



Protease and helicase activities of hepatitis C virus genotype 4, 5, and 6 NS3–NS4A proteins

Marie-Josée Massariol, Songping Zhao, Martin Marquis, Diane Thibeault, Peter W. White *

Boehringer Ingelheim (Canada) Ltd., Department of Biological Sciences, 2100 Cunard St., Laval, QC, Canada H7S 2G5

ARTICLE INFO

Article history:

Received 15 October 2009

Available online 26 November 2009

Keywords:

NS3 protease

NS3 helicase

JFH1

HCV genotype 4

HCV genotype 5

HCV genotype 6

ABSTRACT

The bifunctional NS3 protease–helicase of hepatitis C virus (HCV), together with its cofactor protein NS4A, is an important target for antiviral drugs which can cure HCV infections. HCV strains are divided into six major genotypes based on sequence diversity, and the great majority of reports on NS3 have focused exclusively on genotype 1 proteins. Here we report the cloning, expression, and preliminary characterization of NS3–NS4A gene products from HCV genotypes 4, 5, and 6. This work complements our earlier characterization of genotype 2 and 3 proteins [17]. We compare NS3–NS4A protease and helicase activities of genotypes 4a, 5a, and 6a to those of common reference strains Con1 (genotype 1b) and JFH1 (genotype 2a). The specific activities of the proteases of the newly isolated proteins were similar to those of the reference proteins. Furthermore, the reference inhibitor BILN 2061 had similar activity against all of the proteins except for that of JFH1, which had an apparent K_i that was 11-fold higher relative to Con1. RNA and DNA unwinding activities were also similar for genotypes 1, 4, 5, and 6 proteins, but significantly higher for genotype 2 JFH1. With the availability of these proteins, inhibitors developed based on their activity against genotype 1 can be tested against all the other major genotypes, providing a path to improved treatment for all HCV patients.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Hepatitis C virus (HCV) infection is a serious worldwide health problem, affecting an estimated 170 million people. In 70% of cases infection is not resolved, and chronically infected individuals are at risk of progressive liver disease which can lead to cirrhosis and hepatocellular carcinoma [1]. Prevalence of HCV infection is 1–2% in North America, Europe, and Japan, but reaches up to 15% in some Southeast Asian and African nations, notably Egypt [2], where the infection was spread by use of incompletely sterilized needles for a national parenteral schistosomiasis treatment lasting into the 1980s [3,4].

HCV can evolve rapidly, which has resulted in a wide variation in genomic sequences. Variants have been divided by phylogenetic analysis into at least six major genotypes, differing in nucleotide sequence by 31–33%. Genotypes are further subdivided into subtypes which differ from each other by 20–25% [5]. Geographical analysis of current infections suggests that genotypes 1, 2, 4, and 5 originated in sub-Saharan Africa and genotypes 3 and 6 in South-east Asia [6]. In Europe, North America, and Japan genotypes 1–3 account for the great majority of infections. In contrast, the other genotypes are largely confined to their geographical origins, or to

immigrants from those regions currently residing in Europe or North America [6], but they can be highly prevalent in these populations.

HCV infection can be cured by a 24–48 week treatment with pegylated interferon α and ribavirin. Success of treatment depends in part on the genotype with which a person is infected, although the reasons for this are not yet well understood. Success rates are highest for genotypes 2 and 3 (approximately 80%) and lowest for genotype 1 (approximately 40–50%) [7]. The success rates for treating genotype 4, 5, and 6 infections seem to be intermediate [8]. Several direct-acting antiviral therapies are making their way through clinical trials and offer the promise of improved success rates for treatment as well as a reduction in side effects [9]. Development of these agents has focused primarily on their activity against genotype 1, which in addition to being poorly treated by current therapy accounts for the majority of infections worldwide. Much less is known about the effectiveness of these therapies against other genotypes.

The HCV genome contains approximately 9600 bases and is translated into a single 3000 amino acid polyprotein [10]. The polyprotein is then cleaved into 10 proteins, and four of these cleavages are carried out by the virally encoded NS3–NS4A protein [11]. NS3 is a 631 amino acid bifunctional protein with an N-terminal serine protease domain (180 amino acids) and a C-terminal Superfamily II RNA helicase domain (451 amino acids) [12]. NS4A

* Corresponding author. Fax: +1 450 682 4642.

E-mail address: peter.white@boehringer-ingelheim.com (P.W. White).

is a 54 amino acid hydrophobic protein. Ten central amino acids bind to the protease domain of NS3, forming one strand of a β -sheet [13]. The NS3 protease activity can be evaluated using either the full-length NS3–NS4A protein, which autocleaves itself during heterologous expression in baculovirus or *Escherichia coli*, or the 180 amino acid protease domain, plus an excess of peptide corresponding to the central hydrophobic region of NS4A [14].

The serine protease activity of NS3–NS4A protein has been the subject of intense drug-discovery research for over a decade. Our research led to the discovery of BILN 2061 [15], the first small molecule viral protein-targeted inhibitor to be tested in HCV patients [16] (though later discontinued). We have focused our work on the genotype 1b protein, but have also reported the cloning, expression, and purification of genotype 1a, 2, and 3 NS3–NS4A products, as well as their inhibition by BILN 2061 [17]. Recently, inhibition of genotype 4, 5, and 6 protease domains by BILN 2061 and other inhibitors has been described, but no details of protein sequence, purification, or kinetic characterization were given [18]. Here, we present for the first time the preliminary characterization of NS3–NS4A gene products from clinical isolates of genotype 4, 5, and 6. For reference, we also purified and characterized, under the same conditions, NS3–NS4A proteins from the reference genotype 1b sequence Con1 [19] and from the genotype 2a JFH1 strain, the first HCV sequence isolated which is capable of full viral replication *in vitro* [20].

Materials and methods

Cloning. Human plasma samples were obtained in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). HCV RNA was isolated from these samples using the QiaAmp Viral RNA mini kit (Qiagen) according to the supplier's instructions. Subtypes for each HCV RNA sample were determined by sequencing the NS5B polymerase region of the genome [21].

For genotype 4a NS3–NS4A, isolated viral RNA was reverse transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen) with a mix of HCV-specific primers corresponding to sequences located in NS4B gene (primers 1, 2, and 3, Table 1). A fragment encoding NS3–NS4A plus portions of NS2 and NS4B was amplified from the cDNA using the KOD Hot Start polymerase

(Novagen), together with the same mix of primers used for the cDNA and a mix of forward NS2 primers 4 and 5. The purified fragment was cloned into pCR-Blunt II-TOPO using the Zero Blunt TOPO Cloning kit (Invitrogen). The NS3–NS4A fragment was then amplified using primers 6 and 7, which contain restriction sites for cloning (underlined sequences in Table 1) and ligated (Roche Rapid DNA Ligation kit) into pET 29a + (Novagen), which provides a 28-residue N-terminal sequence containing a hexahistidine tag and a tobacco etch virus (TEV) protease cleavage site.

Genotype 5a and 6a NS3–NS4A were amplified using the SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen). For genotype 5a, NS2 forward primer (primer 12) was used with a mix of NS4B reverse primers (primers 8, 9, 10, and 11). The RT-PCR was purified with PCR purification kit (Qiagen), reamplified using primers 10, 11, and 12 and cloned into pCR-Blunt vector with the Zero Blunt PCR Cloning kit (Invitrogen). The NS3–NS4A was then reamplified from this vector with a specific forward primer (primer 13) and a reverse primer (primer 14) and subcloned into pET 11a + (Novagen).

For genotype 6a, a mix of NS2 forward primers (primers 15, 16, and 17) with a mix of NS4B reverse primers (primers 18 and 19) were used for the RT-PCR. A second round of amplification was performed with the same mix of primers, and the product was cloned in pCR-Blunt II-TOPO. Specific primers 20 and 21 were used to reamplify the NS3–NS4A sequence and subcloned into pET 11a + (Novagen).

The sequences of the cloned coding regions for genotypes 4a, 5a, and 6a NS3–NS4A have been deposited in GenBank: genotype 4a GenBank ID: GU085486; genotype 5a GenBank ID: GU085487; genotype 6a GenBank ID: GU085488. Con1-R3 and JFH1 sequences were amplified from the corresponding subgenomic replicons and subcloned into pET 11a + (Con1-R3) or pET 29a + (JFH1) to give constructs analogous to those described above for genotypes 4, 5, and 6. The Con1-R3 sequence is a replicon-adapted variant of Con1, and this sequence has also been deposited in GenBank: Con1-R3 GenBank ID: GU085489.

Expression and purification of NS3–NS4A proteins. NS3–NS4A proteins were expressed in *E. coli* BL21 (DE3) (Novagen), except for genotype 5a which was expressed in Rosetta (DE3) (Novagen). Protein expression was induced for 18 h at 18 °C, and NS3–NS4A proteins were purified using nickel and poly-uridine affinity resins. Procedures for expression and purification were the same for all five proteins and were very similar to those previously described [17]. Purity was examined using SDS–PAGE (10% bis–Tris gel, Invitrogen) and the enzyme concentration was determined using the Bradford protein assay (Bio-Rad), as well as by active site titration (see below).

Protease activity assay. Kinetic parameters for NS3–NS4A proteins were determined using the depsipeptide fluorogenic substrate anthraniloyl-DDIVP-Abu[C(O)-O] AMY(3-NO₂)-TW-OH. Cleavage reactions were performed in white 96-well plates (Dynatech Microfluor), and were monitored continuously at 23 °C on a PolarStar Galaxy plate reader (BMG), equipped with excitation and emission filters of 320 and 405 nm, respectively. Assays used 0.5–12.5 μ M of substrate, and 5 nM of each enzyme (based on Bradford) in a buffer composed of 50 mM Tris–HCl, pH 8.0, 0.25 M sodium citrate, 0.01% *n*-dodecyl- β -D-maltoside, 1 mM TCEP. Total assay volume was 50 μ L/well. Assays were initiated by addition of enzyme, and kinetic parameters were determined by non-linear regression analysis of initial rate as a function of substrate concentration, using GraphPad Prism software (version 4.02; GraphPad Software, Inc., La Jolla, USA).

Since proteins were 60–80% pure, active site titrations with the fluorogenic substrate Ac-DED(EDANS)EE-Abu[C(O)-O]-ASK(DAB-CYL)-NH₂ were used to determine active protein concentrations. The formation of fluorogenic product was continuously monitored

Table 1
Oligonucleotides used for NS3–NS4A cloning.

Ref. #	Oligonucleotide sequence (5'–3')
#1	GCCAGTTGTAACCCGTGCTCGACTAGTGG
#2	GCTAGCTGTAGGCCATGCTCGACTAGTGG
#3	CTGGTGTTCCTCTTGRGCTTGCTGCCAGCG
#4	TCGGCYAGGTGGGCAATGAGATCTTGCTCGG
#5	TCAGCGAGGCTGGCAATGAAATCTTGCTCGG
#6	GGCCATGGATCGGCCCCCATACGCGCTACGCGCAGCAG
#7	GCGGCCAAGCTTCACTATTAGCACTCTCCATTTCGTGCAACTGTTG
#8	AATTGYCCGCAATGGCRCNGCYTCGTCCA
#9	CCGAGCACTTTCYTTGAATTGCCCGGC
#10	AATTGCYCGCRATRGACGCGCTCGTCCA
#11	GCRATRRRCRGCTGCTCTCCATATAGG
#12	GCCTTGCTCCACCTTGGTAGGCTGACCG
#13	CGCCGCCATGGATCTGCCCCATCACTGCGTACGCGCAGCAGAC
#14	CGCTCTCTCGAGTCACTATTAACTCTCCATCTCATCAAACTGC
#15	TCTCHATGGAGAAGAARGTCATCACYTGG
#16	ATGGAGAAGAARATYATYACTGGGGTGCGGA
#17	GAGCCGGTYATYTTCTCYCCCATGGAGAAG
#18	AYRAAGTTCGCACAGRTGTTCTGCCAGAA
#19	GTGCCAGAATTGCTCAATCTTCGGCCAGG
#20	CGCATGGATCCGGCTCCCATACCCGCTATGCGCAGC
#21	CGCTCTCTCGAGTCACTATTAGCACTCTCCATCTCGTCAAACTGC

Underlined bases correspond to restriction sites used in cloning.

using an Envision Plate Reader (Perkin-Elmer) for 1800 s in reactions using 2 nM enzyme (by Bradford) and 2 μ M substrate in presence of 0–10 nM of a very potent protease inhibitor ($K_i = 30$ pM). The steady state velocity at each inhibitor concentration was determined by linear regression of the steady state rate. This was followed by plotting the ratio of steady state velocity at each inhibitor concentration and velocity in absence of inhibitor as function of inhibitor concentration. Data for each protein were fit to the equation below [22] to yield the active enzyme concentration:

$$\frac{v_s}{v_o} = 1 - \frac{([E]_T + [I]_T + K_i^{app}) - \sqrt{([E]_T + [I]_T + K_i^{app})^2 - 4[E]_T[I]_T}}{2[E]_T}$$

where v_s is steady state velocity at different inhibitor concentration, v_o is velocity in absence of inhibitor, E_T is total concentration of enzyme, I_T is total concentration of inhibitor, K_i^{app} is the apparent dissociation constant.

Inhibition assays. To evaluate inhibition, 5 μ M of substrate anthraniloyl-DDIVP-Abu[C(O)-O] AMY(3-NO₂)-TW-OH was incubated with various concentrations of BILN 2061, in presence of 0.5 nM NS3–NS4A. The incubation was done at 23 °C for 60–70 min. The reaction was terminated by addition of 1 M MES (pH 5.8), and fluorescence of the N-terminal product anthraniloyl-DDIVP-Abu was measured using a PolarStar Galaxy plate reader (BMG). Calculated percent inhibition at each inhibitor concentration was then used to determine the median effective concentration (IC₅₀) by the NLIN procedures of SAS (Statistical Software System; SAS Institute Inc., Cary, NC). K_i values were calculated from IC₅₀ values using the equation $IC_{50} = K_i(1 + [S]/K_m)$ [23].

Helicase and ATPase activity comparison. Both dsDNA and dsRNA unwinding was evaluated. The DNA substrate was prepared by hybridization in 1:1.6 ratio of the following oligonucleotides: 5'-GAATACAAGCTTGGGCTGCAGGTCGACTCTAGAG-Cy3-3' (34-mer) and 5'-BHQ-1CTCTAGAGTCGACCTGCAGCCCAAGCTGTTTTTTTTTTT-3' (45-mer) in 25 mM Hepes, pH 7.5, 500 mM NaCl, 0.1% SDS, 1 mM EDTA. The RNA substrate was prepared similarly using 1:1.6 ratio of the following: 5'-GAAUACAAGCUUGGGCUGCAG GUCG ACUCUAGAG-Cy3-3' (34-mer) and 5'-BHQ-1-CUCUAGAGUC GACCUGCAGCCCAAGCUUGUUUUUUUUUUUUUUUUU-3' (45-mer), in the same buffer as for the dsDNA substrate. RNA and DNA oligonucleotides labeled with Cy3 fluorophore or BHQ (black hole quencher) were obtained from Integrated DNA Technologies. A 10-fold excess of complementary ssDNA trap, complementary to the Cy3 oligonucleotide, was used during the reaction to prevent the reannealing of the unwound DNA or RNA. Final assay conditions were the following: 50 mM PIPES pH 6.5, 2 mM TCEP, 3 mM MgCl₂, 2.5% DMSO, 0.0625% *n*-dodecyl- β -D-maltoside, 5 nM substrate (dsDNA or dsRNA), 100 μ g/mL BSA, 50 nM ssDNA trap, and 2 mM ATP. In assays involving dsRNA substrate, 0.15 U/ μ L of RNA-sin was also added. Final enzyme concentration was 10 or 15 nM for dsDNA or dsRNA unwinding assays, respectively (based on active enzyme concentration). Reactions were performed in 384-well plates, initiated by the addition of enzyme, and monitored continuously for 1800–6000 s at room temperature on an Envision Multi-label Plate Reader (with a 531 nm excitation filter and a 595 nm emission filter). Reaction rates were determined from the linear portion of each progress curve. The percentage of unwinding was estimated by comparison to the fluorescence of unhybridized Cy3 oligonucleotide (DNA or RNA as appropriate).

Nucleic acid-stimulated ATPase activity was measured using the Kinase-GloPlus Luminescent Kinase Assay kit (Promega). ATPase reactions were performed in 384-well plates, using the following conditions: 50 mM PIPES pH 6.2, 2 mM TCEP, 3 mM MgCl₂, 3% DMSO,

0.0625% *n*-dodecyl- β -D-maltoside, 100 μ g/mL BSA, 20 μ M ATP, 10 nM enzyme, and 50 nM dU₂₄. After a 10 min incubation period at room temperature, 10 μ L of the assay mixture was transferred to another 384-well plate containing 10 μ L of Kinase-Glo Plus reagent. Luminescence was measured after 30 min, using a TopCount plate reader. A standard curve for ATP concentration was used to estimate the amount of ATP consumed in each sample.

Results and discussion

Cloning, expression and purification of NS3–NS4A

HCV RNA was obtained from plasma of patients infected with HCV genotype 4a, 5a, or 6a. The entire NS3–NS4A coding sequences were amplified and subcloned into expression vectors to generate constructs with N-terminal poly-histidine tags to facilitate purification. The reference NS3–NS4A from genotype 1b (Con1-R3, a cell culture-adapted Con1 [19]) and 2a (JFH1 [20]), which match our cell culture replicons, were prepared in an identical fashion. Amino acid sequences of the five different proteins are given in Fig. 1.

Proteins were purified using nickel and poly-uridine affinity columns. The poly-histidine tags were not removed since we found this had little effect on protease activity. Purity after the poly-uridine column ranged from approximately 60% (Con1-R3 and 4a) to 80% (JFH1, 5a, and 6a) based on SDS–PAGE analysis. Final yields were 2.0, 0.3, 0.5, 0.3, and 1.5 mg protein/L of culture for Con1-R3, JFH1, 4a, 5a, and 6a proteins, respectively, based on Bradford protein concentrations.

Protease activity of purified NS3–NS4A proteins

Protease assays were performed using a fluorogenic decapeptide substrate modeled on the NS5A–NS5B cleavage site (polyprotein cleavage sites are conserved among all HCV genomes). Steady state initial rates were determined at multiple substrate concentrations, to yield values of k_{cat} and K_m (Table 2). Despite some variation in the individual k_{cat} and K_m values, the values for catalytic efficiency (k_{cat}/K_m) are similar, within 40% of the mean. Values for k_{cat} were calculated using active enzyme concentrations, obtained by titration of a very potent hexapeptide inhibitor. Stock solutions of purified proteins contained 4.3, 4.0, 4.2, 5.4, and 6.9 μ M NS3–NS4A for Con1-R3, JFH1, 4a, 5a, and 6a preparations, respectively. These ranged from 40% to slightly over 100% of the total protein concentration for each preparation estimated by Bradford.

Inhibition of protease activity

IC₅₀ values were determined for the inhibitor BILN 2061 against each of these proteins (Table 2). We previously reported BILN 2061 IC₅₀ values, measured under essentially identical conditions, for genotype 1a, 1b, 2a, 2b, and 3a proteins, although the 1b protein was a different strain from the Con1-R3 sequence used in this work [17]. For a more accurate comparison of inhibitor affinities, approximate K_i values were calculated from IC₅₀ values by taking into account the ratio of substrate concentration to its K_m (see Materials and methods). The genotype 4, 5, and 6 enzymes have similar affinity for BILN 2061 to that observed for genotype 1 protein. Although BILN 2061 is no longer in development, this result shows that some HCV protease inhibitors optimized for activity against genotype 1 may be effective against these three less prevalent genotypes. The value of K_i for genotype 2a is 11-fold higher, although this is still slightly lower than the K_i values we found

			1		60
R3	GT	1b	(1)	APITAYSQQTRG LLGCI ITSLTGRD RNQVE GEVQV VS TATQ SFL ATCVNGV CV WT VY HGAG	
JFH1	GT	2a	(1)	APITAYAQQTRG LLGAI IVSMTGRD RTEQ AGEVQ IL STVS QS FL GT IT SG VLW TV YHGAG	
	GT	4a	(1)	APITAYAQQTRG LF STIV SL TGRD TN NCGEVQ VL STATQ SFL GT IT SGV LM WT VY HGAG	
	GT	5a	(1)	APITAYAQQTRG VLGAI IVSLTGRD RNEA GEVQ VL STATQ TFL GT CT ING VM WT VY HGAG	
	GT	6a	(1)	APITAYAQQTRG LVGT IV SL TGRD KNE VEGEVQV VS TATQ SFL AT S ING VM WT VY HGAG	
			61		120
R3	GT	1b	(61)	SKTL AC PK GP IT Q MY T NVD Q DL V GW Q AP PG AR SL TP CTCG SS DL YL VT KH AD VI P VRR RG	
JFH1	GT	2a	(61)	NKTL AC LR GP VT Q MY S SA EG DL V GW P SP PG TK SL EP CK CA VD LY LT RN AD VI P ARR RG	
	GT	4a	(61)	SKT IS GP KG PN Q MY T NVD Q DL V GW P SP PG VK SL TP CTCG SS DL YL LT RH AD VI P VRR RG	
	GT	5a	(61)	SKTL AC PK GP AV Q MY T NVD K DL V GW P SP PG TR SL TP CTCG SAD LY LT R HAD VI PARR RG	
	GT	6a	(61)	SKTL AC PK GP VC Q MY T NVD K DL V GW P SP PG AR SL TP CTCG SS DL YL LT RE AD VI P ARR RG	
			121		180
R3	GT	1b	(121)	DS R G SL LS SP RF V SV L KG SS GG P LL CP S GH AV GI FRA AV CT RG VAK AV DF VP ES ME TT MR	
JFH1	GT	2a	(121)	DK R G ALL SP RE IST L KG SS GG P VL CP R GH VV GL FRA AV CS RG VAK S DF IP VE TL DV VT R	
	GT	4a	(121)	DT R G ALL SP RE IST L KG SS GG P LL CP M GH AA GI FRA AV CT RG VAK AV DF VP ES LE TT MR	
	GT	5a	(121)	DA R S ALL SP RE IST L KG SS GG P IM CP S GH VV GI FRA AV CT RG VAK AV DF IP VE NLE TT MR	
	GT	6a	(121)	DT R A ALL SP RE IST L KG SS GG P IM CP S GH VV GI FRA AV CT RG VAK S DF IP VE N ME TT MR	
			181		240
R3	GT	1b	(181)	SP V FT D N SS PP AV PQ T FQ VA HL HAPT SG SK TK V PA Y AA Q GY K VL V LN PS VA AT LG F G CA	
JFH1	GT	2a	(181)	SP T F S D NS TP PA VP Q TY Q VG Y L HAPT SG SK TK V PA Y AA Q GY K VL V LN PS VA AT LG F G CA	
	GT	4a	(181)	SP V FT D N ST PP AV PQ T FQ VA HL HAPT SG SK TK V PA Y AA Q GY K VL V LN PS VA AT LG F G CA	
	GT	5a	(181)	SP T F T D NS TP PA VP Q EF Q VG Y L HAPT SG SK TK V PA Y AA Q GY K VL V LN PS VA AT LG F G CA	
	GT	6a	(181)	SP S F T D NS TP PA VP Q TY Q VG Y L HAPT SG SK TR V PA Y AS Q GY K VL V LN PS VA AT LS F G CS	
			241		300
R3	GT	1b	(241)	Y MS K A H CID PN IR TG VR IT TC A PI T Y STY G K FLAD GG C SG AYD II IC D CH S T D ST TI	
JFH1	GT	2a	(241)	Y LS K A H CIN PN IR TG VR IT MT G AI T Y STY G K FLAD GG C AS AYD II IC D CH AV D AT SI	
	GT	4a	(241)	Y MS K AYCID PN IR S GV RT IT TC A P IT Y STY G K FL AD GG C SG AYD II IC D CH S T D ST TI	
	GT	5a	(241)	Y MS K AYCID PN IR TG VR IT IT G AI T Y STY G K FLAD GG C PG AYD VI IC D CH S O D AT TI	
	GT	6a	(241)	Y M R Q AY CV EP NI RT GV RT IT TC AI T Y STY G K FLAD GG C SG AYD II IC D CH S T D PT TV	
			301		360
R3	GT	1b	(301)	L GIG T VL DQ AE TAG AR L V LA TAT PP GS VT PH P NI EE VA LS ST GE IP FF Y K AI PI ET IK	
JFH1	GT	2a	(301)	L GIG T VL DQ AE TAG VR L V LA TAT PP GS VT PH P DI EE V GL RE GE IP FF Y K AI PI LS CI K	
	GT	4a	(301)	L GIG T VL DQ AE TAG VR L V LA TAT PP GS VT PH S NI EE VA LP TT GE IP FF Y K AI PI LE LI K	
	GT	5a	(301)	L GIG T VL DQ AE TAG AR L V LA TAT PP GS VT PH P NI EE VA LP SE GE IP FF Y K AI PI LA LI K	
	GT	6a	(301)	L GIG T VL DQ AE TAG VR L V LA TAT PP GS VT PH P NI TE T AL PT GE IP FF Y K GI PI LE Y IK	
			361		420
R3	GT	1b	(361)	G GR H LIF CH SK KK C DE LA AK L S GL CL NA VAY RG LD VS VI PT SG DV IV VA DAL MT GF TG	
JFH1	GT	2a	(361)	G GR H LIF CH SK KK C DE LA AL R G M GL NA VAY RG LD VS II PA Q GD V V VA D AL MT GY TG	
	GT	4a	(361)	G GR H LIF CH SK KK C DE LA KQ L T SL CL NA VAY RG LD VS VI PT SG DV V CA DAL MT GF TG	
	GT	5a	(361)	G GR H LIF CH SK KK C DE LA KQ L TAL GV NA VAY RG LD VA VI PTAG D V VC S D AL MT GF TG	
	GT	6a	(361)	G GR H LIF CH SK KK C DE LA KL L S GL CL NA VAY RG LD VS VI PT SG DV V CA DAL MT GY TG	
			421		480
R3	GT	1b	(421)	D FDS V IDC N CV ITQ V DF SL DP TF TI ET IT TV PQ DA VS RS QRR GT GR GR MC IT YR VF TP GE	
JFH1	GT	2a	(421)	D FD S VIDC N VA VTQ AV DF SL DP TF TI ET ITQ TV PQ DA VS RS QRR GT GR GR CT Y R V TS GE	
	GT	4a	(421)	D FD S VIDC N SV ITQ V DF SL DP TF TI ET IT TV PQ DA VS RS QRR GT GR GR LC IT YR V TS GE	
	GT	5a	(421)	D FD S VIDC N SA VTQ V DF SL DP TF TI ET IT TV PQ DA VS RS QRR GT GR GR HC IT YR V TS GE	
	GT	6a	(421)	D FD S VIDC N VA VTQ V DF SL DP TF SI ET IT TV PQ DA VS RS QRR GT GR GR KP CV YR VS Q GE	
			481		540
R3	GT	1b	(481)	R PS G MF D SV VL CE YD AG CA W Y EL TP AE TT SV RL RAY LNT P GL PVC Q D H LE F W EG V T GL T	
JFH1	GT	2a	(481)	R AS G MF D SV VL CE YD AG CA W Y DL TP AE TT V RL RAY FNT P GL PVC Q D H LE F W EA V T GL T	
	GT	4a	(481)	R PS G IF D TS VL CE YD AG CA W Y EL TP AE TT SV RL RAY FNT P GL PVC Q D H LE F W ES V T GL T	
	GT	5a	(481)	R PS G MF D SV VL CE YD SG CA W Y DL AP AE TT SV RL RAY LNT P GL PVC Q D H LE F W EG V T GL T	
	GT	6a	(481)	R PS G MF D SV VL CE AY DT CS W Y EL TP SE TT SV RL RAY LNT P GL PVC Q D H LE F W EG V T GL T	
			541		600
R3	GT	1b	(541)	H IDA H FL SQ AK Q CE N F YL VAY QAT VC ARA K AP PP SW D MM W K CL IR L K PT LH GP TP LL Y	
JFH1	GT	2a	(541)	H IDA H FL SQ AK Q CE N F AY L VAY Q AT VC ARA K AP PP SW D AM W K CL AR L K PT LAG PT PL Y	
	GT	4a	(541)	Q ID G H FL S Q AK Q CE N F YL VAY QAT VC AR AL AP PP SW D TM W K CL IR L K PT LH GP TP LL Y	
	GT	5a	(541)	N IDA H FL SQ AK Q CE N F YL VAY QAT VC VR AK AP PP SW D TM W K CL LR L K PT L T G PT PL Y	
	GT	6a	(541)	H IDA H FL SQ AK Q CE N F AY L VAY Q AT VC ARA K AP PP SW D VM W K CL IR L K PT L T G PT PL Y	
			601		660
R3	GT	1b	(601)	R LG AV Q NE V IT TH PT TK Y IM AC MS AD LEV VT ST W VL VGG V LA AL AA Y CL T GS V V TV GR I	
JFH1	GT	2a	(601)	R LG P IT NE V IT TH PT TK Y I AT CM AD LEV MT ST W VL AG GV LA AL AA Y CL AT GC V SI IG R L	
	GT	4a	(601)	R LG AV Q NE V IT TH PT TK Y IM AC MS AD LEV VT ST W VL VGG V LA AL AA Y CL SV G S V TV GR V	
	GT	5a	(601)	R LG P VQ NE IT L TH PT TK Y IM AC MS AD LEV IT ST W VL VGG V VA AL AA Y CL TV G SV AT V GR I	
	GT	6a	(601)	R LG P VQ NE IT L TH PT TK Y IM TC MS AD LEV IT ST W VL VGG V LA AL AA Y CL SV G CV IV CG R I	
			661		685
R3	GT	1b	(661)	I LS G RP AI IP D REV LY RF DE ME EC	
JFH1	GT	2a	(661)	H V N Q R V V VA PD KE V LY EAF DE ME EC	
	GT	4a	(661)	I LS G Q P AI IP DREV LY Q Q FDE ME EC	
	GT	5a	(661)	I LS G K P AI IP DREV LY Q Q FDE ME EC	
	GT	6a	(661)	T LT G K P AV V PDREV LY Q Q FDE ME EC	

Fig. 1. Amino acid sequence alignment of the five HCV NS3–NS4A sequences used in this study. Amino acids which are not identical for all five sequences are highlighted. Amino acids are numbered continuously; the cleavage site between Thr631 of NS3 and Ser-632, which is amino acid 1 of NS4A, is marked by an arrow. R3 is a replicon-adapted variant of Con1 [19] with Con1 residues Arg-109, Ser-534, and Lys-665 (in NS4A) substituted by Lys, Gly, and Arg, respectively.

for two other genotype 2 proteins [17]. Our IC₅₀ values for BILN 2061 against each of these genotypes are very similar to those reported by another group [18], even though the assay conditions

were different. Because enzyme *K_m* values were not provided in that report, this comparison is only approximate, based on IC₅₀ rather than *K_i* values.

Table 2

Kinetic and inhibition parameters for NS3–NS4A protease activity.

Genotype	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{M}^{-1} \text{s}^{-1}$)	BILN 2061 IC_{50} (nM)	BILN 2061 K_i (nM)
1b-Con1	2.1 ± 0.2	46 ± 3	3.7×10^5	4.7 ± 0.6	1.4
2a-JFH1	3.5 ± 0.2	48 ± 3	2.3×10^5	37 ± 3	15
4a	1.3 ± 0.1	12.6 ± 0.2	1.6×10^5	4.8 ± 0.9	1.0
5a	6.6 ± 1.4	102 ± 14	2.6×10^5	5.3 ± 1.1	3.0
6a	2.8 ± 0.3	50 ± 6	3.0×10^5	10.2 ± 1.6	3.7

K_m and k_{cat} values are reported as mean \pm standard deviation, determined in two independent experiments. IC_{50} values correspond to the mean \pm standard deviation from five independent experiments. K_i values were calculated from IC_{50} values as described in Materials and methods.

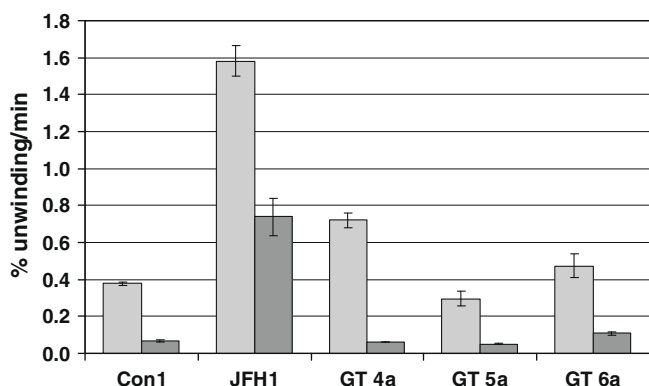


Fig. 2. Helicase activity of purified NS3–NS4A proteins. Light bars for dsDNA substrate unwinding; dark bars for dsRNA unwinding. Values correspond to the mean of four values from two independent experiments, and error bars represent standard deviations.

Helicase activity of NS3–NS4A proteins

Although the protease domain has been the primary focus of drug-discovery research on the NS3 protein, we also wished to verify that the C-terminal helicase domains of our NS3–NS4A proteins had ATP-dependent nucleic acid unwinding activity. All five proteins had similar ATP hydrolysis activity, with JFH1 being approximately 2-fold more active than the others (data not shown). Results for unwinding of dsDNA and dsRNA substrates are shown in Fig. 2. Since Michaelis–Menten kinetic parameters cannot be calculated for helicases under our assay conditions, activity was quantified using the linear portion of each reaction timecourse. As has been described previously [24], activity on RNA substrates was lower than on DNA substrates. These initial rates were generally similar for all the proteins, except that the JFH1 protein had a 7- to 14-fold higher RNA unwinding rate, as well as a 2- to 5-fold higher DNA unwinding rate, than any of the other proteins. Using His-tagged NS3 proteins in the absence of NS4A, a previous report also showed somewhat greater DNA unwinding activity (20–40%) for a different genotype 2a helicase compared to genotype 1b [24]. The JFH1 non-structural region replicates very efficiently [25], and it has been shown that the NS3 helicase portion of the genome is, in part, a determinant of this enhanced replication activity [26,27]. Additional experiments would be required to demonstrate that the efficient unwinding reported here contributes to the high replication capacity of JFH1.

Conclusion

We report here the first characterization of protease and helicase activities from genotype 4, 5, and 6 HCV NS3–NS4A proteins.

These proteins can now be used to determine the potential for NS3 protease inhibitors, usually optimized for activity against genotype 1, to be effective treatments for patients infected with genotypes 4, 5, and 6. Since our proteins also possess RNA unwinding activity, they will serve a similar purpose for NS3 helicase inhibitors, should this class of drug enter development.

Acknowledgments

We thank Muriel Lechevallier and Dr. Yves Benhamou, and their patients, for providing plasma samples. We also thank Mireille Potentier for help with helicase assays, Roger Maurice for help with protease assays and George Kukolj for careful reading of the manuscript.

References

- [1] J.H. Hoofnagle, Course and outcome of hepatitis C, *Hepatology* 36 (2002) S21–S29.
- [2] M.J. Alter, Epidemiology of hepatitis C virus infection, *World J. Gastroenterol.* 13 (2007) 2436–2441.
- [3] S.M. Kamal, I.A. Nasser, Hepatitis C genotype 4: what we know and what we don't yet know, *Hepatology* 47 (2008) 1371–1383.
- [4] C. Frank, M.K. Mohamed, G.T. Strickland, D. Lavanchy, R.R. Arthur, L.S. Magder, T. El Khoby, Y. Abdel-Wahab, E.S. Aly Ohn, W. Anwar, I. Sallam, The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt, *Lancet* 355 (2000) 887–891.
- [5] P. Simmonds, J. Bukh, C. Combet, G. Deleage, N. Enomoto, S. Feinstone, P. Halfon, G. Inchauspe, C. Kuiken, G. Maertens, M. Mizokami, D.G. Murphy, H. Okamoto, J.M. Pawlotsky, F. Penin, E. Sablon, I. Shin, L.J. Stuyver, H.J. Thiel, S. Viazov, A.J. Weiner, A. Widell, Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes, *Hepatology* 42 (2005) 962–973.
- [6] P. Simmonds, Genetic diversity and evolution of hepatitis C virus – 15 years on, *J. Gen. Virol.* 85 (2004) 3173–3188.
- [7] A.M. Di Bisceglie, J.H. Hoofnagle, Optimal therapy of hepatitis C, *Hepatology* 36 (2002) S121–S127.
- [8] M.H. Nguyen, E.B. Keeffe, Prevalence and treatment of hepatitis C virus genotypes 4, 5, and 6, *Clin. Gastroenterol. Hepatol.* 3 (2005) S97–S101.
- [9] V. Soriano, M.G. Peters, S. Zeuzem, New therapies for hepatitis C virus infection, *Clin. Infect. Dis.* 48 (2009) 313–320.
- [10] D. Moradpour, F. Penin, C.M. Rice, Replication of hepatitis C virus, *Nat. Rev. Microbiol.* 5 (2007) 453–463.
- [11] A. Grakoui, D.W. McCourt, C. Wychowski, S.M. Feinstone, C.M. Rice, Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polypeptide cleavage sites, *J. Virol.* 67 (1993) 2832–2843.
- [12] N. Yao, P. Reichert, S.S. Taremi, W.W. Prosser, P.C. Weber, Molecular views of viral polypeptide processing revealed by the crystal structure of the hepatitis C virus bifunctional protease–helicase, *Structure* 7 (1999) 1353–1363.
- [13] J.L. Kim, K.A. Morgenstern, C. Lin, T. Fox, M.D. Dwyer, J.A. Landro, S.P. Chambers, W. Markland, C.A. Lepre, E.T. O'Malley, S.L. Harbeson, C.M. Rice, M.A. Murcko, P.R. Caron, J.A. Thomson, Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide, *Cell* 87 (1996) 343–355.
- [14] D. Thibeault, M.J. Massariol, S. Zhao, E. Welchner, N. Goudreau, R. Gingras, M. Llinas-Brunet, P.W. White, Use of the fused NS4A peptide–NS3 protease domain to study the importance of the helicase domain for protease inhibitor binding to hepatitis C virus NS3–NS4A, *Biochemistry* 48 (2009) 744–753.
- [15] D. Lamarre, P.C. Anderson, M. Bailey, P. Beaulieu, G. Bolger, P. Bonneau, M. Bos, D.R. Cameron, M. Cartier, M.G. Cordingley, A.M. Faucher, N. Goudreau, S.H. Kawai, G. Kukolj, L. Lagace, S.R. LaPlante, H. Narjes, M.A. Poupard, J. Rancourt, R.E. Sentjens, R. St. George, B. Simoneau, G. Steinmann, D. Thibeault, Y.S. Tsantrizos, S.M. Weldon, C.L. Yong, M. Llinas-Brunet, An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus, *Nature* 426 (2003) 186–189.
- [16] H. Hinrichsen, Y. Benhamou, H. Wedemeyer, M. Reiser, R.E. Sentjens, J.L. Calleja, X. Forns, A. Erhardt, J. Cronlein, R.L. Chaves, C.L. Yong, G. Nehmiz, G.G. Steinmann, Short-term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients, *Gastroenterology* 127 (2004) 1347–1355.
- [17] D. Thibeault, C. Bousquet, R. Gingras, L. Lagace, R. Maurice, P.W. White, D. Lamarre, Sensitivity of NS3 serine proteases from hepatitis C virus genotypes 2 and 3 to the inhibitor BILN 2061, *J. Virol.* 78 (2004) 7352–7359.
- [18] S.D. Seiwert, S.W. Andrews, Y. Jiang, V. Serebryany, H. Tan, K. Kossen, P.T. Rajagopalan, S. Misialek, S.K. Stevens, A. Stoycheva, J. Hong, S.R. Lim, X. Qin, R. Rieger, K.R. Condroski, H. Zhang, M.G. Do, C. Lemieux, G.P. Hingorani, D.P. Hartley, J.A. Josey, L. Pan, L. Beigelman, L.M. Blatt, Preclinical characteristics of the hepatitis C virus NS3/4A protease inhibitor ITMN-191 (R7227), *Antimicrob. Agents Chemother.* 52 (2008) 4432–4441.
- [19] V. Lohmann, F. Korner, J. Koch, U. Herian, L. Theilmann, R. Bartenschlager, Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line, *Science* 285 (1999) 110–113.

- [20] T. Wakita, T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H.G. Krausslich, M. Mizokami, R. Bartenschlager, T.J. Liang, Production of infectious hepatitis C virus in tissue culture from a cloned viral genome, *Nat. Med.* 11 (2005) 791–796.
- [21] S. Laperche, F. Lunel, J. Izopet, S. Alain, P. Deny, G. Duverlie, C. Gaudy, J.M. Pawlotsky, J.C. Plantier, B. Pozzetto, V. Thibault, F. Tosetti, J.J. Lefrere, Comparison of hepatitis C virus NS5b and 5' noncoding gene sequencing methods in a multicenter study, *J. Clin. Microbiol.* 43 (2005) 733–739.
- [22] J.F. Morrison, C.T. Walsh, The behavior and significance of slow-binding enzyme inhibitors, *Adv. Enzymol. Relat. Areas Mol. Biol.* 61 (1988) 201–301.
- [23] Y. Cheng, W.H. Prusoff, Relationship between the inhibition constant (K₁) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction, *Biochem. Pharmacol.* 22 (1973) 3099–3108.
- [24] A.M. Lam, D. Keeney, P.Q. Eckert, D.N. Frick, Hepatitis C virus NS3 ATPases/helicases from different genotypes exhibit variations in enzymatic properties, *J. Virol.* 77 (2003) 3950–3961.
- [25] T. Kato, T. Date, M. Miyamoto, A. Furusaka, K. Tokushige, M. Mizokami, T. Wakita, Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon, *Gastroenterology* 125 (2003) 1808–1817.
- [26] A. Murayama, T. Date, K. Morikawa, D. Akazawa, M. Miyamoto, M. Kaga, K. Ishii, T. Suzuki, T. Kato, M. Mizokami, T. Wakita, The NS3 helicase and NS5B-to-3'X regions are important for efficient hepatitis C virus strain JFH-1 replication in Huh7 cells, *J. Virol.* 81 (2007) 8030–8040.
- [27] M. Binder, D. Quinkert, O. Bochkarova, R. Klein, N. Kezmic, R. Bartenschlager, V. Lohmann, Identification of determinants involved in initiation of hepatitis C virus RNA synthesis by using intergenotypic replicase chimeras, *J. Virol.* 81 (2007) 5270–5283.