JMB

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

Daniel O. Cicero, Gaetano Barbato, Uwe Koch, Paolo Ingallinella, Elisabetta Bianchi, M. Chiara Nardi, Christian Steinkühler, Riccardo Cortese, Victor Matassa, Raffaele De Francesco, Antonello Pessi and Renzo Bazzo*

Istituto di Ricerche di Biologia Molecolare P. Angeletti (IRBM), via Pontina Km 30.600, 00040 Pomezia (Rome) Italy

The interactions of peptide inhibitors, obtained by the optimization of N-terminal cleavage products of natural substrates, with the protease of human hepatitis C virus (HCV) are characterized by NMR and modelling studies. The S-binding region of the enzyme and the bound conformation of the ligands are experimentally determined. The NMR data are then used as the experimental basis for modelling studies of the structure of the complex. The S-binding region involves the loop connecting strands E2 and F2, and appears shallow and solvent-exposed. The ligand binds in an extended conformation, forming an antiparallel β-sheet with strand E2 of the protein, with the P1 carboxylate group in the oxyanion hole.

© 1999 Academic Press

*Corresponding author

Introduction

The hepatitis C virus (HCV) is an enveloped virus formed by a positive-sense RNA genome of 9.6 kb with a single open reading frame encoding a precursor polyprotein of about 3000 amino acid residues. The polyprotein is processed by the action of cellular and virally encoded proteases. Two different proteolytic activities encoded by HCV NS2 and NS3 proteins are responsible for the processing of the non-structural proteins of the virus. NS3 is a bifunctional protein. It exhibits a proteinase activity within the N-terminal 180 amino acid residues, while the remaining two-thirds of the protein shows a helicase activity (Gallinari *et al.*, 1998). The proteinase activity is required to process the precursor polyprotein at four site-specific junctions, NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B (for reviews, see Bartenschlager, 1997; Neddermann *et al.*, 1997). In order to perform its physiological task, the action of a cofactor, NS4A, enhances the activity of NS3 protease in all the cleavages, via the formation of an NS3/NS4A complex.

Crystallographic and NMR studies of the enzyme in its free form (Love *et al.*, 1996; Barbato *et al.*, 1999) or complexed with its cofactor (Kim *et al.*, 1996; Yan *et al.*, 1988) have elucidated that it belongs to the chymotrypsin family of serine proteinases. Relevant differences have been found at the N-terminal β-barrel, where the interaction with the activating NS4A peptide takes place. Different hypothesis regarding its mechanism of activation have been made based on the crystallographic structures (Love *et al.*, 1998), on biochemical data (Landro *et al.*, 1997) and on the NMR structure (Barbato *et al.*, 1999). However, although a great harvest of structural information is presently available, there is still debate on important details of the activation mechanism and the different roles played in the process by several agents (Barbato *et al.*, 1999). The structural studies have demonstrated that the C-terminal β-barrel conformation is relevant for the identification and the binding of the substrate (P-side), whereas it is invariant with

---

respect to the absence or the presence of the activating cofactor NS4A peptide.

In the absence of a broadly effective cure for hepatitis C, much effort is devoted to the discovery of inhibitors of the NS3 protease. For the drug discovery effort, it is particularly relevant to obtain experimental information about the structure of the complexes substrate/enzyme and inhibitor/enzyme, since there is no experimental evidence for these aspects to date.

It has been observed that the minimum substrate length for optimal recognition and cleavage efficiency is a decapeptide, of which six residues span the P-side and four the P′-side (Landro *et al.*, 1997; Urbani *et al.*, 1997). Moreover, very recently the N-terminal cleavage products of substrate peptides (P-side) have been observed to act as inhibitors of the NS3 protease (Steinkühler *et al.*, 1998; Llinas-Brunet *et al.*, 1998). Starting from the NS4A/NS4B P-side cleavage site amino acid sequence (Asp-Glu-Glu-Met-Glu-Cys), which is the most potent natural product inhibitor, peptide inhibitors have been selected by sequential optimization of the different position of the natural cleavage product (Ingallinella *et al.*, 1998). The peptides examined in this work can be considered as a representative group of acid product inhibitors, spanning the P-side, with a free carboxyl group in P1, which has been shown (Steinkühler *et al.*, 1998; Ingallinella *et al.*, 1998) to be essential for optimum binding activity. The hexapeptide Ac-Asp-Glu-Dif-Glu-Cha-Cys-OH (for abbreviations, see Table 1) turned out to be one of the most potent and for this reason it has been assumed as our structural probe to investigate the S-binding region of NS3 and to identify the range of interactions that are exploited by the ligand.

Since the complex between NS3 and NS4A is in an intermediate-rate exchange regime between free and bound forms under the NMR experimental conditions, the NMR study is performed in the absence of the NS4A peptide. The absence of the NS4A cofactor in our case does not invalidate our observations, since it has been clearly shown from the crystallographic (Love *et al.*, 1996; Kim *et al.*, 1996; Yan *et al.*, 1998) and the solution structures (Barbato *et al.*, 1999) that the C-terminal β-barrel, which is the domain where the interaction with the P-side of the substrate takes place, is similar to the crystallographic

structure obtained in presence of NS4A (r.m.s.d. 0.56 Å). Moreover, biochemical (Landro *et al.*, 1997) and structural data (Barbato *et al.*, 1999) suggest that the action of NS4A is exerted on the leaving group of the substrate (P′-side), thereby leaving the C-terminal β-barrel virtually unaffected.

A complete and direct experimental determination of the three-dimensional structure in solution of the inhibitor-NS3 complex turned out to be impossible, due to the short lifetime of the complexes and the limited stability of the samples. Therefore, the structure characterization has involved application of modelling techniques to the experimental information obtained by NMR. The analysis of the ligand-binding sites has been made possible by monitoring the changes induced in NS3 spectra upon binding the different ligands. For this purpose, we have considered as ligands a series of structurally related peptides that differ by the deletion of one or more residues to allow the elucidation of the specific role played by each amino acid residue in the binding with the protease.

The ligand-bound conformation has been obtained by transfer NOE experiments. These observations, combined with the solution structure of free NS3 (Barbato *et al.*, 1999), allowed modelling of the three-dimensional structure of the inhibitor-NS3 complex.

This characterization shows that the S-binding region comprises the E2 and F2 strands, entirely located in the C-terminal β-barrel of the NS3 protein. This region appears to be rather shallow, relatively featureless and solvent-exposed, with the peptide picking up several interactions in different parts of the protein in an additive fashion. The ligand, similar to several other proteinases belonging to the same family (Marquart *et al.*, 1983; Bode *et al.*, 1987; Fujinaga *et al.*, 1987) is in an extended conformation (transfer NOE data), and lies across the surface of the protein with the P1 side-chain protruding into the crevice between the two β-barrels where the elements of the catalytic triad are positioned and where the free carboxyl group makes the interactions that are most relevant to stabilize the binding. Optimum binding takes place for the hexapeptide, which has been proposed from S.A.R. studies to exploit two ''anchors'' (Ingallinella *et al.*, 1998). The P1 anchor is the one just described, involving the carboxylate group of P1, which is the more relevant in terms of binding energy. The second is the P5-P6 anchor, which is of an acidic nature and gives rise to an electrostatic interaction with the enzyme that is important for binding. Our data provide a structural basis for the existence of this dual anchoring mechanism. The peptide-binding mode is unaltered by the absence of one of the two anchors, although with a great loss in affinity.

Table 1. Inhibitors of NS3 protease

Inhibitor[a]		IC50 (µM)
Ac-Asp-Glu-Dif-Glu-Cha-Cys-OH	**1**	0.05
Ac-Glu-Dif-Glu-Cha-Cys-OH	**2**	1.4
Ac-Dif-Glu-Cha-Cys-OH	**3**	30
Ac-Glu-Cha-Cys-OH	**4**	460
Ac-Asp-Glu-Dif-Ile-Cha-OH	**5**	46

[a] Abbreviations used: Dif, 3,3-diphenylalanine; Cha, β-cyclohexylalanine.

Results and Discussion

Characterization of the S-binding region of NS3 protease

The binding of peptide inhibitors was followed by observation of ^{15}N and ^{1}H-amide chemical shift changes in two-dimensional ^{15}N-heteronuclear single-quantum correlation (^{15}N-HSQC) spectra upon addition of a ligand to a ^{15}N-labelled NS3 sample. In a standard experiment, a titration of the enzyme with the inhibitor is performed, recording an ^{15}N-^{1}H 2D experiment after each addition of the peptide. In the case of fast exchange relative to the NMR time-scale between the bound and free forms, it is possible to determine the final position of the peaks in the complex, using the assignment of the free enzyme as the starting point and following the displacement of the peaks after each addition.

Table 1 shows the peptides used in the present investigation. The list includes the P-optimized (Ingallinella et al., 1998) hexapeptide 1, pentapeptide 2 lacking P6, tetrapeptide 3 lacking both P5 and P6 and tripeptide 4, spanning from P1 to P3. Additionally, pentapeptide 5, lacking P1, was also investigated.

The great sensitivity of the chemical shift to local nuclear environment is well established (Chen et al., 1993). The magnetic shielding of the nuclear environment, including effects due to peptide bond anisotropy, electrostatic interactions, aromatic ring currents and hydrogen bonding, can be regarded as the main determinant of the chemical shift of a nuclear spin in a protein. The formation of the complex with a ligand will inevitably cause changes in the environments of nuclei in amino acid residues at the interface, resulting in chemical shift changes. It cannot be excluded that these effects can extend beyond the direct contact area, as determined in the case of the cyclosporin A/cyclophilin complex (Spitzfaden et al., 1992). For this reason, a consistent and comparative analysis of the whole set of observed chemical shifts is needed to establish a correct definition of the area of the protein involved in direct interactions with the ligand and of regions indirectly perturbed due to conformational rearrangements.

In order to represent the change in chemical shift of the whole NH group, we used a single quantity, Δ, that averages the effects on the H and N chemical shift (Grzesiek et al., 1996). This index is calculated using the following expression:

$$\Delta = [(\Delta\delta_{NH}^2 + \Delta\delta_N^2/25)/2]^{1/2} \qquad (1)$$

where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the differences in chemical shift between the free and complexed enzyme for the H and N, respectively. In Figure 1, we report a small region of the series of proton-nitrogen correlation spectra obtained in the titration of the enzyme with hexapeptide 1. The resonance

shifts induced by the complex formation are indicated for Arg123 and Ala164.

In Figure 2(a) we illustrate the changes in chemical shifts between the complex formed by hexapeptide 1, NS3 and the free enzyme.

Two well-defined regions show significant changes: the first around the catalytic serine residue (Lys136 to Gly141) and the second from Arg155 to Val170. Two other positively charged residues, Arg123 and Arg109, experienced shifts. The large shifts observed in the NH groups near the catalytic serine residue (Ser139) clearly indicate the occupancy of the active site. These two regions define a contact surface between the peptide and the protein, the S-binding region, which is depicted in Figure 3. According to the structure determined for NS3 in solution (Barbato et al., 1999), the S region involves mainly the strands E2 and F2, and the loop connecting them. It defines a rather flat and positively charged surface for the interaction with the P-region of the substrate.

Additionally, the regions around the other two members of the catalytic triad of NS3 (His57 and Asp81) also experienced significant shifts, which is a clear evidence of a perturbation of the conformation around the catalytic residues. This observation can be regarded as a supportive evidence for the hypothesis that the alignment of the catalytic triad may be influenced by the occupancy of the S-region of the enzyme (Barbato et al., 1999).

As already proposed (Ingallinella et al., 1998), optimal binding of this type of inhibitor seems to require a dual anchor: an acid anchor constituted by residues P6 and P5, and a P1 anchor. We tried to establish the region of the backbone of the enzyme involved in the recognition of these three important residues of the hexapeptide inhibitor. To this purpose, hexapeptide 1 was taken as a parent compound and the complexes with a series of structurally related peptides differing one from the other by the deletion of one or more residues were examined in a comparative fashion.

To perform this comparison, the observed chemical shifts are to be normalized at 100% complex formation. This situation can be difficult to achieve for less potent compounds, for which a large ligand excess can induce precipitation of the enzyme. However, in the case of complexes formed by related ligands, it is reasonable to assume that there will be a number of common effects. Under this assumption, we have developed a computational algorithm to extrapolate the shifts to enzyme saturation (see Methods).

Figure 2(b) shows the comparison between complexes formed by hexapeptide 1/NS3 and pentapeptide 2/NS3. In this case, the difference is measured using the same definition for Δ, but calculated using the difference in chemical shift between the two complexes. The deletion of residue P6 seems to have only a marginal impact on the backbone conformation of the complex. There is only a small change in Thr160, which is the residue preceding the positively charged Arg161 for which we lack
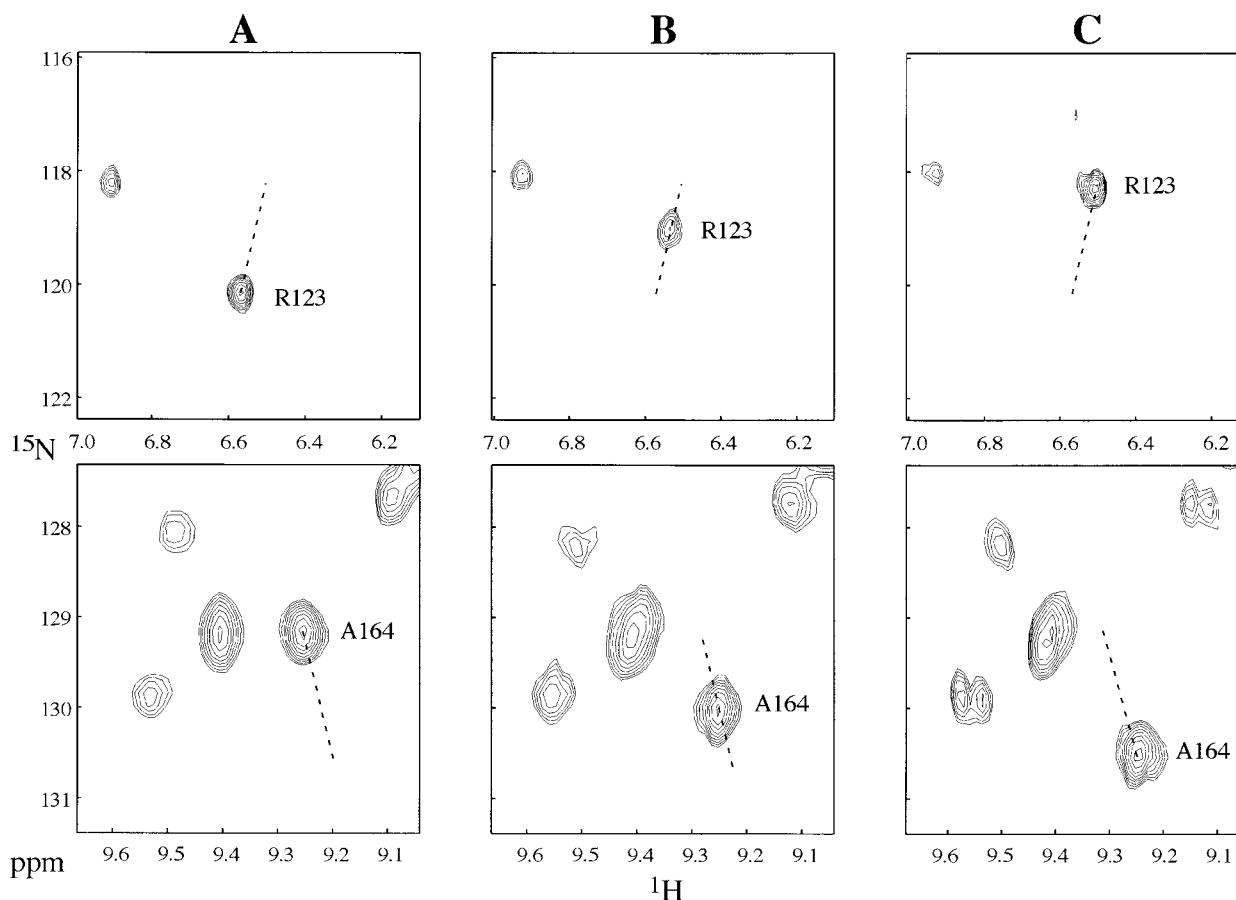
Figure 1. Selected region of the series of proton-nitrogen correlation spectra of NS3 protease titrated with peptide 1. The resonance shifts induced by the ligand binding are indicated. (a) Free NS3; (b) partially saturated (45%) NS3; (c) saturated NS3.

the NH assignment. On this basis, we can only hypothesize that the presence of P6 alters the conformation of the region Thr160-Arg161. However, the overall similarity between these two complexes clearly indicates that the absence of P6 does not change the binding mode.

The impact of P5 deletion is sketched in Figure 2(c). Here, we compare the shifts of the amide group between complexes formed by pentapeptide 2/NS3 and tetrapeptide 3/NS3. A significant shift for Gly162 can be noticed, along with some differences presumably due to conformational adjustments in the region around this residue. As observed in the case of P6 deletion, the overall changes in backbone amide groups of NS3 upon complexation with tetrapeptide 4 are not very marked, except for Gly162. This will again imply that the complete absence of one of the two anchors (the P6-P5 anchor), although reducing the activity 1700-fold (Ingallinella *et al.*, 1998), does not prevent the positioning of P1 in the S1 region, as observed for hexapeptide 1 and pentapeptide 2. The shift observed for Gly162 affects mainly the proton of the NH group, and it is a marked downfield shift (around 0.8 ppm). This observed shift suggests the plausible formation of a new H-bond upon binding of the inhibitor.

The effect of the P4 deletion could not be established, because tripeptide 4 presented no specific interaction with NS3, indicating that at least in the present series of optimised inhibitors, the absence of the P6-P5 anchor cannot be counterbalanced by less than four residues.

The influence of deletion of P1 on the backbone amide shifts of NS3 was then examined. It was earlier recognized that the main contribution to the binding energy comes from the P1 residue (Steinkühler *et al.*, 1998), as three orders of magnitude reduction in activity was observed by deletion of P1. Notably, results with shortened analogues showed that a free carboxyl group was a better C-terminal group than a carboxyamide group. This result opened a new possibility for the binding of pentapeptide 5: it could be the case that the entire binding mode of the peptide is shifted down by one residue, with the former P2 residue binding now in the S1 pocket, explaining the preference for the carboxylate group. However, our results clearly demonstrate that this is not the case. A comparison between the shifts of backbone amide groups of the complex between pentapeptide 5, NS3 and the free enzyme is depicted in Figure 2(d) . The region around the catalytic serine residue is affected less significantly by the binding than in the case of
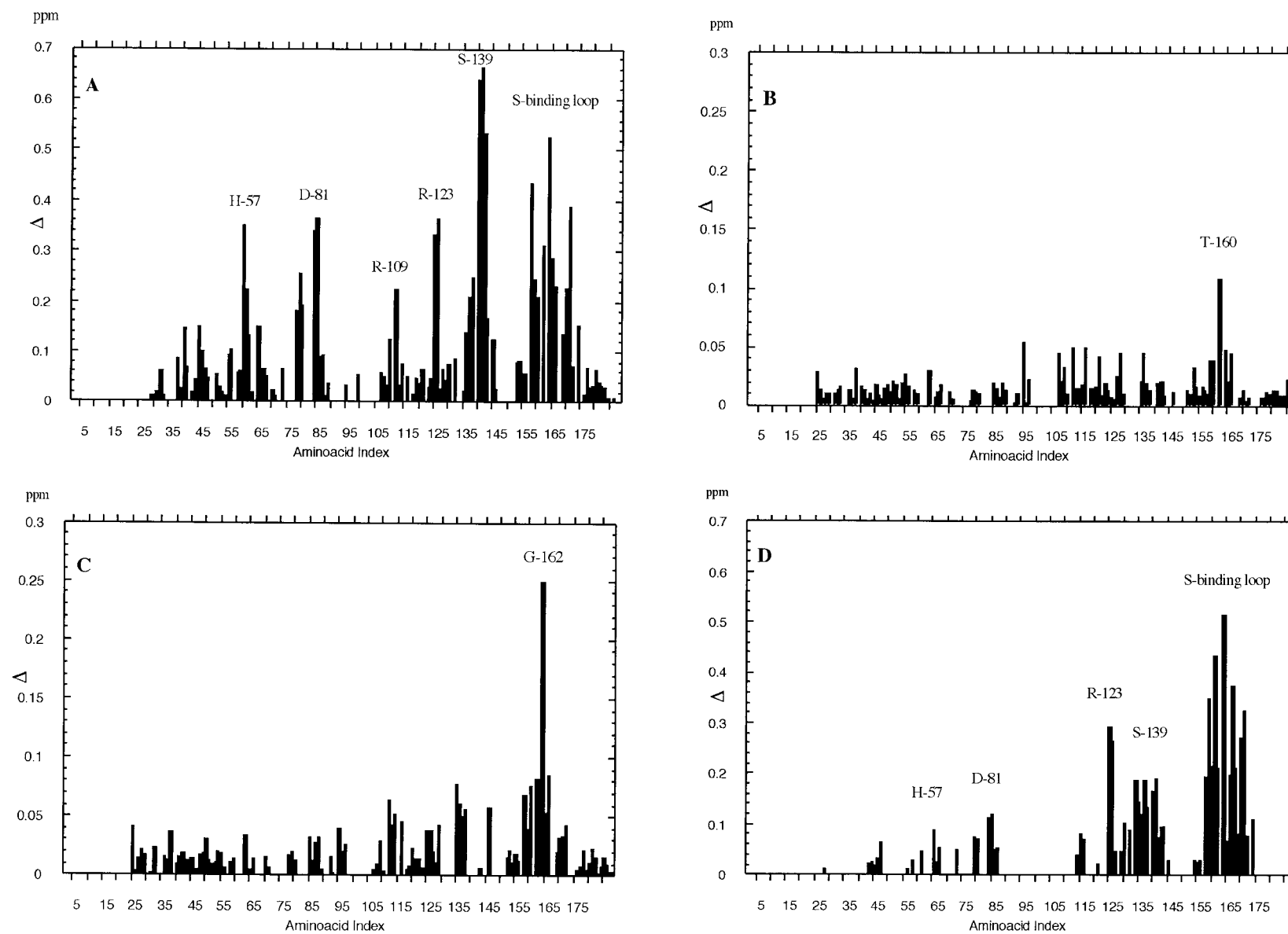
Figure 2. NH chemical shift differences (Δ) between: (a) the complex NS3/1 and free NS3; (b) complexes NS3/1 and NS3/2; (c) complexes NS3/2 and NS3/3; (d) complex NS3/5 and free NS3. The line at $\Delta = 0.1$ is assumed to represent the level of significant difference.
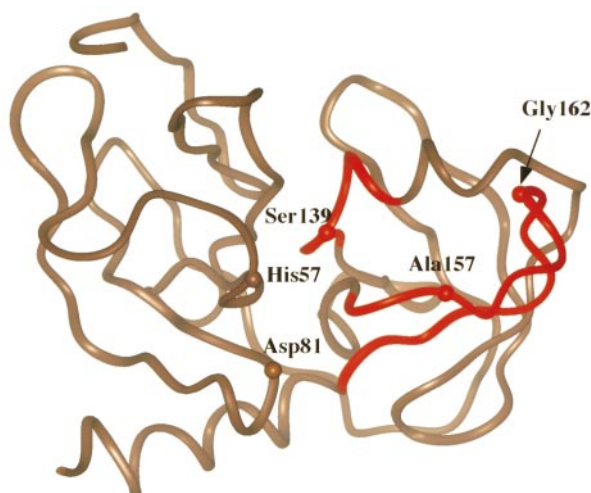
Figure 3. The structure in solution of NS3. The S-binding region, as determined by mapping the ligand binding on the protein backbone, is indicated in red.

complex formation with hexapeptide 1. On the other hand, the region from Ala156 to Val170 shows the most important effect. A comparison of this region between the complex hexapeptide 1/NS3 and pentapeptide 5/NS3 (data not shown) reveals a high degree of similarity for the chemical shifts of the amide NH group. These results imply that the residues are occupying the region from S6 to S2, and that the absence of the P1 anchor does not change the binding mode of the inhibitor.

Residues around Asp81 and His57 are not perturbed by complex formation with pentapeptide 5. This result suggests that the occupancy of the S-1 pocket, or the presence of the carboxylate group in the oxyanion cavity, is necessary to affect the relative position of the catalytic triad.

Bound peptide inhibitor structure and modelled interaction with NS3

The affinity of the peptide inhibitors for the enzyme strongly depends on the ionic strength (Ingallinella *et al.*, 1998). Whereas it has not been possible to achieve sample conditions for NMR where protease-inhibitor complexes could exhibit the long average lifetimes ($\tau > 10$ ms) that would be required to directly address the structure in solution of the complex, upon increasing ionic strength (>100 mM salt) the compounds in solution were in relatively fast exchange ($\tau < 1$ ms) between the free and bound forms. In this situation, data on the bound conformation of the inhibitor peptides could be obtained by transfer NOE (TrNOE) techniques (Ni, 1994). A detailed analysis of the bound structure of the most potent peptide of the series (hexapeptide 1) was conducted based on the transfer NOE data (Figure 4(a)) and compared with data obtained for the free peptide (Figure 4(b)).

A comparison of Figure 4(a) and (b) indicates the appearance of new cross-peaks, especially interactions between side-chains and the consecutive NH groups, and a dramatic increase of the intensity of the cross-peaks already present in the free peptide. The NOESY cross-peak volumes were converted into interproton distances, using the P3 Glu NH to P4 Dif Dif-αH cross-peak to correlate volume with distances. These data were used as distance restraints to generate by molecular modelling a starting bound structure, which was subsequently docked into the active site of the protease. Figure 5 shows a model of the interaction of hexapeptide 1 and NS3 based both on the mapping and the trNOE data.

Only one NH-NH contact was readily observed in the TrNOE spectra of a solution containing hexapeptide 1 and NS3, corresponding to a short distance between the NH groups of P6 and P5. The low intensities of NH to NH contacts suggests an extended conformation of the backbone. To measure the relative population of the β-region of the ϕ, ψ space, the R value (R is the intensity of $d_{NN}(i, i + 1)$/the intensity of $d_{\alpha N}(i, i + 1)$] introduced by Waltho *et al.* (1993) can be used. A low value of R is indicative of a β conformation, whereas values for R above 1.0 are indicative of a more populated α-helical type conformation. The R values calculated for hexapeptide 1 in the bound state are indicated in Table 2. These values clearly indicate a rather extended conformation for the backbone from P2 to P5. Additional evidence for this extended conformation came from the observation of a short distance ($\sim$3.0 Å) between the meta protons of one of the two phenyl rings of P4, and the β-protons of P6. Such a short distance is possible only if a β-like conformation is present from the ψ angle of P6 to the ϕ angle of P4.

A different situation was observed for P6, for which a higher value of R was measured (1.06). This value is typical of ϕ, ψ dihedral angles in the α-region. However, we cannot exclude the possibility that at least the ϕ angle is sampling a set of conformations. This residual mobility in the bound conformation is supported by the observation that the difference in chemical shift between the bound and free peptide is rather low (data not shown), suggesting a solvent exposure of the NH-Ac group. These observations are in line with the finding that the complete replacement of the AcNH group by H has little influence on the inhibitory potency of the compound (data not shown).

According to what is discussed above, hexapeptide 1 is very likely to bind NS3 by forming a β-sheet antiparallel to the strand E2 (Figure 6). In this model, P3 and P5 amino acid residues form hydrogen bonds to the corresponding S3 (Ala157) and S5 (Cys159) amino acid residues. At this point, a measurement of the reduction of the exchange rates for the NH groups of Ala157 and Cys159 due to the ligand or a direct observation of saturation transfer would represent a
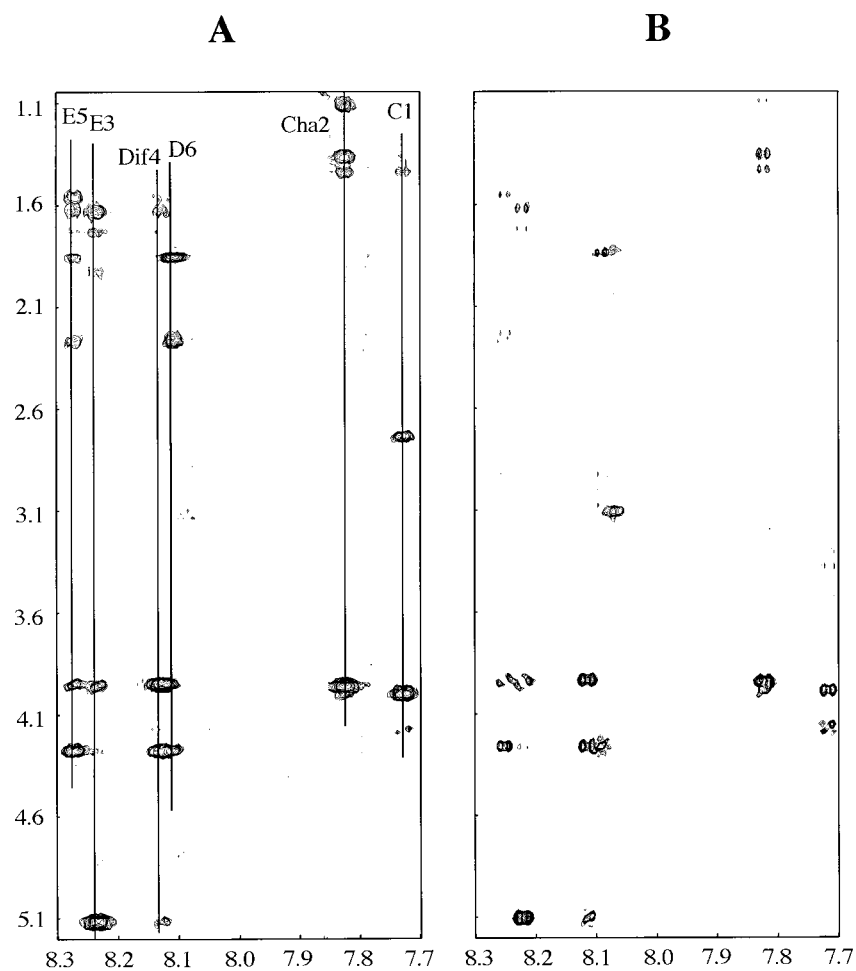
Figure 4. Comparison between (a) the Transfer NOE spectrum on complex NS3/1 and (b) the NOESY spectrum of free peptide 1.

direct proof of the predicted structure. This type of experiment requires a percentage of ligand-bound enzyme relative to the total amount sufficient to give a measurable average effect. In our system, this turned out to be non-feasible in practice, due to the very limited time-stability of a sample of NS3 in presence of an excess of hexapeptide 1. The only measurable effect has remained the frequency shift of NMR signals. However, the proposed hydrogen-bond network is in close agreement with the effect of N-methylation on the activity of these product inhibitors (data not shown). The positions where abolishment of the secondary amide group is most det-
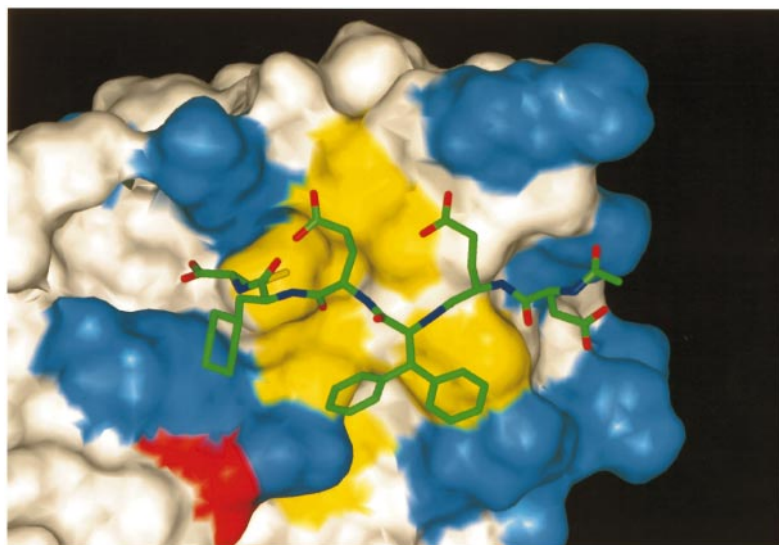

Figure 5. Model of the interaction of hexapeptide 1 and NS3 based on NMR data. Blue represents basic amino acids, red acidic amino acids and yellow hydrophobic amino acids.

Table 2. *R* values determined for the bound structure of hexapeptide 1

P2	P3	P4	P5	P6
0.07	0.04	0.14	0.18	1.06

R is the intensity of $d_{NN}(i, i + 1)$/the intensity of $d_{\alpha N}(i, i + 1)$.

rimental for activity is in fact P3, with an 80-fold increase in the IC_{50}, and P5, with a tenfold increase, whereas no effect is observed on P2 or P4.

The S1 specificity pocket of the NS3 protease is small and lipophilic, comprising Leu135, Phe154 and Ala157. The shape and volume of the relatively small and lipophilic cysteine side-chain is complementary to the pocket. The positively polarised SH hydrogen atom could interact favourably with the π-cloud of the aromatic ring or the sulphur lone pairs can form hydrogen bond-like interactions with the aromatic C-H. Experimental NOEs suggest a *g*⁻ or *t*-conformation of the cysteine side-chain. With respect to the NMR structure of HCV protease, the *g*⁻ conformation seems to be more plausible, since it places the sulfhydryl group into a more hydrophobic environment where the S atom is also close to the edge of the Phe154 side-chain potentially engaged in favourable C—H···S interactions. Similar interactions have been observed in a number of protein and small molecule crystal structures (Burley & Petsko, 1988; Viguera & Serrano, 1995).

The P1 carboxylate group is modelled into the oxyanion cavity forming hydrogen bonds to the backbone amide groups of Ser139 and Gly137 as well as His57 (Figure 7). This mode of binding is analogous to product inhibitors observed crystallographically for other serine proteases (Choi *et al.*, 1991; James *et al.*, 1980; Martin *et al.*, 1992; Nienaber *et al.*, 1993, 1996, Tong *et al.*, 1993). Moreover, resonance shifts of residues around Arg109

and Lys136 suggest that these amino acid residues interact with the P1 carboxylate group. Arg109, and especially Lys136, can stabilise the negatively charged P1 carboxylate group by electrostatic interactions. The importance of Lys136 was verified experimentally by pK_a titration and mutagenesis (Steinkühler *et al.*, 1998). The presence of such basic residues close to the oxyanion hole clearly contributes to the efficiency of product inhibition observed for HCV protease which is unusual for serine proteases.

Interestingly, the S2 site is formed mainly by the His57/Asp81 ion pair (Figure 8). This is particularly relevant, since the loop comprising residues 78-82 contains one member of the catalytic triad (Asp81) that appears strongly solvent-exposed in the free enzyme. In particular, the solvent exposure of the carboxylate group of Asp81 in the free enzyme limits its capability to form efficiently H-bonding with the His57 ring, essential to activate the catalytic machinery. Our results show that upon binding, peptide inhibitors are capable of shielding the dyad Asp81-His57 from the solvent *via* the side-chain of residue P2.

Very recently, a similar example of this type of effect was described. It was observed in the crystal structure of the human cytomegalovirus (HCMV) protease complexed with a peptidomimetic inhibitor based on the natural substrates. This inhibitor spans the P4 to P1′ positions (Tong *et al.*, 1998) and, in this case, the S2 pocket is formed mainly by the catalytic histidine residue (His63). The P2 side-chain lies over the ring of His63 at a distance of ~3.5 Å. Therefore, the shielding effect on the S2 site is analogous to the one we have just described for our enzyme.

This observation seems to be, in general, related to the occupancy of the S-region of the enzyme and therefore it can be extended to natural substrates, reinforcing our proposal for an active role of the substrate in the correct positioning of the
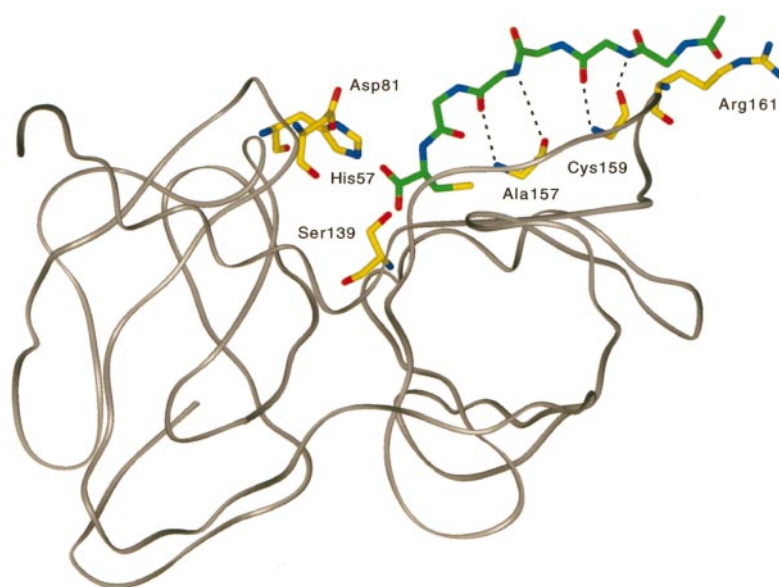

Figure 6. Ribbon representation of the backbone of the lowest-energy NMR-derived solution structure of NS3 complexed with product inhibitor peptide 1. Hydrogen bonds are indicated by broken lines. The carbon atoms of the amino acid residues of HCV-protease forming antiparallel β-sheet hydrogen bonds to P3 and P5 as well as the residues of the catalytic triad are represented in yellow.
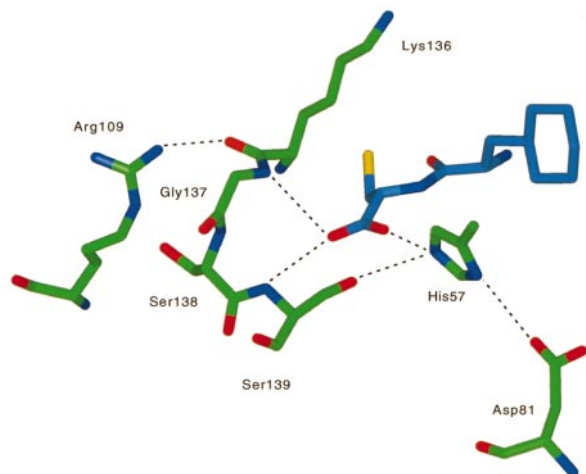
Figure 7. Representation of the active site and the amino acid residues involved in the binding of the P1 carboxylate group. Hydrogen bonds are indicated by broken lines. The carbon atoms of the residues of the protein are represented in green, those of the substrate inhibitor in blue.

Figure 8. The solvent-accessible part of the catalytic ion-pair is represented as a CPK model and the rest of the protease by its solvent-accessible surface (white). A product inhibitor with the P2 Cha side-chains is shown as a stick model.

catalytic triad in this type of complex (Barbato *et al.*, 1999).

One of the more relevant questions arising from the result of the S.A.R. studies described for product inhibitors (Ingallinella *et al.*, 1998) is that of the origin for the observed preference of the bulky diphenylalanine residue in P4. A close inspection of the P4 interaction site reveals a solvent-exposed hydrophobic patch created by the Val158 side-chain. In the crystal and solution structures of HCV protease also, the side-chain of Arg123 is close. Both amino acid residues create a small hydrophobic area that should favour small hydrophobic residues. In order to accommodate the P4 side-chain, a rearrangement of the side-chain of Arg123 is necessary. That would explain the observed shifts in the ^{1}H-^{15}N spectra upon binding of hexapeptide 1. However, there is no specific contact between the diphenyl rings of P4 and NS3 that would suggest a plausible explanation for the increased inhibitory potency of a peptide containing Dif in P4.

Our NMR data on free hexapeptide 1 provide an alternative explanation to this observation. The conformation of the backbone in the region connecting P3 and P4 is extended also for the free peptide, as can be inferred from the intense cross-peak connecting the H$^\alpha$ of Dif and the NH of P3. This conformational preference is very likely due to steric hindrance introduced by the diphenyl group of P4. Moreover, the side-chain conformation of the diphenylalanine in the free state is fixed by steric hindrance of the two aromatic rings. A large coupling constant (11 Hz) is observed between the α and the β protons, indicating a χ1-type conformer in which these two protons are *anti*. The NOE
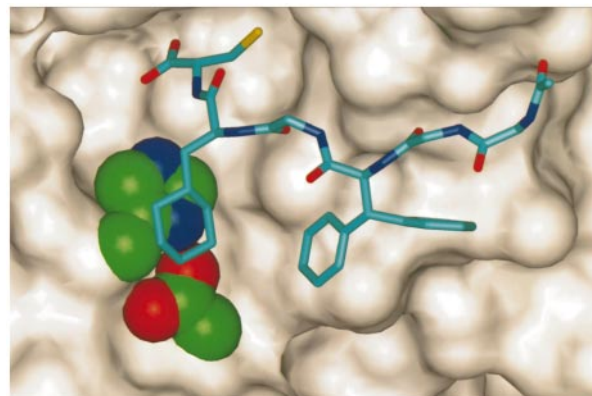
intensities between the β proton and both the α and the NH protons are very similar in the free and in the bound form, suggesting that the side-chain conformation does not change upon binding. These two observations strongly suggest that the preorganization of the conformation of free hexapeptide 1 can lead to an entropic benefit to the resulting binding affinity.

As pointed out above, binding of this type of product inhibitors requires two anchors, P1 and P5-P6. The deletion of the second anchor of the product inhibitors yielded a >100-fold decrease in activity. Since this interaction is electrostatic in nature, it does not depend on the exact nature of the negatively charged moiety. The comparative NMR study previously illustrated shows that a significant shift is induced on the amide resonance of Gly162 by the presence of the Glu residue in P5. In this respect our modelling clearly shows that the P5-Glu side-chain is favourably positioned to form a hydrogen bond to Gly162 and to give rise to a charge-charge interaction with Arg161 and Lys165. Inspection of the surface of NS3 shows that in addition to Arg161 and Lys165 there is a set of basic amino acid residues in the S5-S6 region without nearby neutralising residues, which should further strengthen a favourable electrostatic environment (Figure 5). An acidic P6 amino acid residue seems to profit from the same electrostatically favourable environment. Since this positive electrostatic potential is generated by six basic residues, individual charge/charge interactions are probably not important in this area. The entropic penalty associated with such a rather unspecific interaction due to the reduction of conformational flexibility and desolvation will be reduced compared to an individual charge-charge interaction.

Methods

General

The ^{15}N-labelled NS3 protease was prepared as described in the accompanying paper (Barbato *et al.*, 1999).

Peptide synthesis

Protected amino acids were commercially available from Novabiochem (Läufelfingen), Bachem (Bubendorf), Neosystem (Strasburg) or Synthetech (Albany); peptide synthesis was performed, using DMF as solvent, by Fmoc/t-Bu chemistry (Atherton & Sheppard, 1989) on Tentagel® resin. Individual peptide sequences were assembled on a Millipore 9050 Plus synthesiser, using PyBOP®/HOBt/DIEA (1:1:2, by vol.) activation, fivefold excess of acylants over the resin amino groups and a coupling time of 30 minutes. The peptides were cleaved with 88 % TFA, 5 % phenol, 2 % triisopropylsilane, 5 % water (Solé & Barany, 1992). The peptides were purified by reversed phase-HPLC on a Nucleosyl C-18 column, (250 mm × 21 mm, 100 Å, 7 μm) using water, 0.1 % TFA and acetonitrile, 0.1 % TFA as eluents. Analytical HPLC was performed on a Ultrasphere C-18 column (250 mm × 4.6 mm, 80 Å, 5 μm: Beckman). Purifed (≥98 %) peptides were characterised by mass spectrometry, ^{1}H-NMR and amino acid analysis.

NMR experiments

The sample used to monitor the complex formation by NH chemical shift changes was prepared by adding aliquots of a concentrated solution of the inhibitor (10 mM) in water to a solution of the ^{15}N-labelled enzyme (0.7 mM) in an aqueous buffer composed of 4 % perdeuterated glycerol, 0.1 mM Chaps, 10 mM phosphate buffer. The pH of both solutions was adjusted to 6.2.

The sample for the TrNOE studies was prepared in the same conditions, using a lower enzyme concentration (0.2 mM) and adding a large excess of hexapeptide 1 (1.5 mM). To improve sample stability and exchange conditions, NaCl was added (150 mM).

A sample of free hexapeptide 1 was prepared by dissolving lyophilized freeze-dried hexapaptide 1 (2.0 mM) in the buffer described above.

NMR spectra were recorded on a Bruker AMX-500 spectrometer, equipped for multichannel operation. Two-dimensional ^{1}H-^{15}N HSQC experiments were recorded at 25 °C. Suppression of the water signal was carried out by the use of a scrambling pulse (Messerle *et al.*, 1989). A total of 4096 complex points in t_2 with 256 t_1 increments were acquired. The spectral width used was 16 ppm for ^{1}H and 32 ppm for [15]N.

TrNOE experiments were recorded with a mixing time of 200 ms, at 15 °C. A soft presaturation was used to suppress the water signal. A total of 4096 complex points in t_2 with 512 t_1 increments were acquired.

Data were processed using the programs Nmrpipe (Delaglio *et al.*, 1995) and NMRView (Johson & Blevins, 1994). Lorentzian-Guassian or shifted sine-bell apodization and zero-filling were applied prior to Fourier transformation, and subsequent baseline corrections were applied in one or both dimensions.

Chemical shift corrections for incomplete enzyme saturation

The enzyme saturation was achieved in the titration experiments only with hexapeptide 1, the most potent inhibitor in the series. In all the other cases, a partial enzyme precipitation was induced by the larger excess of inhibitor needed. In order to perform a comparison between the shifts induced by each peptide, data had to be normalized to 100 % of complex formation. We made the assumption that a given number of shifts had to be common to the different inhibitors, due to their close structural relationship. For these peaks, one can define a scaling factor s

$$(\delta_F + \Delta\delta_{C1}) - (\delta_F + s\Delta\delta_{Cn}) \leqslant Tol \qquad (2)$$

where δ_F is the chemical shift of a nuclear spin (^{1}H or ^{15}N) of a given NH group for the free enzyme, $\Delta\delta_{C1}$ and $\Delta\delta_{Cn}$ are the shifts induced for this nucleus upon complex formation with hexapeptide 1 and a different peptide inhibitor, respectively. *Tol* is an arbitrary empirical quantity representing the tolerance. In the present work we used 0.1 ppm for ^{1}H and 0.3 ppm for [15]N. In general, $2n$ inequalities can be established, where n is the number of assigned NH groups. If the initial assumption is correct, a single value of s will make inequality (2) true for a large number of peaks, and there will be only a restricted number of peaks for which this cannot be the case. Our goal in this analysis is to obtain the best possible overall superposition of the maximum number of peaks of the two spectra, within a given tolerance, *Tol*. The empirical parameter *Tol* accounts for the experimental differences between the two samples that cannot be attributed with confidence to the different ligands. Additionally, the value of s will represent our best estimate of the level of enzyme saturation reached in the experiment.

As a first step, the algorithm numbers all the resolved peaks in an arbitrary order. Then the best value for s that minimises the first term in inequality (2) for both ^{1}H and ^{15}N of the first peak is calculated. If (2) is fulfilled, this peak is added to the list of ''active constraints''. A second peak is then analysed together with the first (if it is accepted), and a new value of s is searched that minimises the first term of (2) for the four equations. This analysis is repeated until all the peaks are analysed, and in the end the program outputs the final value of s, and the list of peaks for which inequality (2) cannot be fulfilled. The procedure is then repeated a number of times by picking a different peak arbitrary numbering and checking that the final result is independent of the initial choice. The output of the algorithm constitutes the best objective representation of the ''differences'' between the two complexes, in terms of induced chemical shift changes on the NH resonances of the enzyme backbone.

Computational methods

The structural models of the product inhibitors interacting with the protease were obtained from energy minimization and molecular dynamics *in vacuo*. All calculations were carried out with the program BatchMin and the molecular modelling package InsightII/Discover (Biosym Technologies Inc., San Diego, CA). Hydrogen atoms were included, and the potential energy of the complex was expressed by the force-field MMFF (Halgren, 1996), as implemented in the MacroModel V5.0 distribution of the simulation program Batchmin.

Neither non-bonded nor coulombic cutoffs were used and the dielectric constant was set to 1.0. As a starting point for the simulations, the hexapeptide product inhibitor in an energy-minimised extended conformation was docked into the substrate-binding region of the minimised average solution structure of the enzyme (Barbato *et al.*, 1999). The product inhibitor was appropriately oriented in order to allow the P1 carboxylate group to be bound in the oxoanion hole and the peptide backbone and the E2 strand of the protein to form an anti-parallel β-sheet, with hydrogen bonds between P3 and Ala157 and P5 and Cys159. The bound product inhibitor was subsequently energy-minimised. In order to explore the conformational space accessible to the bound inhibitor, a single molecular dynamics simulation was performed at 1000 K with a time-step of 1 fs and 50 conformations were collected at 1.5 ps intervals. This calculation was performed using the NMR-derived, minimised average solution structure of the free enzyme (Barbato *et al.*, 1999). The conformation of the protein was kept fixed and restraints (100 kJ Å^{-2}) were applied between the P1 carboxylate group and the amide protons of the oxyanion hole. Such restraints were applied as flat-bottomed potentials. A group (20) of low-energy conformations were selected and minimised using molecular dynamics and lowering the temperature in a linear fashion to 400 K, while applying restraints (500 kJ Å^{-2}) derived from the NOEs. Restraints between P3 and Ala157 and between P5 and Cys159 were applied to maintain the anti-parallel β-sheet orientation of the inhibitor and the strand E2. Finally a restrained minimization was applied to a final gradient of 0.05 kJ mol^{-1} and the lowest-energy structures were isolated.

Due to the relatively high number of distance restraints, only a small number of conformations consistent with the experimental data was obtained.

Acknowledgments

References

Atherton, E. & Sheppard, R. C. (1989). *Solid Phase Peptide Synthesis: A Practical Approach* (Richwood, D. & Hames, B. D., eds), IRL Press, Oxford.

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) provides new insights into its activation and catalytic mechanism. *J. Mol. Biol.* **289**, 371-384.

Bartenschlager, R. (1997). Molecular targets in inhibition of hepatitis C virus replication. *Ann. Chem. Chemother.* **8**, 281-301.

Bode, W., Papamokos, E. & Musil, D. (1987). The high resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin C. *Eur. J. Biochem.* **166**, 673-682.

Burley, S. K. & Petsko, G. A. (1988). Weakly polar interactions in proteins. In *Advances in Protein Chemistry* (Anfinsen, C. B., Edsall, J. T., Richards, F. M. & Eisenberg, D. S., eds), vol. 39, pp. 125-192, Academic Press, Inc., San Diego, CA.

Chen, Y., Reizer, J., Saier, M. H., Jr, Fairbrother, W. J. & Wright, P. E. (1993). Mapping of the binding interfaces of the proteins of the bacterial phosphotransferase system, Hpr and II Aglc. *Biochemistry*, **32**, 32-37.

Choi, H. K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G. & Wengler, G. (1991). Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature,* **345**, 37-43.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeier, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR,* **6**, 277-293.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M. & James, M. N. G. (1987). Crystal and molecular structures of the complex of alpha chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution. *J. Mol. Biol.* **195**, 397-404.

Gallinari, P., Brennan, D., Nardi, M. C., Brunetti, M., Tomei, L., Steinkühler, C. & De Francesco, R. (1998). Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. *J. Virol.* **72**, 6578-6769.

Grzesiek, S., Stahl, S. J., Wingfield, P. T. & Bax, A. (1996). The CD4 determinant for down regulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry,* **35**, 10256-10261.

Halgren, T. A. (1996). Merk molecular force field. *J. Comput. Chem.* **17**, 490-519.

Ingallinella, P., Altamura, S., Bianchi, E., Taliani, M., Ingenito, R., Cortese, R., De Francesco, R., Steinkuehler, C. & Pessi, A. (1988). Potent inhibitors of human hepatitis C virus NS3 protease are obtained by optimising the cleavage products. *Biochemistry,* **37**, 8906-8914.

James, M. N. G., Sielecki, A. R., Brayer, G. D. & Delbaere, L. T. J. (1980). Structures of product and inhibitor complexes of *Streptomyces griseus* protease at 1.8 A resolution. A model for serine protease catalysis. *J. Mol. Biol.* **144**, 43-88.

Johnson, B. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR,* **4**, 603-614.

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. & Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell,* **87**, 343-355.

Landro, J. A., Raybuck, S. A., Luong, Y. P. C., O'Malley, E. T., Haberson, S. L., Morgenstern, K. A., Rao, G. & Livingston, D. J. (1997). Mechanistic role of an NS4A peptide cofactor with the truncated NS3 protease of hepatitis C virus: elucidation of the NS4A stimulatory effect via kinetic analysis and inhibitor mapping. *Biochemistry,* **36**, 9340-9348.

Llinas-Brunet, M., Bailey, M., Fazal, G., Goulet, S., Halmos, T., Laplante, S., Maurice, R., Poirer, M., Poupart, M. A., Thibeault, D., Wernic, D. & Lamarre, D. (1998). Peptide-based inhibitors of the hepatitis C virus serine protease. *Biorg. Med. Chem. Letters,* **8**, 1713-1718.

Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T. & Hostomska, Z. (1996). The crystal

structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell,* **87**, 331-342.

Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., Margosiak, S., Dagostino, E. & Hostomska, Z. (1998). The conformation of hepatitis C virus NS3 proteinase with & without NS4A: a structural basis for the activation of the enzyme by its cofactor. *Clin. Diagnost. Virol.* **10**, 151-156.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, P. (1983). The geometry of the reactive site of the peptide groups in trypsin trypsinogen and its complexes with inhibitors. *Acta Crystallog. sect. B,* **39**, 480-490.

Martin, P. D., Robertson, W., Turk, D., Huber, R., Bode, W. & Edwards, B. F. P. (1992). The structure of residues 7-16 of the Aα-chain of human fibrinogen bound to bovine thrombin at 2.3 Å resolution. *J. Biol. Chem.* **267**, 7911-7920.

Messerle, B. A., Wider, G., Otting, G., Weber, C. & Wüthrich, K. J. (1989). Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with water solution. *J. Magn. Reson.* **85**, 608-613.

Neddermann, P., Tomei, L., Steinkühler, C., Gallinari, P., Tramontano, A. & De Francesco, R. (1997). The non structural proteins of the hepatitis C virus: structure and functions. *Biol. Chem.* **378**, 469-476.

Ni, F. (1994). Recent development in transferred NOE methods. *Prog. NMR Spectrosc.* **26**, 517-606.

Nienaber, V. L., Breddam, K. & Birkoft, J. J. (1993). A glutamic acid specific serine protease utilizes a novel histidine triad in substrate binding. *Biochemistry,* **32**, 11469-11475.

Nienaber, V. L., Mersinger, L. J. & Kettner, C. A. (1996). Structure based understanding of ligand affinity using human thrombin as a model system. *Biochemistry,* **35**, 9690-9699.

Solé, N. A. & Barany, G. (1992). Optimization of solid-phase synthesis of [Ala]-dynorpin A. *J. Org. Chem.* **57**, 5399-5403.

Spitzfaden, C., Weber, H.-P., Braun, W., Kallen, J., Wider, G., Widmer, H., Walkinshaw, M. D. & Wüthrich, K. (1992). Cyclosporin A - cyclophilin complex formation. A model based on X-ray and NMR data. *FEBS Letters,* **300**, 291-300.

Steinkühler, C., Biasol, 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.

Tong, L., Wengler, G. & Rossmann, M. G. (1993). Refined structure of Sindbis virus core protein and comparison with other chymotripsin-like serine proteinase structures. *J. Mol. Biol.* **230**, 228-247.

Tong, L., Qian, C., Massariol, M.-J., Déziel, R., Yoakim, C. & Lagacé, L. (1998). Conserved mode of peptidomimetic inhibition and substrate recognition of human cytomegalovirus protease. *Nature Struct. Biol.* **5**, 819-826.

Urbani, A., Bianchi, E., Narjes, F., Tramontano, A., De Francesco, R., Steinkhüler, C. & Pessi, A. (1997). Substrate specificity of hepatitis C virus NS3 protease. *J. Biol. Chem.* **272**, 9204-9209.

Viguera, A. R. & Serrano, L. (1995). Side chain interactions between sulfur containing amino acids and phenylalanine in α helices. *Biochemistry,* **34**, 8771-8779.

Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J. & Wright, P. E. (1993). Peptide models of protein folding initiation sites. *Biochemistry,* **32**, 6337-6347.

Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J., Sardana, M., Steinkuehler, C., Tomei, L., De Francesco, R., Kuo, L. C. & Chen, Z. (1988). Complex of NS3 protease and NS4A peptide of BK strain hepatitis C Virus: a 2.2 Å resolution structure in a hexagonal crystal form. *Protein Sci.* **7**, 837-847.

Edited by P. E. Wright