The substrate specificity of a purified protein encompassing the hepatitis C virus NS3 serine protease domain was investigated by introducing systematic modifications, including non-natural amino acids, into substrate peptides derived from the NS4A/NS4B cleavage site. Kinetic parameters were determined in the absence and presence of a peptide mimicking the protease co-factor NS4A (Pep4A). Based on this study we draw the following conclusions: (i) the NS3 protease domain has an absolute requirement for a small residue in the P1 position of substrates, thereby confirming previous modelling predictions. (ii) Optimization of the P1 binding site occupancy primarily influences transition state binding, whereas the occupancy of distal binding sites is a determinant for both ground state and transition state binding. (iii) Optimized contacts at distal binding sites may contribute synergistically to cleavage efficiency.

The N-terminal third of the hepatitis C virus (HCV) NS3 protease domain interacts with the viral protein NS4A. The latter is a 54-residue protein that has been shown to bind to the N-terminal region of the protease via a central hydrophobic domain spanning residues 21–34 (7–13). NS4A acts as a co-factor of the protease enhancing cleavage at all sites and being an absolute requirement for processing of the NS4B/NS5A junction in vivo (7). Several studies have shown that a peptide encompassing the central hydrophobic domain of NS4A is sufficient for eliciting activation of the protease (12, 14–17).

Serine proteases contact the P1 residue of their substrates through characteristic specificity pockets. The residues flanking the specificity pocket are important determinants of substrate recognition. Homology modelling of the S1 specificity pocket of the NS3 protease has predicted the presence of a phenylalanine as a prominent feature, thus rendering the pocket rather small and hydrophobic (18). These characteristics have led to the prediction of the preference for small, hydrophobic residues, ideally cysteine residues, in the P1 position of NS3 substrates. Radiolabeling of the single cleavage products has subsequently confirmed these predictions, yielding the consensus sequence (D/E)XXXXC(A/S) for all trans cleavage sites, with X being any amino acid and the scissile bond being located between Cys and Ala or Ser (2, 18). The homology model has been used to successfully redesign the enzyme’s specificity, thereby increasing its validity. Very recently, the three-dimensional structure of the protease has been solved by two different groups (20, 21), confirming the presence of a phenylalanine ring pointing into the pocket. The sequences of the four NS3 cleavage sites are listed in Table I, indicating that the intramolecular cleavage site between NS3 and NS4A differs from the consensus having a Thr in the P1 position.

Substrate specificity of the NS3 protease has been investigated by several groups in a qualitative way using transient transfection (22, 23), in vitro translation (24), or intracellular processing of fusion proteins in Escherichia coli (25). Availability of quantitative data using peptidic substrates has been so far hampered by difficulties in expressing and purifying sufficient amounts of enzymatically active recombinant NS3 protease. We (17, 26) and others (15, 20, 21, 27–33) have recently described efficient heterologous expression and purification of the enzyme and were able to define optimized conditions for the determination of protease activity (26).

In the present work we investigate the substrate specificity of the NS3 protease domain by introducing systematic modifications in a peptide substrate derived from the sequence of the NS4A/NS4B junction. Activity on modified substrates has been determined both in the presence and absence of the NS4A co-factor. The results are discussed in the light of our previous homology model of the S1 pocket of the enzyme.

**Materials and Methods**

Enzyme Preparation—The protease domain of the HCV Bk strain NS3 protein encompassing residues 1027–1206 of the viral polyprotein was purified from E. coli as described previously (26). The enzyme was homogeneous as judged from silver-stained SDS-polyacrylamide gel electrophoresis and >95% pure as judged from reversed phase HPLC performed using a 4.6 × 250-mm Vydac C4 column. The enzyme preparations were routinely checked by mass spectrometry done on HPLC purified samples using a Perkin Elmer API 100 instrument and N-terminal sequence analysis carried out using Edman degradation on an Applied Biosystems model 470A gas-phase sequencer. Both techniques indicated that in more than 90% of the enzyme molecules the N-terminal methionine and alanine had been removed, yielding an enzyme starting with Pro2. Enzyme stocks were made 50% in glycerol and kept in aliquots at −80 °C until use. Control experiments had shown that this freezing procedure does not affect specific activity of the enzyme.

Peptides and HPLC Assays—Peptide synthesis was performed on a NovaSyn Gem flow synthesizer by Fmoc/t-Bu chemistry. Protecting groups were as follows: Nα-Fmoc, Asp(Ot-Bu), Glu(Ot-Bu), Tyr(t-Bu), and Nα-norvaline.
Substrate Specificity of HCV Protease

FIG. 1. Estimation of active site concentration. To 100 nM protease in 50% glycerol, 2% CHAPS, 10 mM dithiothreitol, 50 mM Tris, pH 7.5, 100 μM fluorogenic ester substrate Ac-DED(Edans)EEAbuC[COO]ASK(Dabcyl)-NH₂ was added. The reaction was stopped at timed intervals by addition of 1% trifluoroacetic acid. The amount of cleavage product was quantified by HPLC. Extrapolation to zero time gave 94 ± 10 nM active sites. Control experiments indicated no detectable spontaneous hydrolysis of the ester bond to occur within the incubation time of the experiment.

Estimation of Active Site Concentration—To determine the number of active molecules in the enzyme preparation 100 μM of the fluorogenic ester substrate Ac-DED(Edans)EEAbuC[COO]ASK(Dabcyl)-NH₂ was added to 100 nM enzyme solution added from a 5 μM stock, previously quantitated by amino acid analysis. 50-μL samples were withdrawn at timed intervals and immediately quenched by addition of 50 μL of 1% trifluoroacetic acid. A total of 48 data points were collected within 3 min of reaction. The pre-steady state burst expected to equal the concentration of active sites in the enzyme preparation was determined by extrapolation of product formation to zero time.

RESULTS

Estimation of Active Site Concentration—We first addressed the question of whether the activity of our protease preparation was attributable to a fully active enzyme or only to a fraction of enzymatically active protease molecules. To this purpose we used the fact that serine proteases follow a two-step mechanism: acylation of the active site serine residue with concomitant release of the C-terminal fragment of the substrate is followed by deacylation and the release of the N-terminal fragment. The former acylation reaction is usually the rate-limiting step for amide bond scission, whereas for ester substrates the acylation reaction is usually fast as compared with deacylation. Thus, for ester substrates the pre-steady-state release of the C-terminal leaving group creates an initial burst that equals the number of active sites encountered by the substrate (35).

To determine the number of active sites in our enzyme preparation 100 μM fluorogenic ester substrate was added to 100 nM enzyme, the reaction was stopped at timed intervals and samples were analyzed by HPLC (Fig. 1). Extrapolation of the linear phase of the reaction indicated a burst of 94 ± 10 nM demonstrating that 94% of the protease molecules in our preparation were enzymatically active.

Substrate Specificity—We investigated the substrate specificity of the NS3 protease by introducing several modifications into a decaamer peptide whose sequence was based on the NS4A-NS4B junction. The choice for this sequence was determined by the relative ease of synthesis (the other two NS3 trans-cleavage sites have two cysteine residues that are prone to oxidation), by the fact that the corresponding junction is cleaved with a relatively high efficiency in the context of the polyprotein, and taking into account that previous mutagenesis studies have shown that this junction is intermediate with respect to its sensitivity to mutations (22). All experiments were done in buffer containing 50% glycerol and 2% CHAPS. These conditions have been worked out to yield optimum activity (26). However, glycerol might exert differential effects on the binding of polar and non-polar substrates. Since we found that glycerol activation is a very complex phenomenon, affecting both Kₘ and k_cat values of substrates, and in addition having a dramatic effect on the dissociation constant of the NS3-4A complex, no attempts were made to study the influence of this agent on the interaction of the enzyme with different substrates. Our data therefore have to be interpreted with the caveat that they have been obtained in the presence of this not entirely passive solvent.

P1 Substitutions—The consensus sequence of the NS3-dependent junctions within the HCV polyprotein points to con-
Substrate Specificity of HCV Protease

Table II

Kinetic parameters of cleavage of P1-modified decamer peptides corresponding to the NS4A/NS4B cleavage site DEMEEC ASHL

The following abbreviations were used: Cacm, S-acetamidomethyl cysteine; Dapa, diaminopropionic acid; hCys, homocysteine; Hyp, hydroxyproline; NVaL, norvaline; Tha, thiazolidino-alanine. Data are mean ± S.D. of at least three different determinations.

| P1 residue | K_m | k_cat | h_cat/K_m | K_m | k_cat | h_cat/K_m |
|------------|-----|-------|-----------|-----|-------|-----------|
| Cys        | 105 ± 20 | 0.26 ± 0.03 | 47.6 | 42.8 ± 4.1 | 1.40 ± 0.07 | 545 |
| hCys       | 312 ± 59 | 0.10 ± 0.01 | 5.30 | 54.1 ± 12.8 | 0.31 ± 0.01 | 95.5 |
| Alg        | 205 ± 29 | 0.08 ± 0.01 | 6.50 | 140 ± 16 | 0.53 ± 0.02 | 63.1 |
| Abu        | 288 ± 30 | 0.03 ± 0.01 | 1.74 | 96.7 ± 28.5 | 0.17 ± 0.01 | 29.6 |
| Thr        | 244 ± 55 | 0.005 ± 0.001 | 0.34 | 235 ± 37 | 0.21 ± 0.01 | 15.1 |
| NVaL       | ND^a | ND | ND | 156 ± 27 | 0.15 ± 0.01 | 16.2 |
| Val        | ND^a | ND | ND | 153 ± 54 | 0.009 ± 0.001 | 0.93 |
| Ala        | Not cleaved | | | | | |
| Pro        | Not cleaved | | | | | |
| Phe        | Not cleaved | | | | | |
| Ser        | Not cleaved | | | | | |
| Hyp        | Not cleaved | | | | | |
| Abu        | Not cleaved | | | | | |
| Cacm       | Not cleaved | | | | | |
| Leu        | Not cleaved | | | | | |
| Dapa       | Not cleaved | | | | | |
| Tha        | Not cleaved | | | | | |

^a ND, not determined (h_cat/K_m < 0.05 M⁻¹ s⁻¹).

Table III

K_m values of P1-modified decamer peptides derived from the NS4A/NS4B cleavage site DEMEEC ASHL

Data are mean ± S.D. from three different determinations.

| P1 residue | NS3 K_m | NS3 + Pep4A K_m |
|------------|---------|----------------|
| Ala        | 323 ± 25 | 384 ± 23 |
| Pro        | 217 ± 42 | 197 ± 21 |
| Phe        | 202 ± 27 | 200 ± 22 |
| Ser        | 91 ± 5 | 22 ± 2 |

K_m = 97 ± 28 μM and K_m = 133 ± 15 μM. Next, we determined the K_m values of decamer peptides having alanine, proline, phenylalanine, or serine in P1, and that were therefore not cleaved (Table III). Interestingly, the K_m values differ only by a factor of 2–8 from the K_m value determined for the wild type substrate, indicating that the P1 residue makes relatively minor contributions to ground state binding. This is true even for a bulky substituent such as phenylalanine.

P6 and P1’ Substitutions—To investigate the role of the side chain length of the P6 residue we substituted the aspartic acid residue present in the NS4A-NS4B sequence by a glutamic acid. This substitution, although affecting slightly k_cat and K_m individually, had no significant effect on the overall cleavage efficiency (Table IV). Neutralization of the negative charge by introduction of an asparagine residue decreased the cleavage efficiency by a factor of 5. This effect was attributable mainly to an increase in K_m. When the charge in the P6 position was inverted through introduction of a lysine residue, a pronounced decrease in k_cat/K_m was observed, which was again attributable to an impairment in ground state binding of the resulting substrate, as judged from the increase of the respective K_m values. All these effects were less pronounced in the presence of the Pep4A cofactor.

Substitution of the P1’ alanine residue by a serine residue, which is found in this position in other substrates, moderately (2–5-fold) decreased k_cat/K_m (Table IV). Conversely, introduction of a phenylalanine residue in P1’ decreased k_cat/K_m by 2 orders of magnitude in the absence of the co-factor and 25-fold in its presence.

Substrate Alanine Scanning—To investigate the relative contribution to cleavage efficiency of positions other than P1,
P6, and P1’ we performed an alanine scanning experiment. The experiment was repeated also in the presence of saturating amounts of Pep4A. Table V summarizes the results. Only the substitution of the P1 cysteine resulted in complete abolishment of cleavage both in the presence and absence of the co-factor. Introduction of alanine residues in other positions had only slight effects on cleavage efficiency. The largest effect (a 7-fold decrease in efficiency) was observed for the P3 position.

Inverse Alanine Scanning of a “Minimalist” Substrate—The picture that emerges from our data confirms the importance of the consensus P6 and P1’ residues, and to a lower extent also of the P6 residue, in determining cleavage efficiency of a given substrate. Still, there are remarkable differences in the kinetic behavior of the single cis cleavage sites which, nevertheless contain the same consensus P6, P1, and P1’ residues (22, 23). We have recently shown that these differences can be reproduced using decamer peptide substrates (26). Thus, there must exist additional determinants. Failure to detect them in the above mentioned alanine scanning experiment might indicate that it is the sum of several minor contributions that modulates the recognition of substrates containing the consensus residues in P6, P1, and P1’. To start to address this issue we synthesized a polyalanine peptide containing the consensus P6, P1, and P1’ residues. Based on the results of the alanine scanning, pointing to the P3 residue as important contribution to cleavage efficiency we decided also to fix a glutamic acid in this position. We produced using decamer peptide substrates (26). Thus, there must exist additional determinants. Failure to detect them in the above mentioned alanine scanning experiment might indicate that it is the sum of several minor contributions that modulates the recognition of substrates containing the consensus residues in P6, P1, and P1’.

The parent peptide Ac-DAAEACAAAAPKY was cleaved, but cleavage was slowed down 70–85-fold with respect to the wild type sequence. Inspection of the sequences of the natural cleavage sites reveals that a negatively charged residue is conserved in the P5 position of two out of four cleavage sites. Re-introduction of the wild type aspartate in our minimalist substrate indeed resulted in a modest (−2-fold) enhancement of cleavage efficiency (Table VI).

In the P4’ position of natural substrates there appears to be a preference for hydrophobic residues (Table I). As a matter of fact, Leu, Tyr, or Trp residues are found in this position. Introduction of Tyr into the P4’ position of the minimalist substrate had no detectable effect, whereas Leu modestly increased cleavage rates. Both P4’ substituted peptides were very insoluble, thus not permitting individual determinations of kcat and kcat/Km. When both P3 Glu and P4’ Leu were re-introduced in the sequence a 20–23-fold enhancement of cleavage efficiency was observed, yielding a substrate that was cleaved with 22–34% efficiency with respect to the wild type sequence.

DISCUSSION

The homology model of the specificity pocket of the NS3 protease predicts that both its shape and its physico-chemical environment be primarily determined by the presence of phenylalanine 213 (according to the chymotrypsin numbering). Furthermore, the pocket was predicted to be very hydrophobic and closed by the aromatic ring of the phenylalanine. While this article was in preparation the crystal structure of NS3 protease has been solved independently by two groups (20, 21). In the published structures the side chain of Phe213 (according to the chymotrypsin numbering).

Inverse Alanine Scanning of a “Minimalist” Substrate—The picture that emerges from our data confirms the importance of the consensus P6 and P1’ residues, and to a lower extent also of the P6 residue, in determining cleavage efficiency of a given substrate. Still, there are remarkable differences in the kinetic behavior of the single cis cleavage sites which, nevertheless contain the same consensus P6, P1, and P1’ residues (22, 23). We have recently shown that these differences can be reproduced using decamer peptide substrates (26). Thus, there must exist additional determinants. Failure to detect them in the above mentioned alanine scanning experiment might indicate that it is the sum of several minor contributions that modulates the recognition of substrates containing the consensus residues in P6, P1, and P1’. To start to address this issue we synthesized a polyalanine peptide containing the consensus P6, P1, and P1’ residues. Based on the results of the alanine scanning, pointing to the P3 residue as important contribution to cleavage efficiency we decided also to fix a glutamic acid in this position. We extended the peptide to the P6’ residue, which being a tyrosine facilitated HPLC detection of cleavage products via monitoring of tyrosine fluorescence. We further introduced a lysine residue in P7’.

The parent peptide Ac-DAAEACAAAAPKY was cleaved, but cleavage was slowed down 70–85-fold with respect to the wild type sequence. Inspection of the sequences of the natural cleavage sites reveals that a negatively charged residue is conserved in the P5 position of two out of four cleavage sites. Re-introduction of the wild type aspartate in our minimalist substrate indeed resulted in a modest (−2-fold) enhancement of cleavage efficiency (Table VI).

In the P4’ position of natural substrates there appears to be a preference for hydrophobic residues (Table I). As a matter of fact, Leu, Tyr, or Trp residues are found in this position. Introduction of Tyr into the P4’ position of the minimalist substrate had no detectable effect, whereas Leu modestly increased cleavage rates. Both P4’ substituted peptides were very insoluble, thus not permitting individual determinations of kcat and kcat/Km. When both P3 Glu and P4’ Leu were re-introduced in the sequence a 20–23-fold enhancement of cleavage efficiency was observed, yielding a substrate that was cleaved with 22–34% efficiency with respect to the wild type sequence.

DISCUSSION

The homology model of the specificity pocket of the NS3 protease predicts that both its shape and its physico-chemical environment be primarily determined by the presence of phenylalanine 213 (according to the chymotrypsin numbering). Furthermore, the pocket was predicted to be very hydrophobic and closed by the aromatic ring of the phenylalanine. While this article was in preparation the crystal structure of NS3 protease has been solved independently by two groups (20, 21). In the published structures the side chain of Phe213 is indeed very insoluble, thus not permitting individual determinations of kcat and kcat/Km. When both P3 Glu and P4’ Leu were re-introduced in the sequence a 20–23-fold enhancement of cleavage efficiency was observed, yielding a substrate that was cleaved with 22–34% efficiency with respect to the wild type sequence.

**TABLE IV**

**Kinetic parameters of cleavage of P6 and P1’-modified decamer peptides corresponding to the NS4A/NS4B cleavage site**

Residues in boldface type indicate modifications with respect to the wild type sequence. Data are mean ± S.D. of at least three different determinations.

| Peptide   | K_m (μM) | k_cat (μM⁻¹ min⁻¹) | k_cat/K_m (μM⁻¹ s⁻¹) | K_m (μM) | k_cat (μM⁻¹ min⁻¹) | k_cat/K_m (μM⁻¹ s⁻¹) |
|-----------|----------|-------------------|----------------------|----------|-------------------|----------------------|
| DEMECASHL | 105 ± 20 | 0.26 ± 0.03       | 47.6                 | 42.8 ± 4.1| 1.40 ± 0.07       | 545                  |
| EEMECASHL | 254 ± 44 | 0.64 ± 0.04       | 42.0                 | 166 ± 15 | 2.89 ± 0.20       | 289                  |
| NEMECASHL | 630 ± 62 | 0.37 ± 0.25       | 9.80                 | 188 ± 4  | 2.90 ± 0.46       | 257                  |
| KEMECASHL | 1050 ± 430| 0.26 ± 0.10        | 4.10                 | 449 ± 91 | 2.65 ± 0.02       | 99.5                 |
| DEMECASHL | 273 ± 21 | 0.32 ± 0.01       | 19.5                 | 246 ± 54 | 1.68 ± 0.22       | 112                  |
| DEMECPASHL| ND       | ND                | 0.41                 | 156 ± 54 | 0.20 ± 0.02       | 21.3                 |

ND, not determined.

**TABLE V**

**Substrate alanine scanning**

Residues in boldface type indicate modifications with respect to the wild type sequence. Data are mean ± S.D. of at least three different determinations.

| Peptide     | K_m (μM) | k_cat (μM⁻¹ min⁻¹) | k_cat/K_m (μM⁻¹ s⁻¹) | K_m (μM) | k_cat (μM⁻¹ min⁻¹) | k_cat/K_m (μM⁻¹ s⁻¹) |
|-------------|----------|-------------------|----------------------|----------|-------------------|----------------------|
| DEMECASHL   | 105 ± 20 | 0.26 ± 0.03       | 47.6                 | 42.8 ± 4.1| 1.40 ± 0.07       | 545                  |
| ASMECASHL   | 125 ± 28 | 0.10 ± 0.02       | 13.3                 | 55.0 ± 11| 0.90 ± 0.06       | 271                  |
| DAMECASHL   | 569 ± 150| 0.86 ± 0.46       | 25.2                 | 143 ± 27 | 2.73 ± 0.84       | 317                  |
| DEAECASHL   | 409 ± 24 | 0.30 ± 0.02       | 12.2                 | 124 ± 22 | 2.13 ± 0.30       | 286                  |
| DEPAECASHL  | 790 ± 205| 0.33 ± 0.06       | 7.0                  | 190 ± 35 | 2.81 ± 0.66       | 246                  |
| DEMECAASHL  | 171 ± 23 | 0.26 ± 0.01       | 22.0                 | 148 ± 44 | 2.50 ± 0.60       | 292                  |
| DEMECAAL    | 191 ± 11 | 0.67 ± 0.04       | 58.4                 | 144 ± 5.6| 3.08 ± 0.15       | 356                  |
| DEMECAAL    | 231 ± 39 | 0.85 ± 0.16       | 61.5                 | 130 ± 15 | 3.61 ± 0.96       | 462                  |
| DEMECASHA   | 360 ± 37 | 0.58 ± 0.12       | 37.1                 | 219 ± 10 | 3.00 ± 0.94       | 226                  |
in uncleaved peptides. Most likely these side chains are too hydrophilic to favorably interact with the hydrophobic milieu of the S1 pocket. Increasing the side chain length of the P1 residue in homocysteine and in allylglycine resulted in substrates that were still reasonably well cleaved. In contrast, incorporation of NVal, having the same side chain length resulted in a more pronounced impairment of cleavage efficiency. This could be related to a favorable interaction of the SH or the allyl group with the phenylalanine ring in the pocket which is expected not to occur in the case of the methyl group in NVal.

The kinetics of Thr and Val substituted peptides, the only two branched residues for which we could detect cleavage, deserve some further comment. The peptide substrate containing Val in P1 was detectably cleaved only in the presence of Pep4A and with a relative efficiency that was 15-fold lower than that observed for Thr in the P1 position. As a matter of fact the isopropyl branch seems detrimental to productive transition state binding as judged by the preference of NVal over Val. Fig. 2 shows a schematic view of the S1 pocket together with a cysteine docked into the pocket and a comparison of the conformation of this cysteine with the conformers of Thr and Val most commonly found in proteins. In this conformation the Val side chain is more likely to encounter steric hindrances in contacting the pocket than would be expected for Thr (Fig. 2). Alternatively, it could be assumed that for both Thr and Val only a methyl group of the side chain will point into the S1 pocket, whereas the other branch (a hydroxy group in Thr and a methyl group in Val) will point out of the pocket. In this view, the fact that Thr is preferred over Val could indicate that its hydroxyl group will make some contacts outside the pocket that cannot be made by the methyl group of Val.

It is interesting to compare our data obtained with peptidic substrates to previous reports in which point mutations were introduced into polyprotein substrates. Kolomeyko et al. (22) have shown that susceptibility to mutations depends on the sequence context, the NS4A/NS4B cleavage site being intermediate between the least sensitive NS3/NS4A junction and the most sensitive NS5A/NS5B cleavage site. Among several substitutions, only Arg and Asp in the P1 position of the NS4A/NS4B junction resulted in complete abolishment of cleavage while a gradient following the order Asn Cys Leu was observed under conditions of short metabolic labeling pulses. Remarkably, in this experiment Leu proved to be even superior to Cys as P1 residue. In further experiments Bartenschlager et al. (23) found that in the context of a NS4A/NS4B junction cleavage was reduced but still well detectable upon substitution of the P1 Cys with Phe, Ser, Thr, and Ala. Clearly, these findings are at variance with our data. However, both the experimental context and the nature of the substrates that have been used might explain these differences. In fact, in transient transfection experiments, even using short labeling times, the accumulation of considerable amounts of substrate and cleavage products is inevitable leading to deviation from true initial rates. This fact compresses differences in cleavage efficiencies. Furthermore, it is possible that NS3 is more active on polyprotein substrates than it is using a peptidic substrate, thereby being less discriminating against suboptimal P1 residues. Differences in specific activity using either polyprotein or peptidic substrates have been reported for other proteases such as CMV protease (38) or tPA (39). Using in vitro translated substrates based on the NS5A/NS5B junction we have estimated the specific activity of added purified protease to be in the order of $k_{cat}/K_m = 200,000$ M$^{-1}$ s$^{-1}$. Nevertheless, we wanted to rule out that the relatively low activities we were observing using peptidic substrates were due to an only partially active enzyme. We have shown that we were indeed working with an enzyme population that was composed of more than 90% of enzymatically active molecules, indicating that the activities we measured under our experimental conditions were intrinsic features of the enzyme.

We found that optimized S1 pocket occupancy was a major determinant of $k_{cat}$, whereas the P1 residue exerted a less pronounced effect on ground state binding of substrate peptides. Incorporation in the P1 position of residues that reduced cleavage of the resulting peptide to undetectable levels (at least 100-fold) resulted in unaltered or only up to 8-fold decreased affinities, as judged from the respective $K_i$ values. Interestingly, this turned out to be true even for residues that were expected to cause significant steric or conformational perturbations such as phenylalanine or proline. This behavior sheds some light on the mechanisms of substrate recognition by the NS3 protease: apparently ground state binding of the substrate is mediated by multiple interactions involving distal residues,

---

**TABLE VI**

**Inverse substrate alanine scanning**

Residues in boldface type indicate modifications with respect to the minimalist sequence DAAEACAAAAPYK. Data are mean ± S.D. of at least three different determinations.

| Peptide       | $K_m$  | $k_{cat}$ | $k_{cat}/K_m$ | $K_m$  | $k_{cat}$ | $k_{cat}/K_m$ |
|---------------|--------|-----------|---------------|--------|-----------|---------------|
|               | $\mu M$| min$^{-1}$| $s^{-1}$      | $\mu M$| min$^{-1}$| $s^{-1}$      |
| DMEECASHLPYK  | 80.1 ± 10.0| 0.5 ± 0.1 | 104  | 40.5 ± 5.2| 3.5 ± 0.5 | 1460         |
| DAAEACAAAAPYK | 1000 ± 216| 0.07 ± 0.01| 1.2 | 330 ± 119| 0.42 ± 0.08| 21           |
| DAAEACAAAAPYK | 490 ± 85 | 0.08 ± 0.01| 2.7 | 130 ± 76 | 0.38 ± 0.06| 48           |
| DAAEACAAAAPYK | ND      | ND        | 0.8 | ND      | ND        | 14           |
| DAAEACAAAAPYK | ND      | ND        | 2.3 | ND      | ND        | 48           |
| DAAEACAAAAPYK | 249 ± 26 | 0.34 ± 0.01| 22.7| 42.6 ± 10| 1.26 ± 0.07| 492          |

ND, not determined.

---

**FIG. 2. Schematic view of the model of the S1 pocket of the NS3 protease.** Solid lines indicate the backbone of residues flanking the S1 pocket. Cysteine is docked into the pocket with its side chain pointing in the direction of Phe$^{211}$. Cys, Thr, and Val are shown in the conformations most commonly found in proteins.

---

$^2$ R. De Francesco unpublished observations.
whereas the efficiency with which the bound substrate will proceed through the transition state is strongly influenced by the nature of the residue in the P1 position. This dual requirement is probably needed to endow the enzyme with the high degree of specificity necessary to accomplish its physiological role of mediating the generation of the mature HCV replication machinery.

In the P6 position a conserved negative charge is present in all cleavage sites. From our data it appears that, at least for the NS4A/NS4B junction there is a preference but no stringent requirement for this negative charge. This finding is in agreement with what has been found by others through introduction of point mutations in polyprotein substrates (22, 23). In fact, it has been reported that, in the context of different cleavage sites, extensive mutagenesis of the P6 position has little if any effect on cleavage efficiency. The absolute conservation of this residue especially in the light of the pronounced variability of the HCV genome might therefore indicate that it serves some more subtle function.

We attempted to identify additional crucial determinants of recognition by the protease by both “classical” alanine scanning and by “inverse alanine scanning” using a polyalanine substrate with 4 non-alanine residues. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimalist polyalanine substrate with 4 non-alanine positions. Re-introduction of wild type sequence residues in the minimal...
Substrate Specificity of the Hepatitis C Virus Serine Protease NS3
Andrea Urbani, Elisabetta Bianchi, Frank Narjes, Anna Tramontano, Raffaele De Francesco, Christian Steinkühler and Antonello Pessi

J. Biol. Chem. 1997, 272:9204-9209.
doi: 10.1074/jbc.272.14.9204

Access the most updated version of this article at http://www.jbc.org/content/272/14/9204

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 38 references, 21 of which can be accessed free at http://www.jbc.org/content/272/14/9204.full.html#ref-list-1