Affinity selection of a camelized V\textsubscript{H} domain antibody inhibitor of hepatitis C virus NS3 protease

F.Martin, C.Volpari, C.Steinkuhler, N.Dimasi, M.Brunetti, G.Biasiol, S.Altamura, R.Cortese, R.De Francesco and M.Sollazzo\textsuperscript{1}

Istituto di Ricerche di Biologia Molecolare (IRBM) P.Angeletti, Via Pontina Km 30,600, 00040 Pomezia (Rome), Italy

\textsuperscript{1}To whom correspondence should be addressed

The HCV genome encodes, within the NS3 gene, a serine protease whose activity specifically cleaves the viral poly-protein precursor. Proteolytic processing of HCV poly-protein precursor by the viral NS3 protease is essential for virion maturation and designing specific inhibitors of this protease as possible anti-viral agents is a desirable and practical objective. With a view to studying both the function of HCV NS3 protease and to designing inhibitors of this enzyme, we directed our interest towards engineering macromolecular inhibitors of the viral protease catalytic activity. We describe here the affinity-selection and biochemical characterization of one inhibitor, cV\textsubscript{H}E2, a ‘camelized’ variable domain antibody fragment, isolated from a phage displayed synthetic repertoire, which is a potent and selective inhibitor of proteolysis by the NS3 enzyme. In addition to being useful as a biological probe to study the function of HCV protease, this inhibitor can serve as a potential pharmacophore model to design anti-virals. Moreover, the results suggest a way of engineering improved human-derived small recognition units tailored for enzyme inhibition.

Keywords: anti-viral/camelized antibody/hepatitis C virus/phage display/protease inhibitor

Introduction

Hepatitis C virus (HCV) is commonly accepted to be the major etiological agent of both parenterally transmitted and sporadic non-A non-B hepatitis (Choo et al., 1989; Kuo et al., 1989). It is estimated that this virus infects 100–200 million people worldwide and can lead to chronic and potentially life-threatening infections; liver cirrhosis and primary liver cell carcinoma can result in premature death in a substantial number of infected individuals (Chien et al., 1992). Protective immunotherapy for HCV infection is not available and interferon treatment is of only limited efficacy (Weiland, 1994). For these reasons, a thorough dissection of the HCV life cycle and the identification of targets for anti-viral therapies are important goals. The viral genome is ~9400 nucleotides long encoding a precursor polyprotein (Figure 1) of 3011 amino acids (Kato et al., 1990; Choo et al., 1991; Takamizawa et al., 1991; Grakoui et al., 1993). The enzymatic cleavage of the polyprotein precursor produces structural proteins and enzymes that are essential to the life cycle of HCV (Tomei et al., 1993). These proteins are released from the precursor polyprotein by cellular proteases in association with membranes of the endoplasmic reticulum (Hijikata et al., 1991), whereas all proteolytic cleavages downstream of the non-structural gene 3 (NS3) are catalyzed by a serine protease contained within the N-terminal region of NS3 (Figure 1). Cleavage activity at three of the four sites affected by NS3 occurs in \textit{trans}, whereas processing at the carboxy terminus of NS3 is a \textit{cis} event (Tomei et al., 1993). Substrate cleavage mediated by NS3 protease occurs between Cys/Ser or Cys/Ala and between Thr/Ser in \textit{cis} (Pizzi et al., 1994; Failla et al., 1996).

The product of the NS3 gene is a multidomain protein of 70 kDa that, in addition to the protease domain at the N-terminus, contains an RNA helicase at its C-terminus (Kim et al., 1995). It was demonstrated that a 20 kDa N-terminal fragment of NS3, in association with the viral polypeptide cofactor, NS4A, is capable of performing all cleavages in both \textit{in vitro} translation and transfection experiments with an efficiency indistinguishable from that of the wild type enzyme (Failla et al., 1995). The X-ray structure of the NS3 protease domain, devoid of co-factor (Love et al., 1996) and with the bound NS4A peptide (Kim et al., 1996), was determined. By analogy with data that have emerged from studies with flaviviruses (Chambers et al., 1990), it was presumed that inhibition of the NS3 protease activity could lead to the production of non-infectious viral particles, and hence this enzyme has become one of the main targets for anti-viral drug design.

In recent years, protease inhibitors, both low molecular weight compounds and proteinaceous molecules, have emerged as very important pharmaceutical agents (Neurath, 1989). By displaying on filamentous phage mutants of natural proteinaceous inhibitors, new specificities have been selected with high potency (Roberts et al., 1992; Dennis and Lazarus, 1994; Rottgen and Collins, 1995; Wang et al., 1995; Markland et al., 1996a,b). In this context, small protein scaffolds on to which functions can be engineered have emerged as a means of generating conformationally defined structures with potential as pharmacophores (Sollazzo et al., 1995; Zhao et al., 1995). In nature, immunoglobulins are unsurpassed as molecules for number of infected individuals (Chien et al., 1992). Protective immunotherapy for HCV infection is not available and interferon treatment is of only limited efficacy (Weiland, 1994). For these reasons, a thorough dissection of the HCV life cycle and the identification of targets for anti-viral therapies are important goals. The viral genome is ~9400 nucleotides long encoding a precursor polyprotein (Figure 1) of 3011 amino acids (Kato et al., 1990; Choo et al., 1991; Takamizawa et al., 1991; Grakoui et al., 1993). The enzymatic cleavage of the polyprotein precursor produces structural proteins and enzymes that are essential to the life cycle of HCV (Tomei et al., 1993). These proteins are released from the precursor polyprotein by cellular proteases in association with membranes of the endoplasmic reticulum (Hijikata et al., 1991), whereas all proteolytic cleavages downstream of the non-structural gene 3 (NS3) are catalyzed by a serine protease contained within the N-terminal region of NS3 (Figure 1). Cleavage activity at three of the four sites affected by NS3 occurs in \textit{trans}, whereas processing at the carboxy terminus of NS3 is a \textit{cis} event (Tomei et al., 1993). Substrate cleavage mediated by NS3 protease occurs between Cys/Ser or Cys/Ala and between Thr/Ser in \textit{cis} (Pizzi et al., 1994; Failla et al., 1996).

The product of the NS3 gene is a multidomain protein of 70 kDa that, in addition to the protease domain at the N-terminus, contains an RNA helicase at its C-terminus (Kim et al., 1995). It was demonstrated that a 20 kDa N-terminal fragment of NS3, in association with the viral polypeptide cofactor, NS4A, is capable of performing all cleavages in both \textit{in vitro} translation and transfection experiments with an efficiency indistinguishable from that of the wild type enzyme (Failla et al., 1995). The X-ray structure of the NS3 protease domain, devoid of co-factor (Love et al., 1996) and with the bound NS4A peptide (Kim et al., 1996), was determined. By analogy with data that have emerged from studies with flaviviruses (Chambers et al., 1990), it was presumed that inhibition of the NS3 protease activity could lead to the production of non-infectious viral particles, and hence this enzyme has become one of the main targets for anti-viral drug design.

In recent years, protease inhibitors, both low molecular weight compounds and proteinaceous molecules, have emerged as very important pharmaceutical agents (Neurath, 1989). By displaying on filamentous phage mutants of natural proteinaceous inhibitors, new specificities have been selected with high potency (Roberts et al., 1992; Dennis and Lazarus, 1994; Rottgen and Collins, 1995; Wang et al., 1995; Markland et al., 1996a,b). In this context, small protein scaffolds on to which functions can be engineered have emerged as a means of generating conformationally defined structures with potential as pharmacophores (Sollazzo et al., 1995; Zhao et al., 1995). In nature, immunoglobulins are unsurpassed as molecules for generating ligands for practically any target, especially as their variable regions can be displayed on the surface of filamentous phage, increasing their potential even further (reviewed by Winter et al., 1994). Antibody variable domains can provide a very useful source of structural templates for designing low molecular weight lead compounds based on their complementarity-determining regions (CDR) (Dougall et al., 1994; Sollazzo, 1995). In order to generate pharmacophore models, smaller sized antibody-derived fragments could be the molecules of choice as their binding activities can be focused on a smaller surface area.

Attempts were made to reduce the size of the minimum antibody fragment required for antigen binding. This led to the design of a ‘minimized’ V\textsubscript{H} domain molecule, the \textit{minibody} (Pessi et al., 1993; Tramontano et al., 1994) and the ‘camelized’ antibody (cV\textsubscript{H}), a modified human V\textsubscript{H} domain obtained through mimicking camel heavy chains for use as a small recognition unit (Davies and Riechmann, 1994; Riechmann, 1996).
Molecules specific for hapten, protein and peptide ligands with affinities in the micro-nanomolar range were isolated from a repertoire of phage-displayed cVH with randomized CDR3 (Davies and Riechmann, 1995). These results have been further validated by determining the X-ray structures of the natural Camelidae molecules (Desmyter et al., 1996; Spinelli et al., 1996), which provided the ultimate experimental evidence that V\textsubscript{H}-derived molecules represent the smallest natural immunoglobulin based recognition unit and open up new possibilities for generating and engineering macromolecules with highly specific binding properties.

From the analysis of the X-ray structure of NS3 protease, it was predicted that the design of low molecular weight inhibitors would be a very challenging task owing to the relatively featureless appearance of the substrate binding groove (Kim et al., 1996). With a view to studying both the function of HCV protease and the future design of small molecule inhibitors of its enzymatic activity, we sought to engineer macromolecules that would bind NS3 protease and to study their effects on the catalytic activity of this protease and ultimately on the viral life cycle. We describe here the isolation from a repertoire of camelized human antibody domain of one variant that binds to the HCV NS3 protease domain and inhibits selectively and competitively its enzymatic activity. This macromolecular inhibitor may serve as a probe for further characterizing the role this enzyme plays in HCV replication and as a basis for the development of anti-viral compounds. Also, the results reported here may provide clues for engineering improved small recognition unit enzyme inhibitors.

Materials and methods

**Microbiological and recombinant DNA techniques**

Microbiological and recombinant DNA methods were carried out according to standard protocols (Ausubel et al., 1994) or as recommended by suppliers. Oligonucleotides were synthesized using an Applied Biosystems (Foster City, CA) 380B synthesizer. Phage manipulation and *Escherichia coli* electroportation were carried out as described previously (Martin et al., 1994; Davies and Riechmann, 1995). Nucleotide sequences were determined using Sequenase (United States Biochemical, Cleveland, OH) according to the supplier’s recommendations. The design of cVH wild type has been described previously (Davies and Riechmann, 1994, 1995).

**Construction, expression and partial purification of NS3-myc protease**

The choice of fusing the myc epitope to the C-terminus of NS3 protease 20 kDa fragment was dictated by the assumption that modification of the N-terminus could interfere with interaction with NS4A co-factor (Failla et al., 1995; Kim et al., 1996). To construct NS3-myc, plasmid pT7-NS3 (1027–1206) encoding the HCV BK strain NS3 protease domain from residue 1027 to 1206, was used as a template for PCR amplification. Fusion of the myc epitope EFEKQKLISQDQLG to the C-terminus of the protease domain was obtained by using oligo-myc and NS3up as PCR primers (oligo-myc: GCTAGCCCGTCTGCCGGAGATCGTTGTGCTCGATTTCCGAGATTCCGTAGATGTTTCCATGG; NS3up: GCATACATGCGCCCATCAGCC). The PCR fragment was digested with *Nde*I and cloned into pT7–7 vector digested with *Nde*I and Smal. The resulting pT7-NS3-myc was fully sequenced to ensure that no mutations had been introduced by PCR. To induce expression of the protease, *E.coli* BL21 cells were transformed with pT7NS3-myc and grown at 37°C in LB medium to a density of 0.9 OD; 400 µM IPTG were added and cells were grown for a further 4 h at 23°C. Cells were then harvested by centrifugation and pellets were resuspended in 25 mM sodium phosphate buffer, pH 6.5, 50% glycerol, 10 mM DTT, 0.5% CHAPS, 10% glycerol (equilibration buffer) operating at 2 ml/min. The protease was eluted with equilibration buffer containing 1 M NaCl and dialyzed 1:100 for 8 h at 4°C against equilibration buffer. After this step the protease was 60% pure as judged by Coomassie-stained SDS–PAGE. The protease-containing solution was made 50% in glycerol content and kept at −20°C until use. The protease was immobilized on anti-myc coated Dynabeads (see below) and extensively washed with SLB. In order to check whether the immobilized protein was still enzymatically active, an aliquot of protease-bound Dynabeads was incubated for 6 h in SLB buffer containing 100 µM substrate peptide Ac-DEEMECASHLPYK. Cleavage of this peptide was verified by subsequent HPLC analysis of the solution.

**NS3 protease assays**

Assays were performed in 50 mM Tris, pH 7.5, 1% CHAPS, 15% glycerol, 0.1 mM DTT using 20 nM recombinant NS3 protease purified from *E.coli* as described previously (Steinkuhler et al., 1996a). The protease was preincubated for 15 min with a 14-mer peptide corresponding to the central domain of the protease co-factor NS4A with the sequence GSVVIVGRIILSGR (Tomei et al., 1996). To 60 µl of assay mix, up to 6 µl of cVH-E2 solutions (or corresponding amounts

---

**Fig. 1.** HCV polyprotein schematic structure and processing. Cleavages mediated by cellular proteases are marked with asterisks. Cleavage between NS2 and NS3 (◇) is mediated by the NS2/NS3 protease activity. NS3 protease cleaves between NS3 and NS4A (cis cleavage), NS4A and NS4B, NS4B and NS5A and NS5A and NS5B (arrowheads). Amino acid number and substrate sequence at which cleavage occurs are indicated.

<table>
<thead>
<tr>
<th>H2N</th>
<th>C</th>
<th>E1</th>
<th>E2/NS1</th>
<th>NS2</th>
<th>NS3</th>
<th>NS4A</th>
<th>NS4B</th>
<th>NS5A</th>
<th>NS5B</th>
<th>COOH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1658 DLEVVTST</td>
<td>◇</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1706 DEMEEFACAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>DCSTPCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2414</td>
<td>EDVVCCSM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*F.Martin et al.*
of buffer) were added and incubated for a further 30 min at 25°C. Reactions were started by adding 40 μM substrate peptide Ac-DEMEECASHLPYK-NH₂ and stopped by adding 40 μl of 1% TFA at <20% conversion. Samples were analyzed by HPLC using a Merck–Hitachi chromatograph equipped with an autosampler, column oven and fluorescence detector. Samples of 45–90 μl were injected on to a reversed-phase HPLC cartridge column (LiChrospher C18, 5 μm, 0.4 × 7.5 cm, Merck) equilibrated with 90% solvent A (H₂O–0.1% TFA) and 10% solvent B (acetonitrile–0.08% TFA) and operating at a flow rate of 2.5 ml/min. A 10–40% gradient of solvent B at 5%/min was used to separate cleavage fragments. Peaks were detected by monitoring tyrosine fluorescence (excitation at 260 nm, emission at 305 nm). Cleavage products were quantified by integrating chromatograms with respect to samples in which 100% conversion was achieved by 12 h of incubation in the presence of 2 μM protease.

IC₅₀ values were calculated by fitting inhibition data to Equation 1 using Kaleidagraph software:

\[ \text{Activity} \% = \left( \frac{\text{maximum activity}}{1 + \left( \frac{[I]}{IC_{50}} \right)^2} \right) \]

where \([I]\) is the cV₇₄E2 inhibitor concentration, maximum activity is that of the enzyme in the absence of inhibitor and \(S\) is the slope factor of the curve.

Reversibility of inhibition was assessed by dilution experiments. Briefly, NS3 protease was preincubated with cV₇₄E2 at a concentration of 4–5×IC₅₀ as described above. After 30 min, half the sample was diluted 10-fold and in both samples activity was determined and compared with the activity of a sample incubated with a buffer blank and a sample incubated from the beginning with a 10-fold diluted inhibitor concentration. Inhibition was defined as reversible if more than 75% of activity was recovered upon dilution. Inhibition mechanisms were determined by performing substrate titration experiments using concentrations of substrate peptide between 15 and 250 μM (\(K_{m}/3.3\) and 5×\(K_{m}\)) in the absence and presence of 125 and 250 nM of cV₇₄E2. Initial rates of cleavage were determined on samples with <20% conversion. Kinetic parameters were calculated from least-squares fit of initial rates as a function of substrate concentration with the help of Kaleidagraph or SigmaPlot software, assuming Michaelis–Menten kinetics. \(K_{m}\) and \(K_{i}\) values were calculated by re-fitting the data to a modified Michaelis–Menten equation:

\[ V = V_{max}S/\left( K_m \left( 1 + \left[ I/K_i \right] \right) + S \left( 1 + \left[ I/K_{i0} \right] \right) \right) \]

Affinity selection

The library used for the selection was a kind gift from Dr L. Riechmann (MRC, Cambridge, UK). Briefly, the library was constructed by inserting randomized (NNG/C) oligonucleotides in the region corresponding to CDR3 of prototype cV₇₄E2 between residues 98 and 103 (Brookhaven Protein Data Bank, 1VHP), shown in Figure 2. The library contained a total of ~10¹⁰ transformants having insertions ranging from 5 to 12 residues in length (for details, see: Davies and Riechmann, 1995). Dynabeads M-280 sheep anti-mouse antibody (Dynal) were coated with 90 pmol of anti-myc mAb 9E10 according to the supplier’s instructions. A saturating amount of myc-tagged NS3, 400 pmol in selection buffer (SLB: 15% glycerol, 0.1 mM DTT, 50 mM NaCl, 0.5% CHAPS, 10 mM Tris, pH 7.5,) was used for immobilization on to the anti-myc mAb 9E10. After a 2 h incubation at room temperature, the unbound NS3 was eliminated by five washes (100 μl, 1 min each) with SLB.

The beads were then blocked with 3% (w/v) E.coli bacterial extract in SLB, which was more efficient than non-fat milk or BSA as a scavenger reagent, and washed once again before use. Then ~10¹¹ TU (transducing units) of the cV₇₄E2 library in 100 μl SLB were added to the beads and affinity-selected for 3 h at 23°C in the presence of 3% E.coli bacterial extract and 10¹¹ UV-killed f1 phage particles as a blocking agent. Subsequently, the beads were washed 10 times and resuspended in 100 μl of SLB. The suspension was loaded on to 700 μl of a 30% sucrose cushion and the beads were collected by applying a magnetic field (Dynal, MPC). This step was repeated a second time before acidic elution of the bound phage with 0.1 M Tris–glycine buffer (pH 2.2). Following neutralization of pH and phage amplification in TG1 cells, the selection was repeated four times using the same input of purified phage as in the first round. A depletion step on to the anti-myc mAb 9E10-coated beads was carried out after each enrichment cycle to counter-select for phage interacting with the matrix devoid of protease. As a measure of the efficacy of this counter-selection, we estimate that ~50% of input phage population is recovered after the depletion step.

Expression and purification of cV₇₄H

The V₇₄ encoding expression plasmids were transformed in TG1 E.coli strain bearing the plasmid pDMI1 which expresses LacI (kind gift from Dr Buijard, University of Heidelberg). Cells were grown in fortified SB medium (Power et al., 1992) in a bench-top fermenter (MFS, SAVI, Italy) at 27°C, the stirring rate was set at 500 r.p.m., the pH was kept between 7.0 and 7.5 and the air flow was fixed at 0.5 Nl/min (1/min of nitrogen equivalents). When the culture A₆₀₀ reached 2.0, expression was induced by addition of 500 μM of IPTG and the air flow was increased to 1 Nl/min, culture was continued for 6 h (corresponding to a final A₆₀₀ of 5.0) and cells were collected by centrifugation. The pellet was immediately resuspended in 50 ml (1/20 of the original volume) in 30 mM Tris, pH 7.4, 1 mM EDTA, 20% sucrose and left on ice for 30 min. Cells were centrifuged (20 min at 2500 g) and supernatant 2 was kept. The pellet was resuspended in 50 ml of pre-cooled 5 mM MgSO₄ and incubated for another 30 min at 4°C. The solution was centrifuged and the supernatant was pooled with supernatant 2, filtered and purified using IMAC (Hochuli et al., 1987).

Gel filtration chromatography

Gel filtration experiments were run on a Pharmacia FPLC system, with a Superdex-75 analytical column equilibrated in TBST containing 10% glycerol and 0.1 mM DTT. The calibration run was done with a Pharmacia Biotech LMW gel filtration calibration kit, which includes the following markers: dextran blue, bovine serum albumin, ovalbumin, chymotrypsinogen and RNase corresponding to 200, 67, 43, 25 and 13.7 kDa, respectively.

In vitro translation of NS3 and substrate

DNA fragments derived from HCV-BK strain cDNA were inserted downstream of the 5’ untranslated region of encephalomyocarditis virus and under the T7 promoter in the pCite-1 vector (Novagen) in the appropriate translational reading frame and followed by a termination codon. The plasmids pCiteNS3–4Acut and pCiteNS5ABAC51, expressing the HCV proteins NS3-4A with a mutated cleavage site and NS5AB from residue 1965 to residue 3010, respectively, have been described previously (Steinkuhler et al., 1996a,b; Tomei et al., 1996).
In vitro transcription was done with T7 RNA polymerase (Promega). The transcripts were translated for 1 h at 30°C in the presence of [35S]methionine using RNA-dependent rabbit reticulocyte lysate (Promega). Cleavage of labeled precursor was assayed by SDS–PAGE on 10% gels and exposed on the Phosphorimager from Molecular Dynamics.

Elastase and kallikrein assays

The pancreatic porcine elastase (PPE), its substrate (Me-o-Suc-Ala-Ala-Pro-Val-pNA) and the pancreatic porcine kallikrein (PPK) were purchased from Calbiochem. The PPE substrate Chromozym (p-Pro-Phe-Arg-pNA) was purchased from Boehringer Mannheim. The PPE assay was performed in a 96-well plate, in 50 mM TES, pH 7.5, 500 mM NaCl, 0.05% Triton X-100, 5% DMSO buffer; the enzyme was added to the final concentration of 0.5 µg/ml and substrate to 500 µM. Samples were tested in a 100 µl final volume and after 75 min of incubation at room temperature under constant agitation, the OD was read at 405 nm. The PPK assay was performed in 120 mM Tris, pH 8.0, 0.01% Triton X-100, 5% DMSO; the concentration of the protease and its substrate were 50 nM and 100 µM, respectively. After incubation for 1 h at 25°C under agitation the OD was recorded at 405 nm.

Results

In order to carry out affinity selection of randomized CDR3 variants of cV_{H} domains displayed on filamentous phage, it was fundamental to engineer the target NS3 protease 20 kDa minimal domain that can be expressed in E.coli and purified in large quantities (Steinkuhler et al., 1996a) to allow its immobilization on to a solid-phase matrix in a bioactive conformation. To this end we fused a myc-tag peptide (EFEQKLISQQDLG) to the C-terminus of HCV NS3 domain, cV_{H} of some bona fide NS3 ligands were excised from the RF dsDNA by digestion with XhoI and BstEII and subcloned into pUCV_{H}-myc vector, suitable for expression in the periplasmic space of E.coli (Davies and Riechmann, 1995). Next, the cV_{H} proteins bearing the myc epitope and the (His)_{6} tag genetically fused to their C-termini were purified by IMAC (Hochuli et al., 1987) before testing. As a preliminary screen for inhibition activity, we used the in vitro translation assay in sequence and length (5–12 residues) and displayed as fusion with the pH1 protein of fdCAT (Davies and Riechman, 1995). Four rounds of selection were carried out and subtractive steps were included between rounds. A depletion step consists of incubating the amplified library with the matrix, including all reagents but devoid of NS3 20 kDa fragment. Because NS3 protease requires detergent and kosmotropic agents for optimal activity (Steinkuhler et al., 1996a), affinity-selection steps were carried out for 4 h at room temperature, in the presence of 15% glycerol and 1% CHAPS detergent, under mild reducing conditions (0.1 mM DTT and in a nitrogen atmosphere), to maintain adequate conditions for the enzymatic activity but avoiding disruption of the V_{H} disulfide linkage. Under these buffer conditions the protease is active for 3–4 h, judged adequate to carry out the affinity selection.

Following selection, 42 phage clones from the fourth round were cultured, concentrated by PEG precipitation, their ssDNA extracted and the CDR3 amino acid sequence was determined by DNA sequencing (Table I). Some of these putative NS3 ligands were present as multiple isolates; particularly striking was clone cV_{H}E2 that represented the major component (20/42) of the phage population. Its CDR3, considerably rich in proline residues, shows a suggestive similarity to other isolates, albeit the lengths of their CDR3 differ greatly. Because the library was constructed using randomized oligonucleotides according to the scheme NNG/C, the relatively high percentage of Pro residues observed here is not due to a bias in the repertoire, as suggested also by the sequence of other isolates recovered from a different study (Davies and Riechmann, 1995). We decided to study some of these clones further with the aim of identifying protease inhibitors. To this end, the DNA fragments encoding the variable region of the isolated cV_{H} of some bona fide NS3 ligands were excised from the phage RF dsDNA by digestion with XhoI and BstEII and subcloned into pUCV_{H}-myc vector, suitable for expression in the periplasmic space of E.coli (Davies and Riechmann, 1995). Next, the cV_{H} proteins bearing the myc epitope and the (His)_{6} tag genetically fused to their C-termini were purified by IMAC (Hochuli et al., 1987) before testing. As a preliminary screen for inhibition activity, we used the in vitro translation assay

Fig. 2. NMR structure of cV_{H} prototype molecule. Stereo drawing of the camelized VH-P8 domain NMR structure (Riechmann et al., 1996), PDB entry 1VHP (Bernstein et al., 1977), showing the close spatial proximity of CDR3 (magenta) and residue Ile 47 (green). The figure was drawn using RIBBONS (Carson, 1987).
Fig. 3. In vitro translated 20 kDa NS3 inhibition assay. E.coli expressed and IMAC-purified cVH proteins (reported on the y-axis) were tested at 1.5 µM in triplicate samples for their ability to inhibit in vitro translated 20 kDa NS3 protease by monitoring the percentage conversion (x-axis) of the in vitro translated NS5AB substrate.

Table I. Amino acid sequence of CDR3 from affinity-selected cVH molecules

<table>
<thead>
<tr>
<th>Name</th>
<th>CDR3 sequence</th>
<th>Number of isolates/42</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>EPRIPRPPS</td>
<td>20</td>
</tr>
<tr>
<td>E4</td>
<td>STTLSTSPITK</td>
<td>6</td>
</tr>
<tr>
<td>E5</td>
<td>APHIPRALTPS</td>
<td>1</td>
</tr>
<tr>
<td>E17</td>
<td>PRAPQI</td>
<td>1</td>
</tr>
<tr>
<td>E24</td>
<td>RPRAPOV</td>
<td>4</td>
</tr>
<tr>
<td>M2</td>
<td>PKPQPOY</td>
<td>1</td>
</tr>
<tr>
<td>M3</td>
<td>AQVDSPGSPFT</td>
<td>1</td>
</tr>
<tr>
<td>M4</td>
<td>ADHVEAFLS</td>
<td>1</td>
</tr>
<tr>
<td>M5</td>
<td>SKCSYHPPYTS</td>
<td>1</td>
</tr>
<tr>
<td>M12</td>
<td>SKKSSDSVFP</td>
<td>1</td>
</tr>
<tr>
<td>M17</td>
<td>HSSSPQGTGVT</td>
<td>1</td>
</tr>
<tr>
<td>M20</td>
<td>HPSAKPSST</td>
<td>1</td>
</tr>
<tr>
<td>M22</td>
<td>NAPNEPGLDRS</td>
<td>1</td>
</tr>
<tr>
<td>M23</td>
<td>GYDYGAW</td>
<td>1</td>
</tr>
<tr>
<td>M24</td>
<td>EPRHARARRS</td>
<td>1</td>
</tr>
<tr>
<td>N5a</td>
<td>KSSLPD</td>
<td>1</td>
</tr>
</tbody>
</table>

*aN5 is a non-selected cVH.

(Steinkuhler et al., 1996b) in which the in vitro transcribed RNAs encoding the 20 kDa NS3 fragment and one of the substrates, NS5AB, are incubated in the presence of fixed amounts of cVH proteins. The purified proteins encoded by several constructs were tested at 1.5 µM in triplicate samples and only cVH E2 was revealed as an inhibitor of the protease activity (Figure 3). At this concentration, most of the molecules tested showed little or no activity (0–30% inhibition), within the range of non-specific inhibition given by the non-selected, negative control cVH N5 and therefore were not pursued further. On the other hand, cVH E2 appeared to be a relatively potent inhibitor, deserving additional characterization.

In order to determine the apparent molecular weight of cVH E2 under NS3 protease assay buffer conditions, we carried out gel filtration FPLC in this buffer, except that the glycerol concentration was lowered to 10% (the maximum compatible with chromatography). The protein migrates as a dimeric species with an apparent molecular weight of 25 kDa (Figure 4), whereas other selected and unselected molecules elute as monomeric species, thereby suggesting that the CDR3 sequence of cVH E2 affects the physical state of this molecule (see Discussion). The FPLC elution profile did not change even in the absence of glycerol and CHAPS.

An accurate estimate of the potency of cVH E2 inhibitor was carried out through titration experiments by monitoring the residual enzyme activity as a function of cVH E2 concentration using a small fluorescent substrate at a concentration of 40 µM. It is worth noting that the IC₅₀ of cVH E2 was not affected by the co-factor NS4A and was estimated to be 300 nM (Figure 5). Because cVH E2 activity is not inhibited by NS4A peptide and vice versa, it appears that the selected ligand can bind both forms of the enzyme, both as a free species and as a heterodimer complex. It was also important to establish the reversibility of the inhibition activity, assessed by an assay in which we pre-incubated the enzyme with 1.5 µM concentration of the cVH E2 inhibitor (five times its IC₅₀ value) followed by a 10-fold dilution and determination of the residual enzymatic activity (Figure 6). The results were consistent with reversible inhibition as we recovered more than 80% of the enzyme activity upon dilution of the complex. To establish the mechanism of action of the cVH E2 inhibitor, substrate titration experiments were performed in the absence and in the presence of 125 and 250 nM of cVH E2 (Figure 7). By fitting the experimental data to a modified Michaelis–Menten equation (Equation 2), the dissociation constants of the enzyme inhibitor...
Fig. 5. Determination of cVH E2 IC₅₀. Recombinant NS3 20 kDa fragment protease (20 nM) was incubated in the presence of increasing amounts of purified cVH E2 inhibitor (duplicate), 40 µM substrate and 10 µM NS4A cofactor. Percentage of residual activity (y-axis) was determined as function of inhibitor concentration (x-axis) and the IC₅₀ value derived.

Fig. 7. Determination of the inhibition mechanism. To determine the mechanism of inhibition of cVH E2, substrate titration curves were recorded in the absence and in the presence of two different concentrations of inhibitor: 250 nM (circles), 125 nM (squares), no inhibitor (diamonds), and the values of Kᵢ (enzyme–inhibitor complex dissociation constant) and Kᵢᵢ (ternary enzyme–inhibitor–substrate complex dissociation constant) derived by fitting the experimental data to a modified Michaelis–Menten equation.

selectivity by testing cVH E2 on two commercially available serine protease assays, namely porcine elastase and kallikrein. Neither of these enzymes was inhibited to any measurable extent upon incubation with a concentration of cVH E2 corresponding to a 5–30-fold excess of the inhibitor/substrate ratio (data not shown), indicating a substantial level of selectivity.

Finally, because of the lack of cell-based replication systems for HCV, in order to determine if cVH E2 was active on the entire (70 kDa) NS3 gene product (including the helicase domain) and on the natural substrate, we exploited the in vitro translation assay described previously (Tomei et al., 1993) as a surrogate for a biological assay. In this experiment (Figure 8), cVH E2 inhibits 70 kDa-borne proteolytic activity with an apparent IC₅₀ of ~1 µM, 3.3-fold above the value determined by the activity titration assay with the (13-mer) synthetic substrate.

Discussion

Because of the importance of HCV as a human pathogen and the lack of an effective treatment (Weiland, 1994), there is a crucial need to develop inhibitors of viral metabolism. Proteases are fundamental components of physiological regulation and are involved in the pathophysiology of a number of disease states (Neurath, 1989). They are essential to the life cycle of many parasitic pathogens and to processing viral precursor proteins (Korant, 1988; Krausslich and Wimmer, 1988). These observations have led to an ever increasing interest in proteases as potential therapeutic targets. Previous studies have indicated that the HCV NS3 protease activity is required for processing the non-structural proteins NS3, NS4 and NS5 in addition to the NS4A binding site. We also determined that cVH E2 was not cleaved by NS3 after overnight incubations with high concentrations (30 µM) of NS3 (data not shown), demonstrating that the observed inhibition of NS3 by cVH E2 is not due to cleavage of a competitor substrate. Next, we checked
structural determination of NS3 protease domain (Kim et al., 1996; Love et al., 1996). In particular, it was argued that the absence of some conserved loops, in the NS3 protease, renders its substrate binding site relatively shapeless, thus it was anticipated that the making of substrate-based small molecule inhibitors will be a daunting task (Kim et al., 1996). For these reasons, we anticipate that a deeper understanding of the NS3 surface properties and the availability of biological tools for studying it will have a substantial impact on the development of effective inhibitors. Here we report the affinity selection and characterization of cVH E2, a camelized VH domain (Davies and Riechmann, 1994), that is a relatively potent and selective competitive inhibitor of the NS3 protease activity.

The Camelidae is the only taxonomic family known to possess functional heavy-chain antibodies, devoid of light chains (Hamers-Casterman et al., 1993). VH domains lacking their VL partners have been bacterially expressed and shown to be able to bind antigen specifically (Ward et al., 1989). Camelization of a human VH domain (VH-P1) to mimic camelid heavy chains was achieved by introducing the mutations G44E, L45R, and W47G (Davies and Riechmann, 1994, 1996) that are crucial for reducing the affinity for the cognate VL domain (Chothia et al., 1985). Camelizing mutations also made the structural determination of this cVH by NMR spectroscopy possible (Riechmann, 1996), except that residue 47 was mutated in ileucine, which yields better expression and higher stability compared with the G47 containing VH domains (Davies and Riechmann, 1994; 1996). The solution structure of the cVH showed well defined regions of β-structure and less well defined connecting loops and turns (Riechmann, 1996). The general topology of two pleated β-sheets and the conformation of the hypervariable loops H1 and H2 were shown to be very similar to those of VL-associated VH domains (Figure 2).

Usually, the G47I mutation does not affect antigen binding or specificity (Davies and Riechmann, 1995, 1996); however for cVH E2 we observed a 20-fold reduction in potency which is not dependent on dimerization. The cVH E2 protein forms dimeric species independently from the nature of its residue 47 (G or I). In the light of recent structural data, it is conceivable that in the absence of cognate VL, some CDR3 structures can be affected by the nature of residue 47, which in turn may promote dimerization. In fact, the NMR structure of the parental cVH, which has a shorter CDR3 than cVH E2 (Figure 2), demonstrated that it is a stable monomeric protein (Riechmann, 1996). In addition, that dimerization may be dependent on the nature of CDR3 is also supported by the 1.85 Å X-ray structures of antigen-free llama VH domain (Spinelli et al., 1996) and by the 2.5 Å crystal structure of a camel VH in complex with its antigen, lysozyme, as reported recently (Desmyter et al., 1996). Compared with human and mouse VH domains, there are no major backbone rearrangements in the VH framework, but the architecture of the region of VH that interacts with a VL, in a conventional Fab, is different from any previously seen. The camel antibody has a very long CDR3, one half of which contacts the VH region which in conventional immunoglobulins interacts with a VL, whereas the other half protrudes from the antigen binding site and penetrates into the lysozyme active site.

The CDR3 segment that penetrates deeply into the active site of lysozyme provides most of the binding energy; unfortunately, this is unlikely to be the case for the NS3/cVH E2 interaction, as suggested by the analysis of the X-ray structure of the NS3 specificity pocket (Kim et al., 1996). Nonetheless, cVH E2 may still be a useful tool for guiding the synthesis of CDR3 peptido-mimetics as shown successfully, albeit with an expected loss in potency (Smythe and von Itzstein, 1994). In the absence of structural data, there are two considerations that can be made about the mechanism of inhibition for cVH E2: (i) the inhibitor binds NS3 both as a free enzyme and as a heterodimeric complex with NS4A; (ii) the mechanism of action is that of a competitive inhibitor. It is therefore likely that cVH E2 binds to the enzyme active site or to a nearby epitope, thus hampering substrate entry, but in any case its binding site is far from the NS4A binding site. It should also be emphasized that the inhibitor is not cleaved by NS3, which can be a consequence of the absence of a suitable Cys residue within the CDR3 sequence (Desmyter et al., 1996). Taking these results together, it is tempting to propose that to design candidate molecules (good building blocks) for the selection of more potent NS3 inhibitors, it would be wise to engineer CDR3 constructs with a long CDR3, part of which could be randomized to fit the target enzyme active site cleft whereas an invariant component (which would act as an additional framework) could be designed based on the natural camel antibody (Desmyter et al., 1996), to shield the former VH interface, thus avoiding dimerization. In this respect it may also be relevant to point out that by using a minibody scaffold (Pessi et al., 1993) having shorter CDRs (six residues) for the selection of NS3 inhibitors, we were able to recover a ligand of comparable potency but showing a non-competitive inhibition

This study confirms previous findings with selected cVH ligands specific for hapten, protein and peptide whose affinities are in the micro- to nanomolar range (Davies and Riechmann, 1995). On the other hand, by displaying on filamentous phage variants of natural proteinaceous inhibitors such as BPTI (Roberts et al., 1992), Ecotin (Wang et al., 1995), APPi (Dennis and Lazarus, 1994), hPSTI (Rottgen and Collins, 1995), LACI-D1 (Markland et al., 1996), protease inhibitors have been remodelled and new specificities selected with Ks as the target has yielded competitive inhibitors of modest (micromolar K) potency (N.Dimasi, F.Martin, C.Volpari, N.Brunetti, G.Biasiol, S.Altamura, R.Cortese, R.De Francesco, C.Steinkhuler and M.Sollazzo, in preparation). Probably, more potent cVH inhibitors will be achieved by a stepwise approach of affinity maturation, since the complexity of the repertoire explored (10^10–10^12) represents only a subset of all the possible combinations of a 6–12 amino acid long CDR3. In order to improve the potency of our macromolecular lead, cVH E2, it is conceivable to identify by Ala scanning the residues within its CDR3 that are crucial for activity and to optimize them by subsequent cycles of randomization/selection (possibly in the context of a 147 framework format to ensure higher levels of expression). Once a more potent compound is obtained, it may be possible to proceed with the synthesis of the smallest cyclic peptide that conserves significant levels of bioactivity.

These easy to generate and engineer minimal antigen recognition units serve as potential pharmacophoric models to drive low molecular weight compound design (reviewed by Sollazzo et al., 1995) and in addition provide useful biological probes in cell based assays and animal models to validate NS3 and Roberts, B.L., Markland, W., Ley, A.C., Kent, R.B., White, D.W., Guterman, S.K. and Ladner, R.C. (1996b) Proc. Natl Acad. Sci. USA, 87, 9524–9528.


Received November 27, 1996; revised January 24, 1997; accepted January 28, 1997.