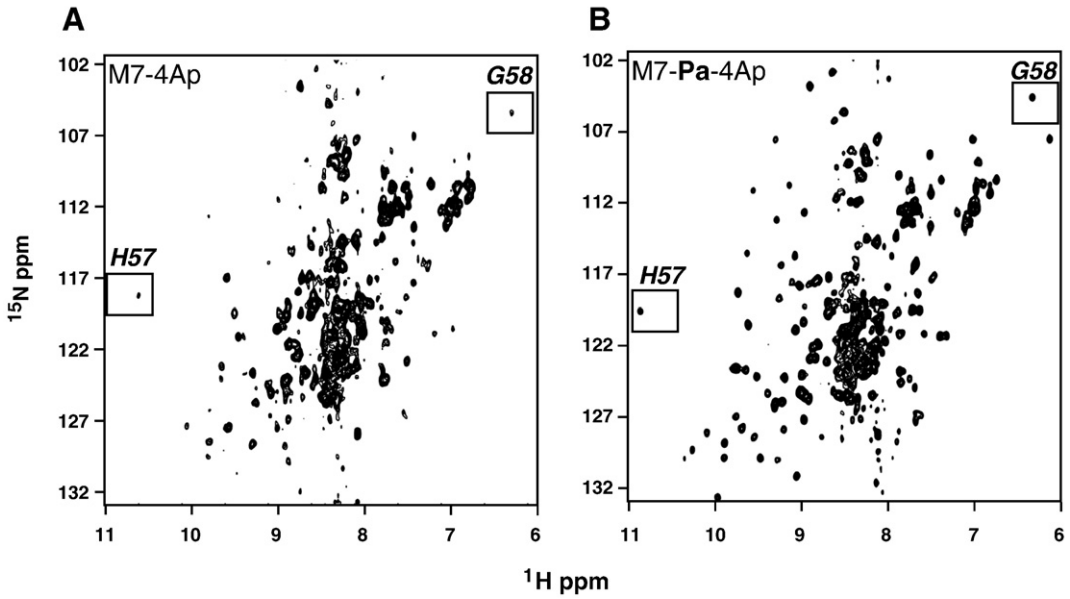

Fig. 6. Interaction of M7 and M7–Pa complex with 4Ap. ^{1}H–^{15}N HSQC spectra on ^{15}N-labeled M7 after addition of 1 equivalent of the 1b genotype 4Ap (A) and after the addition of both 4Ap and Pa (B). The presence of diagnostic signals belonging to the active site, H57 and G58, are boxed. (A) was carried out at 25 °C and in the presence of 4% deuterated glycerol, while (B) at 35 °C and in the absence of glycerol.

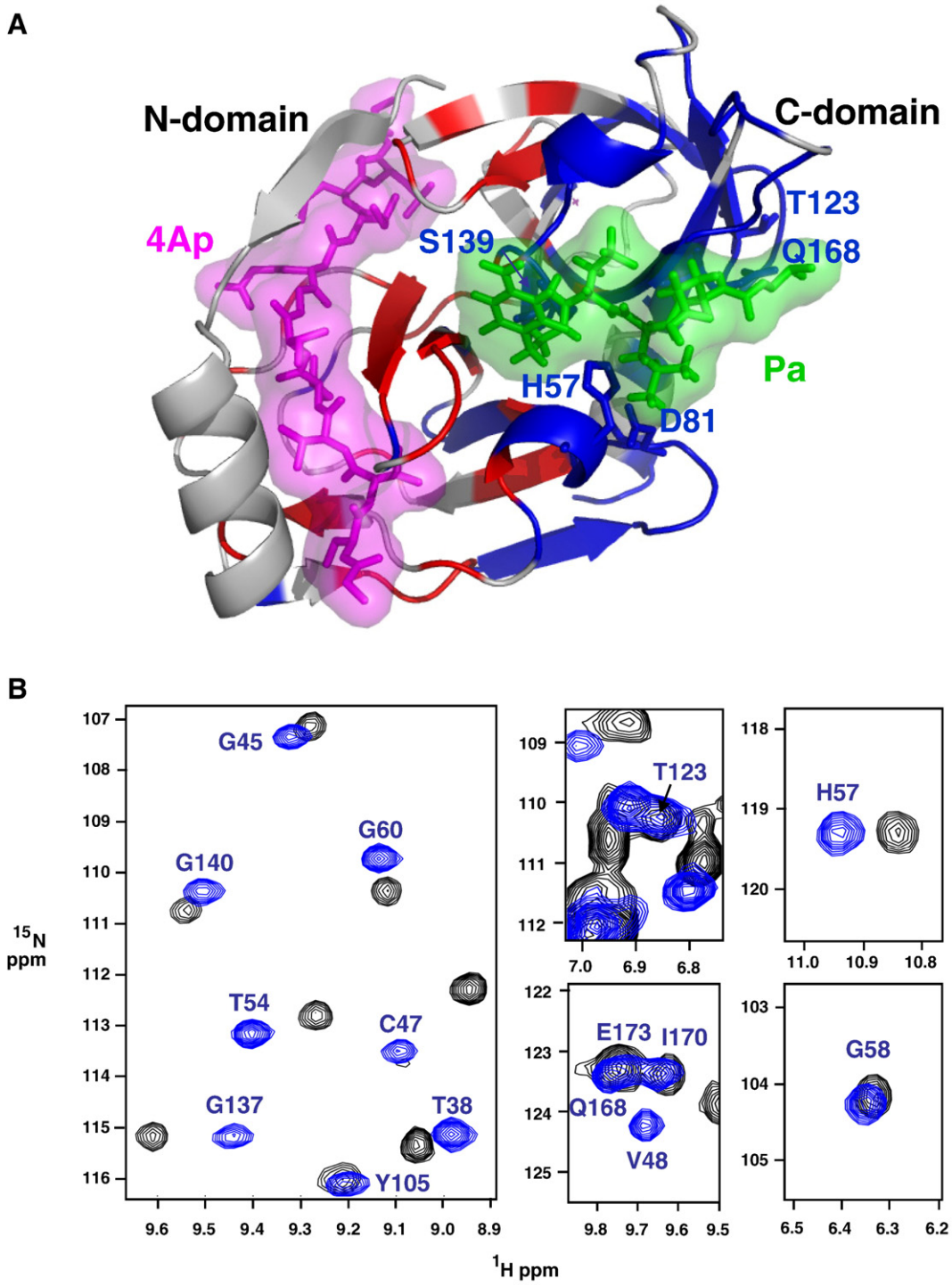

Fig. 7. Chemical shift comparison between M7–Pa and M7–Pa–4Ap. (A) Chemical shift perturbation caused by 4Ap with M7–Pa interaction mapped onto the model for the ternary complex. Residues showing perturbation are colored in red. Unperturbed residues are colored in blue. Residues for which the perturbation could not be established (prolines, unassigned or overlapped residues) are colored in gray. Pa and 4Ap are depicted in green and magenta, respectively. Side chains of residues of the catalytic triad (H57, D81, and S139) and T123–Q168 are evidenced. The model was constructed using the crystal structure of NS3p1b–4Ap (PDB entry 1A1R) and the NMR structure of the complex NS3p1b with an analog of Pa (PDB entry 2K1Q). (B) Selected regions of ^{1}H–^{15}N HSQC spectra of the M7–Pa complex obtained in the absence (blue) or presence (black) of 4Ap. The central and right panels show the shift provoked on T123, Q168, H57 and G58 upon interaction with the cofactor.

Side chain dynamics of key residues in the M7 protease

In view of the fact that **Pa** is able to strongly stabilize NS3p in solution, and since this interaction does not involve residues 123 and 168, responsible for the loss of inhibitor potency on genotype 3a vs. 1b by compounds with a large P2 substituent, we decided to study the conformational properties of these residues in the context of the M7–**Pa** complex. The assumption is that these residues are present in the same conformation in M7 and in M7–**Pa**. The goal of our study was to establish whether there are differences in the dynamics of the side chains of these two residues in the proteases of both genotypes *before* interacting with inhibitors that exploit one or both positions for

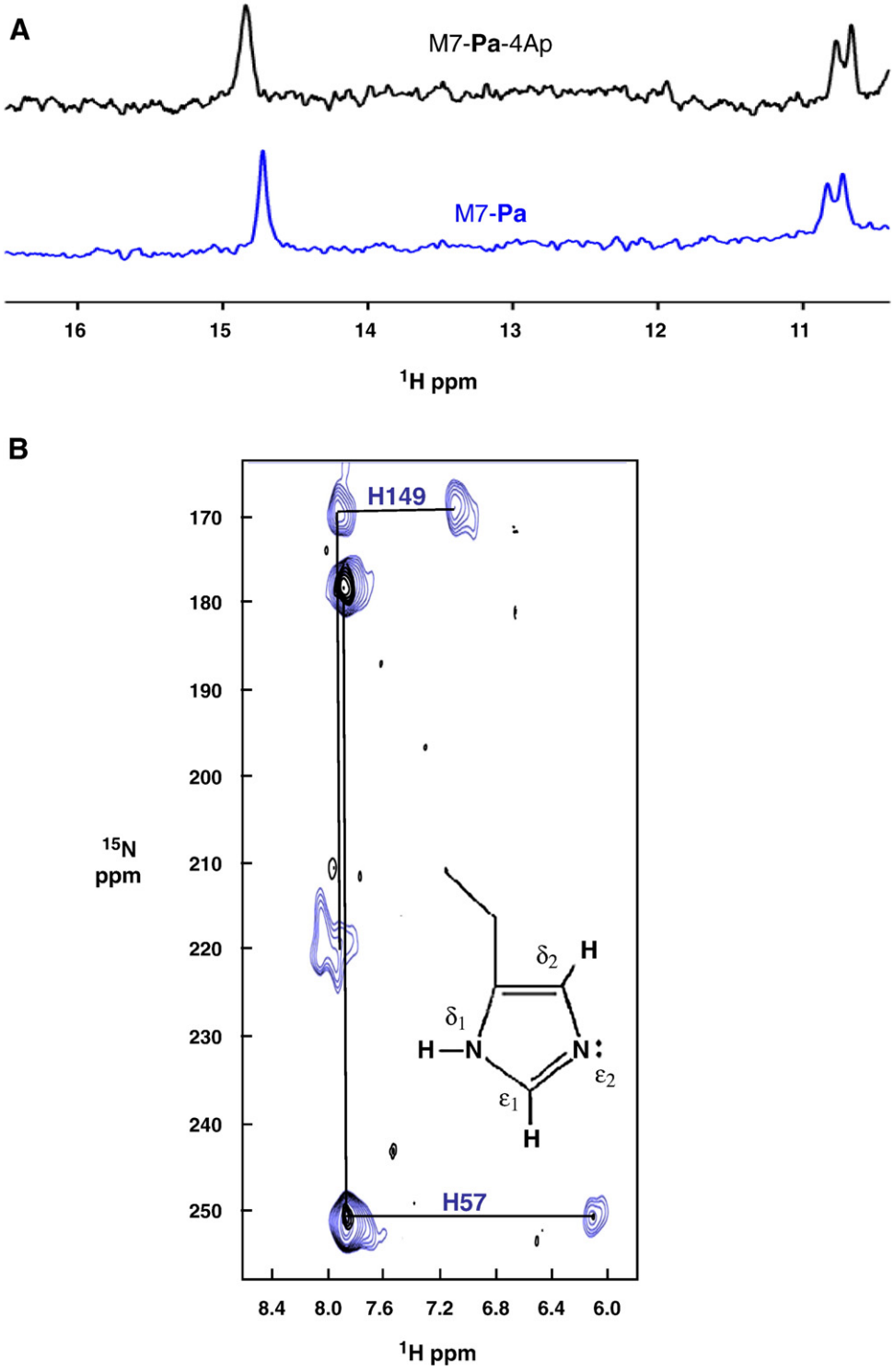

Fig. 8. Effect of cofactor binding to M7–Pa on the imidazole ring nuclei of H57. Spectra of the M7–Pa complex in the absence or presence of the 4Ap peptide are depicted in blue and black, respectively. (A) Chemical shift changes in the low field signal belonging to the Nδ1-attached proton observed in the one dimensional ^{1}H spectra. (B) ^{1}H–^{15}N long range HSQC spectrum for the correlation of Hε1 and Hδ2 with Nδ1 and Nε2. The type of correlation corresponds to an Nδ1-protonated tautomer in both cases. H149 cross-peaks were not visible in the conditions in which the spectrum of M7–Pa–4Ap was acquired.

binding, in order to evaluate the role of entropy in the formation of the complex.

A total of 14 out of 20 threonine side chains in the M7–**Pa** complex, comprising T123, were assigned using standard 2D and 3D experiments. Most of the non-assigned threonines belong to the first 21 amino acids, which gave no detectable signals as in other NS3p complexes (Gallo et al., 2009, Barbato et al., 2000). As a simple measurement of the dynamical behavior of the side chains, we used the intensity of the Hβ–Cβ cross-peaks in the ^{1}H–^{13}C 2D constant-time correlation experiment (Fig. S2A). In this experiment, the intensity of the cross-peak is predominantly dictated by the ^{13}C transverse relaxation time T_2 and therefore the larger the flexibility of

the side chain, the more intense the cross-peak will result. Upon examining these intensities for all threonines, T123, T160 and T177 appear to exhibit the greatest side chain flexibility. As a support for this analysis, peak intensities were related to the Accessible Solvent Area (ASA) of the threonine side chain calculated using the high resolution structure of NS3p1b as a model (Kim et al., 1996). It is expected that solvent-exposed threonines will exhibit more flexible side chains leading to a direct relationship between ASA and the cross-peak intensity. Fig. 9A shows the plot of ASA against peak intensity, indeed revealing the expected correlation, with the three flexible threonine residues being the most solvent-exposed. There are two notable exceptions from this behavior: T22 and T40, both showing a diminished cross-peak intensity with respect to that expected based on their solvent exposure in the structure of free NS3p. T22 is the first observed amino acid for M7: the first 21 residues are broad and not detectable, probably due to conformational exchange. It is possible that T22 shows a diminished intensity due to a contribution of μs–ms movements. T40 is part of the inhibitor binding site, and the presence of **Pa** therefore results in a lower ASA in the complex with respect to that predicted by analysis of the free NS3p structure.

Assignment of the side chain NH_2 groups of glutamine (Q) and asparagine (N) residues was possible for 9 of a total of 12 present in M7 (Fig. S2B). To determine the degree of flexibility that these side chains show, we measured the ^{15}N–{^1H} NOE of the side chain NH_2 group. For residues that do not show local mobility, NOE values close to +0.8 are expected. Conformational flexibility of the side chains will translate into a significant decrease of the NOE value. Fig. 9B shows the heteronuclear NOE for the Q and N residues. Q8 and Q9 were tentatively assigned based on the fact that these two glutamines show chemical shift close to random coil values and that the heteronuclear NOE values are between −2 and −3, indicating a very flexible

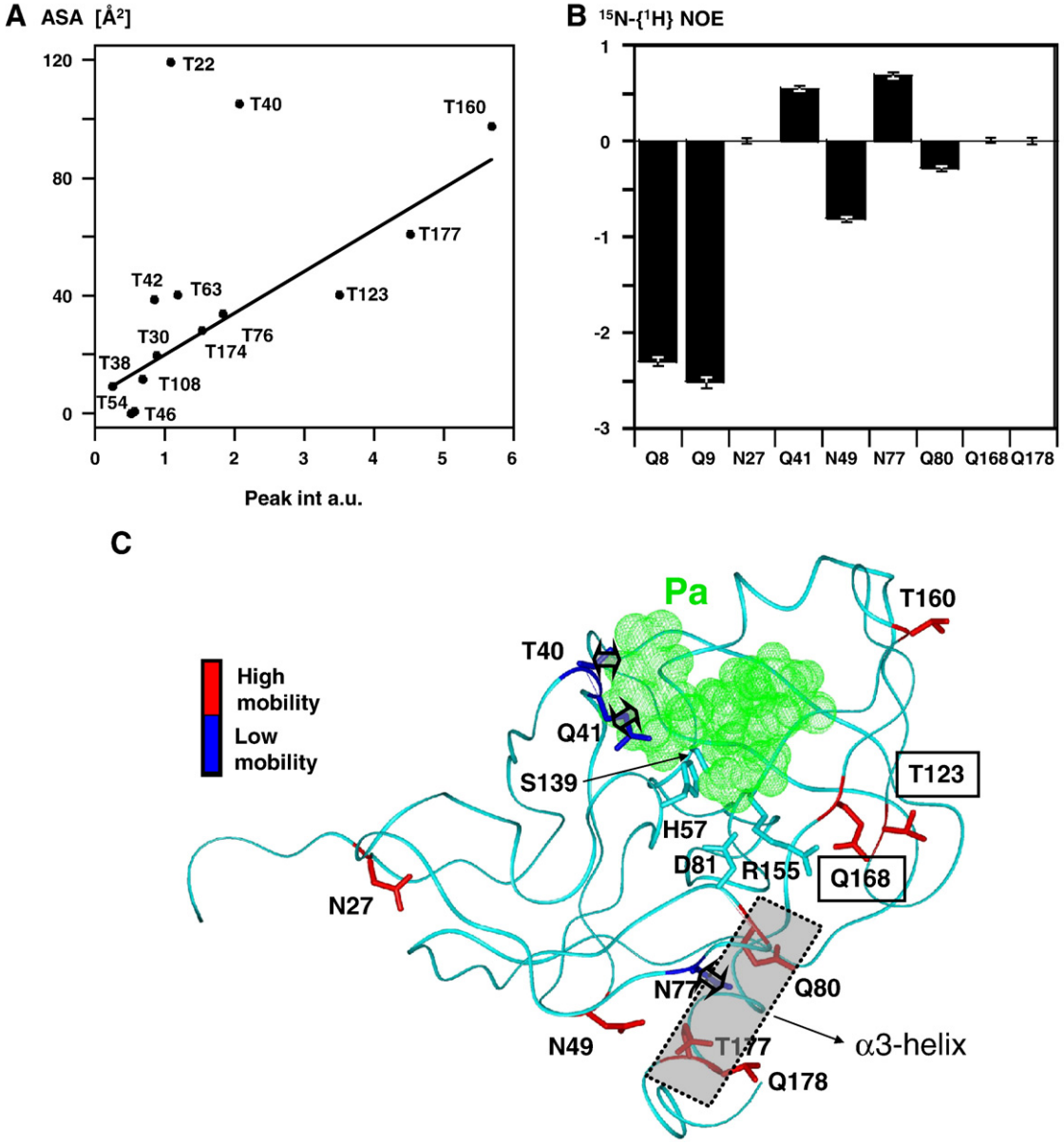

Fig. 9. Side chain dynamics of T123 and Q168 key residues in the M7 protease. (A) Correlation between Accessible Surface Area (ASA) and Hβ–Cβ cross-peak intensity for assigned threonine residues in a constant-time ^1H–^{13}C HSQC spectrum of the M7–Pa complex. (B) Observed heteronuclear ^{15}N–{^1H} NOEs for assigned asparagines and glutamine NH_2 groups in the same complex. (C) Structure location and side chain mobility of selected threonine, glutamine and asparagines residues. T, Q and N residues were classified according to the mobility of their side chains and mapped on a model of the structure M7–Pa complex, based on the structure of an analog inhibitor complexed to NS3p1b (PDB entry 2K1Q). Highly and lowly mobile residues are indicated in red and blue, respectively. The respective side chains are evidenced. S139, H57, D81 and R155 are also shown in light blue. Inhibitor Pa is depicted in green.

situation. These two residues are located in the N-terminal stretch, whose backbone amide NH groups could not be detected in the HSQC spectrum of the M7–**Pa** complex. Other very flexible side chains are those of N49 and Q80, which appear solvent exposed in a model of the M7–**Pa** complex (see below). A heteronuclear NOE close to 0 was observed for N27, Q168 and Q178, providing evidence of significant local mobility of the side chain. The only two side chains of Q and N residues that show some degree of rigidity are those of Q41 and N77.

To rationalize these results, we have indicated the position of some of these residues in the structure of a homology model for the M7–**Pa** complex, built from the recently determined structure of an analogous inhibitor bound to the NS3p1b (Gallo et al., 2009) (Fig. 9C). Q41 as well as T40 are involved in the interaction with the phenyl ring of **Pa**, and this fact explains the rigidity of the two side chains. Likewise, the N77 side chain, which shows a large positive heteronuclear NOE, participates in an interaction between the E1b-F1 β-strand and the α3 helix. This interaction is important to maintain the catalytic aspartate aligned with the histidine residue (Kim et al., 1996; Barbato et al., 1999). All the other N and Q residues resulted flexible, and in the structure are mostly solvent exposed, including N27, N49, Q80 and Q178. Regarding the two key residues for the difference in inhibitor affinity T123 and Q168, they both resulted very mobile. This scenario is expected to be very different for the 1b genotype, for which the network of salt bridges R155–D168–R123 is likely to maintain the three side chains in a more rigid conformation.

Discussion

By using NMR spectroscopy we have performed the first structural characterization of the HCV NS3 protease belonging to the 3a genotype. This characterization was made feasible by engineering seven mutations of surface amino acids that rendered the protein more soluble without changing its behavior against a panel of inhibitors. As observed for NS3p3a, the mutated protein (M7) showed a decreased susceptibility to inhibition by P2–P4 macrocyclic inhibitors like compounds **Ma** or **Mb** compared to the 1b enzyme. In addition, a linear inhibitor without large substituents at the P2 position, like the phenethylamide **Pa**, showed very similar potency on all three NS3 proteins. Binding of the **Pa** inhibitor was also able to structurally stabilize the protease, improving the quality of the NMR spectra with respect to those recorded for the less stable free protease. The phenyl ring of **Pa** confers a particular stabilization on the active site by binding in the S1′ pocket, as previously observed for a complex between NS3p1b and an analog of **Pa** (Gallo et al., 2009).

Two other potent NS3 inhibitors that are currently in clinical trials are VX-950 and SCH-503034, also known as Velaprevir and Boceprevir, respectively. In contrast to the inhibitors studied in this work, BILN 2061, **Ma** and **Mb,** and **Pa**, which are all reversible non-covalent inhibitors, VX-950 and SCH-503034 are reversible covalent inhibitors. Structurally, VX-950 and SCH503034, are ketoamide linear compounds, larger than **Pa**, with a P4 moiety interacting directly with position 123 (Tong et al., 2006; Guo et al., 2006). However, regarding the genotype potency shift, both VX-950 and SCH-503034 exhibit approximately a 10-fold loss of potency against the 3a genotype with respect to the 1b genotype (Flores et al., 2009), in comparison with the more than 200-fold shift displayed by BILN 2061 and the two or three orders of magnitude decrease of K_i for 3a vs. 1b genotype here observed for **Ma** and **Mb** compounds. The reduced sensitivity to the NS3 genotype of VX-950 and SCH-503034 was attributed to a greater flexibility of these acyclic inhibitors, which allow them to better adapt to changes in the enzyme surface. The macrocyclic moiety confers to BILN 2061, and **Ma** and **Mb**, a rigidity that is responsible of the greater potency against genotype 1, but may render it more sensitive to naturally occurring variations near the binding site (Tong et al., 2006; Guo et al., 2006).

In our earlier NS3p1b investigation we could not obtain evidence about interaction between NS3p and peptides derived from its natural cofactor NS4Ap (Gallo et al., 2009). Here, working with the M7 protein and using the peptide cofactor from genotype 1b (4Ap), we have observed an evident interaction and the corresponding stabilization of the active site caused by the peptide. In fact, the cross-interaction between 4Ap with NS3p3a, at least in the conditions of our study, leads to a more clearly observable interaction than when using the homologous 3a cofactor peptide. This is especially noteworthy considering that in the two peptides, 13 out of 20 residues are identical. However, this effect is not unprecedented: no sufficiently stable complex has been so far obtained for the 1b genotype protease with its cognate cofactor peptide for NMR studies (Archer et al., 2002; Cicero et al., 1999), but the 2a genotype NS3 protein was able to interact with the genotype 1b cofactor giving rise to a stable complex in solution (Wright-Minogue et al., 2000), which could be further stabilized by interaction with an inhibitor (Archer et al., 2002). In line with our results, it was recently noted that in one case NS3p3a was over two-fold more active in the presence of the NS4A cofactor of genotype 1b, than in the presence of cofactors from any other genotype, including its proper cofactor. Indeed, this pairing was the most active cross-genotypic combination in 28 heterologous interactions tested (Franco et al., 2008).

Our results strongly indicate that **Pa** and 4Ap produce a very similar conformation of the active site, as judged by the chemical shifts observed for the amide HN nuclei of H57 and G58, the low field signal corresponding to H57 HNδ1, and the two nitrogen nuclei of the imidazole ring of the M7–**Pa** and M7–4Ap–**Pa** complexes. These results strengthen our findings that binding to the S′ pockets, in a similar way to the binding mode of the substrate, can induce an active site conformation that is very similar to that predicted for the protease–cofactor complex in terms of catalytic triad alignment (Gallo et al., 2009). In addition, our data show that the conformational properties of the two residues responsible for the inhibitor potency shift, 123 and 168, are not affected by cofactor interaction, indicating that their conformational properties in the binary complex M7–**Pa** can be extrapolated to the more physiologically relevant ternary complex, M7–**Pa**–4Ap.

Interaction of NS3p with the macrocyclic inhibitor **Ma** yielded conformationally unstable complexes with proteases belonging to genotypes 1b and 3a. However, the active site shows signs of stabilization as suggested by the appearance of the signal belonging to G58. Our chemical shift perturbation data indicate that the structures of the two **Ma** complexes are very similar. This similarity suggests that the different chemical nature of residues 123 and 168 is not favoring a different binding mode of macrocyclic inhibitors. Considering that the two final states do not differ, it is likely that differences in the conformational features of the two proteins that precede the interaction play a significant role in determining the change in affinity of **Ma** for NS3p1b and NS3p3a.

The structural basis for this differential inhibition profile between proteases of the 1b and 3a genotypes is still an open question, although some hints are emerging (Beyer et al., 2001; Courcambeck et al., 2006; Guo et al., 2006; Thibeault el al., 2004; Tong et al., 2006; Lin et al., 2004, 2005; Trozzi et al., 2003). The proteins belonging to the two genotypes show a high degree of amino acid identity (75.8% average, see Fig. S3). As already discussed, some studies have identified residues in position 123 and 168 to be the cause of this potency shift (Tong et al., 2006; Thibeault et al., 2004; Beyer et al., 2001). To shed light on this subject, our purpose was to study how this region is presented to the macrocyclic inhibitor and more specifically how the dynamics of these residues in the free enzyme may affect the binding affinity for the P2–P4 macrocyclic inhibitors. To this end, we used the **Pa**–NS3p complex as a model for the free protein, based on the lack of interaction of **Pa** with residues 123 and 168 as shown by the high-resolution structure of the NS3p1b–**Pa** analog complex

(Gallo et al., 2009). This feature is reflected in the similar inhibitory potency that inhibitor **Pa** exhibits on both genotype enzymes.

Binding of P2–P4 macrocyclic compounds **Ma** and **Mb**, presenting an extra substituent in P2, is expected to involve more directly the region containing amino acids 155–123–168. As no structures of this type of inhibitor bound to the NS3 protease are currently available, and because complexes between this enzyme and **Ma** are not amenable to detailed structural analysis by NMR, we generated a model of the NS3 protease–**Ma** complex. In this model, the linker between the P2 and P4 side chains lies just over Q168 of NS3p3a and over D168 of NS3p1b. This direct interaction explains our observed chemical shift differences for this residue between the M7–**Pa** and M7–**Ma** complexes, and contributes in turn to the decrease in K_i for **Ma** and **Mb**. The same kind of interaction of the side chain of D168 was observed in a complex of NS3p1b with a P2–P4 macrocyclic inhibitor bearing a bulky P2 substituent 4-hydroxyproline (Arasappan et al., 2006). For position 123, the interaction with P2–P4 macrocyclic inhibitors is not direct. In the case of genotype 1b, which presents a pre-organized 155–168–123 region to interact with this kind of inhibitors, there is no significant chemical environment change around this residue upon binding, in line with the lack of chemical shift difference between NS3p1b–**Pa** and –**Ma** complexes. In contrast, in the 3a genotype protease there is a considerable conformational change around residue 123 when the enzyme binds the inhibitor, causing the observed chemical shift difference between M7–**Pa** and M7–**Ma**.

In the context of the S2 pocket of NS3p, interaction of the isoquinoline of macrocycles **Ma** and **Mb** with R155 is expected to force an extended conformation of its side chain, favoring the formation of a salt-bridge R155–D168 in NS3p1b (Courcambeck et al., 2006; Tsantrizios et al., 2003; Guo et al., 2006). The lack of this salt-bridge in NS3p3a, due to the D168Q substitution, may be one of the factors contributing to the decrease in affinity for compounds bearing a bulky P2 substituent (Tong et al., 2006). A second effect is the increase in the solvation of R155 when the partially neutralizing negative charge of D168 is removed (Trozzi et al., 2003). Binding of inhibitors making extensive contacts with R155 side chain will need to pay an energetic price in order to desolvate the charged group, which will be more solvated in the absence of a salt bridge with D168. However, in the context of NS3p3a protease R123 is also mutated into a neutral amino acid (threonine), preserving the net charge of the region. This in turn explains why the BILN2061 inhibitor shows a slightly higher affinity for a NS3p3a/1b chimera in which both amino acids 123 and 168 were mutated than for the single-mutated D168Q enzyme (Thibeault et al., 2004).

Our results point to a third effect, involving the flexibility of T123 and Q168 side chains in the 3a protein. In addition to a non-ideal conformation for inhibitor binding, one has to consider also the conformational freedom that will be lost upon inhibitor binding, particularly for Q168 in the case of compounds **Ma** and **Mb**, or T123 for inhibitors like hexapeptides (Beyer et al., 2001). From the secondary structure analysis performed here (Fig. 3C), the overall structure of the M7–**Pa** complex is very similar to that of the corresponding 1b enzyme, suggesting that the reason for the different inhibitor profiles of the two proteins by the P2–P4 macrocycles may reside at the level of side chain chemical and conformational properties. Therefore, it is the dynamical behavior of the side chains of these two key residues, T123 and Q168, which adds another piece of information to aid the construction of an atomic model that explains the differences in inhibitory potency of some compounds on the 1b vs. 3a NS3 protease. Fig. 10 and Table 1 summarize our model for the loss of entropy produced in the 3a but not in the 1b genotype NS3p when binding inhibitors that exploit residues 123 and/or 168 for the interaction. In the case of NS3p3a, side chains of residues T123, Q168, and possibly R155 are flexible. When the complex is formed, the large P2 substituent of the inhibitor interacts with R155 and Q168,

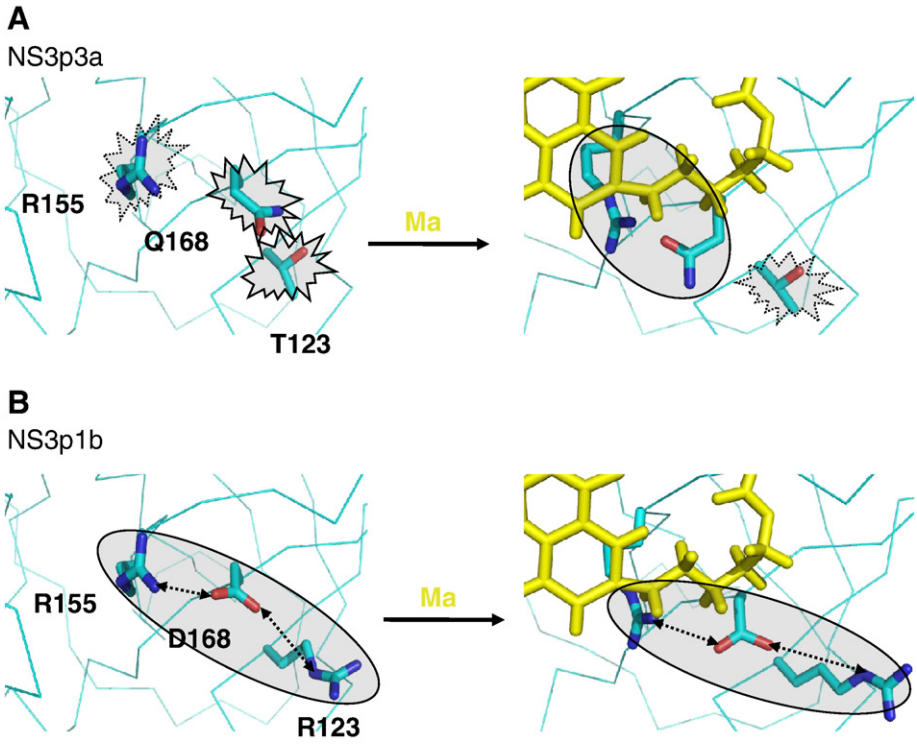

Fig. 10. Model for the loss of entropy produced in the NS3p from genotype 3a but not from genotype 1b when binding P2–P4 macrocyclic inhibitors. (A) NS3p3a in the free protein and in the complex with Ma. Side chains of residues R155, Q168 and T123 are evidenced. Ma is depicted in yellow. In the free protein residues Q168, T123 and, possibly, R155 are flexible. In the complex with Ma, R155 and Q168 directly interacting with the inhibitor are in a rigid conformation. (B) The same for NS3p1b. The interactions between R155, D168 and R123 are evidenced. These residues are rigid in the apo-NS3p1b and continue to be rigid upon binding Ma.

Table 1
Model for the loss of entropy associated with the genotype 1b vs. 3a NS3p inhibitor potency shift.

Inhibitor	Potency shift between 1b and 3a genotypes	Interaction with residues 123/168	Side chain residues 123/168 dynamics	
			NS3p1b	NS3p3a
Pa	No	No	123 rigid/168 rigid	123 flexible/168 flexible
Ma, Mb	Yes	123 indirect/168 direct	123 rigid/168 rigid	123 not known/168 rigid

which experience a transition from a mobile to a rigid conformation. Therefore, binding of these compounds will impose a decrease in side chain mobility with a consequent entropic penalty. The situation is very different for the genotype 1b, where such an entropic price is not necessary for binding of **Ma** or **Mb** to the enzyme, because the two consecutive salt bridges will maintain the side chains of R123, D168, and R155 in rigid positions.

In conclusion, our results have generated a plausible explanation for the observed loss in potency when comparing inhibitors of the NS3 proteases from HCV genotypes 1 and 3, with ramifications for the evaluation and improvement of potentially therapeutic pan-genotype compounds currently under development. For example, NS3p1b presents a rigid and flat surface around the S2 and S4 pockets that ultimately led to the identification of inhibitors with flat and bulky substituents in P2 and P4 positions. Such a strategy produces less potent inhibitors of the 3a genotype NS3 protease and additional favorable interactions are needed to counterbalance the entropic penalty imposed by flexible side chains of this region in NS3p3a.

Acknowledgments

We thank John Butcher and Nigel Liverton (MRL West Point, USA) for supplying compounds **Ma** and **Mb**, Frank Narjes (IRBM) for supplying compound **Pa**, Raffaele DeFrancesco (IRBM) for fruitful discussions, and the IRBM for the grant that supported part of this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2010.05.035.

References

Arasappan, A., Njoroge, F.G., Chen, K.X., Venkatraman, S., Parekh, T.N., Gu, H., Pichardo, J., Butkiewicz, N., Prongay, A., Madison, V., Girijavallabhan, V., 2006. P2–P4 macrocyclic inhibitors of hepatitis C virus NS3-4A serine protease. Bioorg. Med. Chem. Lett. 16, 3960–3965.

Archer, S.J., Camac, D.M., Wu, Z.J., Farrow, N.A., Domaille, P.J., Wasserman, Z.R., Bukhtiyarova, M., Rizzo, C., Jagannathan, S., Mersinger, L.J., Kettner, C.A., 2002. Hepatitis C virus NS3 protease requires its NS4A cofactor peptide for optimal binding of a boronic acid inhibitor as shown by NMR. Chem. Biol. 9, 79–92.

Bachovchin, W.W., 1986. ^{15}N NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: evidence for a moving histidine mechanism. Biochemistry 25, 7751–7759.

Barbato, G., Cicero, D.O., Nardi, M.C., Steinkühler, C., Cortese, R., De Francesco, R., Bazzo, R., 1999. The solution structure of the N-terminal proteinase domain of the hepatitis C virus (HCV) NS3 protein provides new insights into its activation and catalytic mechanism. J. Mol. Biol. 289, 371–384.

Barbato, G., Cicero, D.O., Cordier, F., Narjes, F., Gerlach, B., Sambucini, S., Grzesiek, S., Matassa, V.G., De Francesco, R., Bazzo, R., 2000. Inhibitor binding induces active site stabilization of the HCV NS3 protein serine protease domain. EMBO J. 19, 1195–1206.

Beyer, B.M., Zhang, R., Hong, Z., Madison, V., Malcolm, B.A., 2001. Effect of naturally occurring active site mutations on hepatitis C virus NS3 protease specificity. Proteins 43, 82–88.

Bukh, J., Miller, R.H., Purcell, R.H., 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin. Liver Dis. 15, 41–63.

Cicero, D.O., Barbato, G., Ingallinella, P., Bianchi, E., Nardi, M.C., Steinkühler, C., Cortese, R., Matassa, V., De Francesco, R., Pessi, A., Bazzo, R., 1999. Structural characterization of the interactions of optimized product inhibitors with the N-terminal proteinase domain of the hepatitis C virus (HCV) NS3 protein by NMR and modelling studies. J. Mol. Biol. 289, 385–396.

Cicero, D.O., Contessa, G.M., Paci, M., Bazzo, R., 2006. HACACO revisited: residual dipolar coupling measurements and resonance assignments in proteins. J. Magn. Reson. 180, 222–228.

Cornilescu, G., Delaglio, F., Bax, A., 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR 13, 289–302.

Courcambeck, J., Bouzidik, M., Perbost, R., Jouirou, B., Amrani, N., Cacoub, P., Pèpe, G., Sabatier, J.M., Halfon, P., 2006. Resistance of hepatitis C virus to NS3-4A protease inhibitors: mechanisms of drug resistance induced by R155Q, A156T, D168A and D168V mutations. Antrivir. Ther. 11, 847–855.

Dahl, G., Sandström, A., Åkerblom, E., Danielson, U.H., 2006. Resistance profiling of hepatitis C virus protease inhibitors using full-length NS3. Antivir. Ther. 12, 733–740.

De Francesco, R., Migliaccio, G., 2005. Challenges and successes in developing new therapies for hepatitis C. Nature 436, 953–960.

De Francesco, R., Carfi, A., 2007. Advances in the development of new therapeutic agents targeting the NS3-4A serine protease or the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. Adv. Drug Deliv. Rev. 59, 1242–1262.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293.

Di Bisceglie, A.M., Hoofnagle, J.H., 2002. Optimal therapy of hepatitis C. Hepatology 36, S121–S127.

Franco, S., Clotet, B., Martínez, M.A., 2008. A wide range of NS3/4A protease catalytic efficiencies in HCV-infected individuals. Virus Res. 131, 260–270.

Flores, M.V., Strawbridge, J., Ciaramella, G., Corbau, R., 2009. HCV-NS3 inhibitors: determination of their kinetic parameters and mechanism. Biochim. Biophys. Acta 1794, 1441–1448.

Gale Jr, M., Froy, E.M., 2005. Evasion of intracellular host defence by hepatitis C virus. Nature 436, 939–945.

Gallo, M., Pennestri, M., Bottomley, M.J., Barbato, G., Eliseo, T., Paci, M., Narjes, F., De Francesco, R., Summa, V., Koch, U., Bazzo, R., Cicero, D.O., 2009. Binding of a noncovalent inhibitor exploiting the S′ region stabilizes the hepatitis C virus NS3 protease conformation in the absence of cofactor. J. Mol. Biol. 385, 1142–1155.

Guo, Z., Prongay, A., Tong, X., Fischmann, T., Bogen, S., Velazquez, F., Venkatraman, S., Njoroge, F.G., Madison, V., 2006. Computational study of the effects of mutations A156T, D168V, and D168Q on the binding of HCV protease inhibitors. J. Chem. Theory Comput. 2, 1657–1663.

Hinrichsen, H., Benhamou, Y., Wedemeyer, H., Reiser, M., Sentjens, R.E., Calleja, J.L., Forns, X., Erhardt, A., Cronlein, J., Chaves, R.L., Yong, C.L., Nehmiz, G., Steinmann, G. G., 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.

Holloway, M.K., Liverton, N.J., McCauley, J.A., Rudd, M.T., Vacca, J.P., Ludmerer, S.W., Olsen, D.B., 2007. Preparation of macrocyclic peptides as HCV NS3 protease inhibitors. WO 2007016441 A1.

Johnson, B.A., 2004. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Meth. Mol. Biol. 278, 313–352.

Kay, L.E., Torchia, D.A., Bax, A., 1989. Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 28, 8972–8979.

Kim, J.L., Morgenstern, K.A., Lin, C., Fox, T., Dwyer, M.D., Landro, J.A., Chambers, S.P., Markland, W., Lepre, C.A., O'Malley, E.T., Harbeson, S.L., Rice, C.M., Murcko, M.A., Caron, P.R., Thompson, J.A., 1996. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. Cell 87, 343–355.

Lamarre, D., Anderson, P.C., Bailey, M., Beaulieu, P., Bolger, G., Bonneau, P., Bos, M., Cameron, D.R., Cartier, M., Cordingley, M.G., Faucher, A.M., Goudreau, N., Kawai, S.H., Kukolj, G., Lagace, L., LaPlante, S.R., Narjes, H., Poupart, M.A., Rancourt, J., Sentjens, R.E., St George, R., Simoneau, B., Steinmann, G., Thibeault, D., Tsantrizos, Y.S., Weldon, S.M., Yong, C.L., Llinàs-Brunet, M., 2003. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. Nature 426, 186–189.

Lin, C., Thomson, J.A., Rice, C., 1995. A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine proteinase complex *in vivo* and *in vitro*. J. Virol. 69, 4373–4380.

Lin, C., Lin, K., Luong, Y.P., Govinda Rao, B., Wei, Y.Y., Brennan, D.L., Fulghum, J.F., Hsiao, H.M., Ma, S., Maxwell, J.P., Cottrell, K.M., Perni, R.B., Gates, C.A., Kwong, A.D., 2004. *In vitro* resistance studies of hepatitis C virus serine protease inhibitors, VX-950 and BILN2061. J. Biol. Chem. 279, 17508–17514.

Lin, C., Gates, C.A., Govinda Rao, B., Brennan, D.L., Fulghum, J.R., Luong, Y.P., Frantz, J.D., Lin, K., Ma, S., Wei, Y.Y., Perni, R.B., Kwong, A.D., 2005. *In vitro* studies of cross-resistance mutations against two hepatitis C virus serine protease inhibitors, VX-950 and BILN2061. J. Biol. Chem. 280, 36784–36791.

Liverton, N.J., Holloway, M.K., McCauley, J.A., Rudd, M.T., Butcher, J.W., Carroll, S.S., DiMuzio, J., Fandozzi, C., Gilbert, K.F., Mao, S.S., McIntyer, C.J., Nguyen, K.T., Romano, J.J., Stahlhut, M., Wan, B.L., Olsen, D.B., Vacca, J.P., 2008. Molecular modeling based approach to potent P2–P4 macrocyclic inhibitors of hepatitis C NS3/4A protease. J. Am. Chem. Soc. 130, 4607–4609.

Lu, L., Pilot-Matias, T.J., Stewart, K.D., Randolph, J.T., Pithawalla, R., He, W., Huang, P.P., Klein, L.L., Mo, H., Molla, A., 2004. Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor *in vitro*. Antimicrob. Agents Chemother. 48, 2260–2266.

Malancona, S., Colarusso, S., Ontoria, J.M., Marchetti, A., Poma, M., Stansfield, I., Laufer, R., Di Marco, A., Taliani, M., Verderame, M., Gonzalez-Paz, O., Matassa, V.G., Narjes, F., 2004. SAR and pharmacokinetic studies on phenethylamide inhibitors of the hepatitis C virs NS3/NS4A serine protease. Bioorg. Med. Chem. Lett. 14, 4575–4579.

Mao, S.S., DiMuzio, J., McHale, C., Burlein, C., Olsen, D.B., Carroll, S.S., 2008. A time-resolved, internally quenched fluorescence assay to characterize inhibition of hepatitis C virus nonstructural protein 3-4A protease at low enzyme concentrations. Anal. Biochem. 373, 1–8.

McCoy, M.A., Senior, M.M., Gesell, J.J., Ramanathan, L., Wyss, D.F., 2001. Solution structure and dynamics of the single-chain heptatitis C virus NS3 protease NS4A cofactor complex. J. Mol. Biol. 305, 1099–1110.

Nolte, F.S., 2001. Hepatitis C virus genotyping: clinical implications and methods. Mol. Diagn. 6, 265–277.

Pawlotsky, J.M., 2003. Hepatitis C virus genetic variability: pathogenic and clinical implications. Clin. Liver Dis. 7, 45–66.

Pelton, J.G., Torchia, D.A., Meadow, N.D., Roseman, S., 1993. Tautomeric states of the active site histidines of phosphorylated and unphosphorylated IIIGlc, a signal transducing protein from *Escherichia coli* using two dimensional heteronuclear NMR techniques. Protein Sci. 2, 543–558.

Reed, K.E., Rice, C.M., 2000. The hepatitis C viruses. In: Hagedorn, C.H., Rice, C.M. (Eds.), Current Topics in Microbiology and Immunology. Springer, Heidlberg, Germany, pp. 55–84.

Reiser, M., Hinrichsen, H., Benhamou, Y., Sentjens, R.E., Wedemeyer, H., Calleja, J.L., Cronlein, J., Yong, C.L., Nehmiz, G., Steinmann, G., 2003. Antiviral effect of BILN 2061, a novel HCV serine protease inhibitor, after oral treatment over 2 days in patients with chronic hepatitis C, non-genotype 1. Hepatology 38 Abstract No. 136.

Steinkühler, C., Biasiol, G., Brunetti, M., Urbani, A., Koch, U., Cortese, R., Pessi, A., De Francesco, R., 1998. Product inhibition of the hepatitis C virus NS3 protease. Biochemistry 37, 8899–8905.

Thibeault, D., Bousquet, C., Gingras, R., Lagacé, L., Maurice, R., White, P.W., Lamarre, D., 2004. Sensitivity of NS3 serine proteases from hepatitis C virus genotypes 2 and 3 to the inhibitor BILN2061. J. Virol. 78, 7352–7359.

Tong, X., Guo, Z., Wright-Minogue, J., Xia, E., Prongay, A., Madison, V., Qiu, P., Venkatraman, S., Velazquez, F., Njoroge, F.G., Malcolm, B.A., 2006. Impact of naturally occurring variants of HCV protease on the binding of different classes of protease inhibitors. Biochemistry 45, 1353–1361.

Trozzi, C., Bartholomew, L., Ceccacci, A., Biasiol, G., Pacini, L., Altamura, S., Narjes, F., Muraglia, E., Paonessa, G., Koch, U., De Francesco, R., Steinkühler, C., Migliaccio, G., 2003. *In vitro* selection and characterization of hepatitis C virus serine protease variants to an active-site peptide inhibitor. J. Virol. 77, 3669–3679.

van den Berg, S., Löfdahl, P.A., Härd, T., Berglund, H., 2006. Improved solubility of TEV protease by directed evolution. J. Biotechnol. 121, 291–298.

Wasley, A., Alter, M.J., 2000. Epidemiology of hepatitis C: geographic differences and temporal trends. Semin. Liver Dis. 20, 1–16.

Wright-Minogue, J., Yao, N., Zhang, R., Butkiewicz, N.J., Baroudy, B.M., Lau, J.Y.N., Hong, Z., 2000. Cross-genotypic interaction between hepatitis C virus NS3 protease domains and NS4A cofactors. J. Hepatol. 32, 497–504.