ELSEVIER

International Journal of Mass Spectrometry 176 (1998) 113–124

Rapid mass spectrometric determination of preferred irreversible proteinase inhibitors in combinatorial libraries

John E. Mckendrick[a], Sven Frormann[a], Colin Luo[a], Paul Semchuck[a], John C. Vederas[a,*], Bruce A. Malcolm[b,*]

[a]*The Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada*
[b]*The Department of Biochemistry, Tufts University, Boston, MA 02111, USA*

Received 25 March 1998; accepted 31 March 1998

Abstract

Optimal *N*-iodoacetyldipeptide inactivators of hepatitis A virus 3C proteinase were identified directly from equimolar mixtures of these compounds using electrospray ionization mass spectrometry (ESI-MS). Limiting amounts of proteinase were allowed to react with the library of inhibitors and subsequently analyzed by ESI-MS to determine the mass of the adducts formed. *N*-iodoacetyl-Ser-Phe-NH_2 was found to be the most potent inactivator with a second order rate constant of 840 ± 90 $M^{-1}s^{-1}$. Fragmentation of the complexes by using cyanogen bromide and trypsin followed by liquid chromatography/ESI-MS confirmed the identity of the adduct and allowed inhibitor mass differences of as little as 6 Da to be distinguished in a single experiment. This approach allows the rapid screening and identification of preferred covalent inhibitors or intermediates from combinatorial libraries without deconvolution or resynthesis and should be applicable to irreversible inhibitors of virtually any enzyme that uses a covalent catalysis mechanism. (Int J Mass Spectrom 176 (1998) 113–124) © 1998 Elsevier Science B.V.

Keywords: Inhibitors; Proteinase; Mass Spectrometry; Hepatitis A Virus

1. Introduction

The widespread application of combinatorial chemistry to accelerate drug discovery programs [1] has focused on development of both solid phase [2] and solution phase [3] chemical syntheses of libraries of compounds for biologic screening and lead selection. A number of drug candidates discovered by these procedures are now in clinical trials [4]. The problem of identification of the active compound(s) in a library mixture is often solved either by specific attachment of a covalently bonded tag (e.g. chromophore) or by generation and repetitive screening of sublibraries [5]. Recently, affinity chromatography using capillary electrophoresis directly, as well as in conjunction with on-line mass spectrometry, has been successful in determining which peptidic ligands bind most strongly to receptors [6]. A very attractive, nonchromatographic, alternative approach involves direct selection of the best binding compounds (e.g. sulfonamides with IC_{50} in nanomolar range) from a library mixture by the key enzyme (e.g. carbonic anhydrase), followed by identification of the *noncovalent* ligand–protein complexes by ion cyclotron resonance mass spectrometry [7]. In the present study,

* Corresponding authors.

PII S1387-3806(98)14022-8

Table 1
Selected second order rate constants for iodoacetyldipeptides

IAc-AA_1–AA_2–NH_2	k_2 (M^{-1} s^{-1})
Gly-Pro	9.4 ± 1.6
Gly-Va	5.5 ± 0.4
Gly-Thr	9.8 ± 0.6
Gly-Tyr	24 ± 2
Val-Phe	200 ± 30
Ser-Phe	840 ± 90
Ser-Phg	240 ± 20
Ser-Cha	230 ± 50
Hse-Phe	200 ± 30
Phe-Phe	12 ± 3

Iodoacetamide inhibited HAV 3C with k_2 = 3.1 ± 1.8 M^{-1} s^{-1}.

Table 2
Mixtures of *N*-iodoacetylpeptides prepared for inactivation of HAV 3C and subsequent analysis by electrospray mass spectrometry

Mixture	AA_1	AA_2
	Gly	Val, Pro, Thr, Tyr
1	Gly	Leu-Ser
	Val	Phe
	Iodoacetamide	
2	Ser	Gly, Ala, Ser, Val, Leu, Gln, His, Phe, Tyr, Trp
3	Gly, Ala, Ser, Val, Leu, Gln, Phe, Tyr, Trp	Phe, His
4	Ser, iAsn, Trp	Phg, Phe, Cha, ITyr
5	Hse, Phe, ITyr	Phg, Phe, Cha, ITyr

we report that a cysteine proteinase can select the best peptide-based *irreversible* inactivator from a mixture of analogs, and that the resulting *covalent* inhibitor–enzyme complex can be readily identified by standard electrospray mass spectrometry available in most laboratories.

Studies by a number of groups, including our own, have targeted design of inhibitors of the picornaviral family of 3C proteinases as potential therapeutic agents for pathogens such as human rhinovirus, hepatitis A virus (HAV), poliovirus, and foot and mouth disease virus [8–10]. Recent work shows that peptide-based inhibitors (e.g. aldehydes, fluoromethyl ketones) of these cysteine proteinases can dramatically inhibit viral replication in cell culture [11, 12]. Key goals in the current study were to (1) obtain and optimize inhibitors based on the P′ recognition subsites (carboxy terminal side of the scissile bond) of the HAV 3C proteinase; (2) determine whether the enzyme can effectively discriminate between inherently reactive irreversible inactivators in small libraries; and (3) ascertain whether electrospray ionization on standard mass spectrometers provides a useful method of analysis for such experiments with enzymes of modest size (HAV 3C: 23,877 Da).

2. Results and discussion

Initially the seven *N*-iodoacetyl peptides in mixture 1 (Table 1) were separately synthesized using standard solid phase techniques (RINK amide resin [16]) with *N*-terminal iodoacetylation followed by trifluoroacetic acid cleavage from the resin with concomitant removal of side-chain protecting groups. Each compound, as well as pure *N*-iodoacetamide, was individually evaluated as an irreversible inhibitor of HAV 3C proteinase to determine its second order rate constant of inactivation (Table 2). Displacement of iodide by the active site cysteine thiol group results in irreversible alkylation of the enzyme (Fig. 1).

A minilibrary (mixture 1, Table 1) containing 10 equivalents of each substance was then prepared and used to inactivate 1 equivalent of HAV 3C proteinase. Precipitation of the enzyme–inhibitor complex followed by acquisition of its electrospray mass spectrum and deconvolution of the multiply charged peaks gave a single molecular ion (MH^+) at 24,181 corresponding to the adduct derived by reaction of the enzyme with the most effective inhibitor in the group, *N*-iodoacetyl-Val-Phe-NH_2 ($C_{16}H_{22}N_3O_3$ fragment MW = 304; k_2 = 200 ± 30 M^{-1} s^{-1}) (Fig. 2). Interestingly, this compound was able to completely out-compete *N*-iodoacetyl-Gly-Leu-Ser-NH_2, which

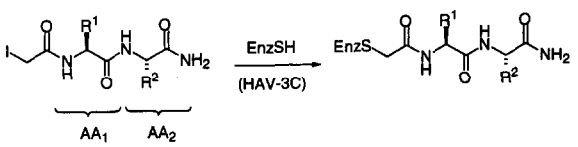

Fig. 1. Alkylation of the active site thiol of HAV 3C using iodoacetyl dipeptides, resulting in irreversible inactivation.

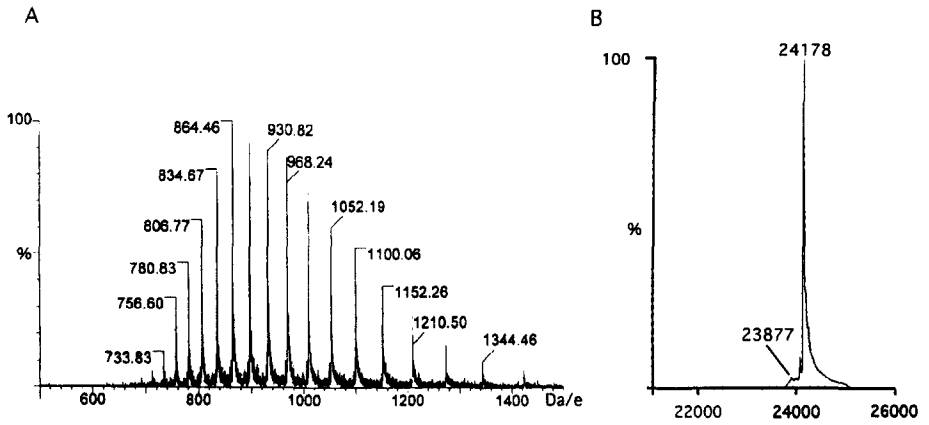

Fig. 2. (A) Raw ion pattern for HAV 3C inhibited by mixture 1. (B) Deconvoluted spectrum showing the exact mass of the isolated protein.

had a rate constant of inactivation only threefold lower ($k_2 \sim 60\ M^{-1}\ s^{-1}$).

Additional libraries (Table 2), containing both natural and unnatural amino acids as their *N*-iodoacetamide derivatives, were synthesized as mixtures and used to inhibit HAV 3C as before (each constituent in tenfold excess). The components in each mixture were chosen such that their mass difference after attachment to the 24-kDa protein could be easily distinguished by electrospray ionization mass spectrometry (ESI-MS) (≥10 u). The results clearly show that the most active inhibitor, *N*-iodoacetyl-Ser-Phe-NH_2 ($k_2 = 840 \pm 90\ M^{-1}\ s^{-1}$), is readily selected by the enzyme from each test mixture that contains it (Fig. 3).

The selection of Ser-Phe as the preferred P′1–P′2 residues is consistent with the natural polyprotein cleavage substrates of HAV 3C, which is known to have the sequence Ser-Phe in positions P′1–P′2 [13, 14]. The extra methylene group present in the iodoacetyldipeptides was thought to have been accommodated in the active site of HAV 3C without major structural changes. It was found that structural changes in the active site were apparent when a crystal structure of HAV 3C with *N*-iodoacetyl-Val-Phe-NH_2 was solved, showing the inhibitor bound in the P′ subsites. Details of the crystallization methodology and solution of the crystal structure will be reported elsewhere [15].

To explore more subtle variations of the Ser-Phe P′1–P′2 scaffold motif, a series of *N*-iodoacetyldipeptides containing related amino acid analogs was synthesized and screened (Table 1, mixtures 4 and 5). Analysis of mixture 4 gave rise to a species of molecular weight 24,172 Da (ΔM 295 Da) suggesting primary inactivation with either *N*-iodoacetyl-Ser-Phe-NH_2 (ΔM 298 Da) or *N*-iodoacetyl-Ser-Cha-NH_2 (ΔM 292 Da) or a mixture of the two, along with a smaller peak corresponding to adducts with *N*-iodoacetyl-iAsn-Phe-NH_2 or *N*-iodoacetyl-iAsn-Cha-NH_2 [Fig. 4(i)]. Despite repeated attempts, resolution was insufficient to conclusively determine whether the major peak was the result of inactivation with *N*-iodoacetyl-Ser-Phe-NH_2 or *N*-iodoacetyl-Ser-Cha-NH_2 because of mass averaging inherent in ESI-MS. Incubation of mixture 5 with HAV 3C resulted in a number of adducts [Fig. 4(ii)]; the most abundant species corresponded to inactivation with *N*-iodoacetyl-Hse-Phe-NH_2, 24180 Da (ΔM 306 Da). Other species present in this mixture were assigned to inactivation by *N*-iodoacetyl-Hse-3-ITyr-NH_2 24326 Da (ΔM 449 Da) and *N*-iodoacetyl-3-ITyr-Phe-NH_2 24,372 Da (ΔM 494 Da).

To determine unequivocally whether the primary adduct from mixture 4 contained *N*-iodoacetyl-Ser-Phe-NH_2 or *N*-iodoacetyl-Ser-Cha-NH_2, as well as to determine whether other inactivated HAV 3C species were present under the major peak of the ESI-MS

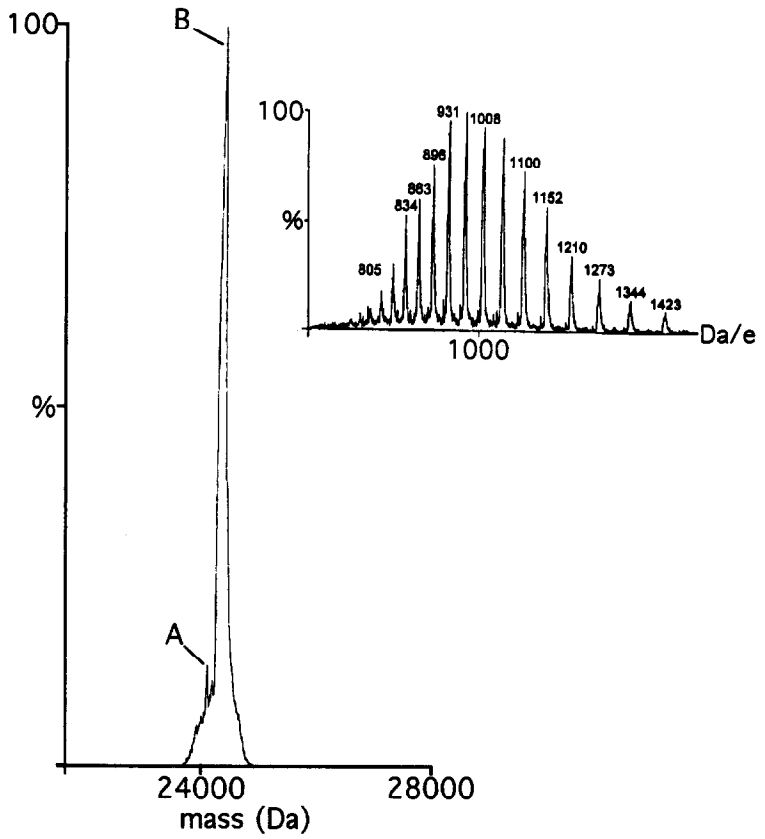

Fig. 3. Deconvoluted spectrum of HAV 3C inactivated with mixture 2. The only detectable complex was that formed by inactivation with *N*-iodoacetyl-Ser-Phe-NH_2 (peak B, 24174 Da), a small amount of uninhibited HAV 3C was also present (A).

spectrum, attempts were made to fragment reproducibly the complexes before ESI-MS analysis. Fragmentation allows greater resolution, enabling the detection of minor adducts that may have had very similar masses to that of the major product. Initially cyanogen bromide, which causes peptide backbone cleavage after methionine residues, was used to fragment the complexes (Fig. 5(A)). To validate the approach, HAV 3C proteinase, inhibited separately with 10 equivalents of either *N*-iodoacetyl-Ser-Phe-NH_2 or *N*-iodoacetyl-Ser-Cha-NH_2, was combined (in a 2:1 molar ratio), fragmented, and analyzed by liquid chromatography (LC)/ESI-MS. Both *N*-iodoacetyl-Ser-Phe-NH_2 and *N*-iodoacetyl-Ser-Cha-NH_2 thiol adducts were present and could be easily assigned (Fig. 6).

Having demonstrated that BrCN digestion followed by LC/ESI-MS could clearly distinguish between adducts differing by only 6 Da, further experiments to determine whether multiple adducts were hidden under major peaks were undertaken. Inactivation of HAV 3C was carried out using two new mixtures, a four component mixture consisting of *N*-iodoacetyl-*X*-Phe-NH_2 (where X = Ser, Tyr, Gly, and His) and an eight component mixture consisting of the previous four component mixture augmented with X = Val, Asn, Gln, and Arg. The activated dipeptides required were synthesized individually and mixed together in equimolar quantities as before. Direct ESI-MS analysis indicated that HAV 3C was almost completely inactivated, with a mass increase corresponding to inactivation with *N*-iodoacetyl-Ser-Phe-NH_2 in both cases.

Digestion of the complexes from both experiments

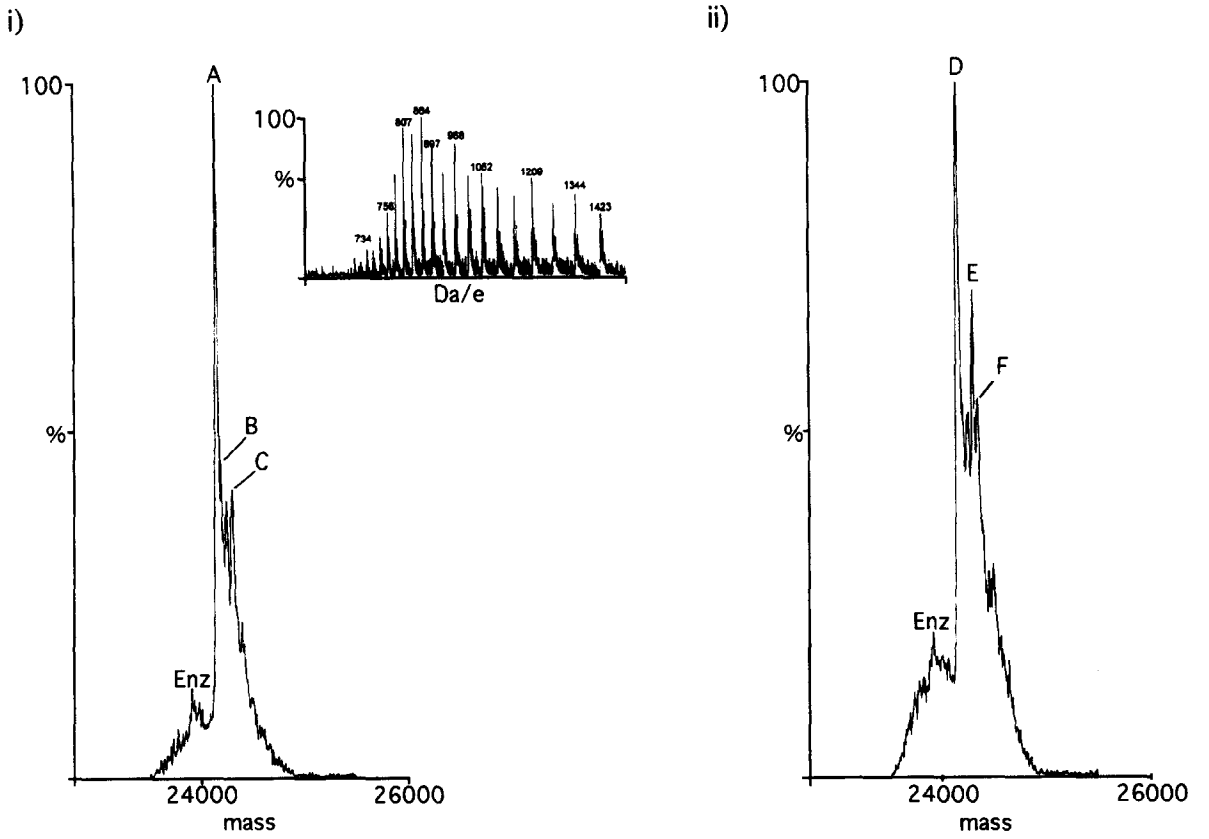

Fig. 4. (i) Deconvoluted mass spectrum of HAV 3C inhibited with mixture 4, the insert shows the smoothed raw data. The major peak A corresponds to Enzyme+acetyl-Ser-Phe/Cha-NH_2 (24,172 Da), the minor peaks B and C correspond to inactivation with *N*-iodoacetyl-iAsn-Phe/Cha-NH_2 (24,210 Da) and *N*-iodoacetyl-Ser-3-ITyr-NH_2 (24,317 Da). (ii) Deconvoluted mass spectrum of HAV 3C inhibited with mixture 5. Peak D corresponds to E+acetyl-Hse-Phe-NH_2 (24,180 Da); peaks E and F correspond to inactivation with *N*-iodoacetyl-3ITyr-Phe-NH_2 (24,326 Da) and *N*-iodoacetyl-Phe-3-ITyr-NH_2 or *N*-iodoacetyl-Phe-3-ITyr-NH_2 (24,372 Da). A small amount of inactivated enzyme was recorded, denoted Enz (24877 Da).

with BrCN followed by LC/ESI-MS analysis confirmed the presence of the expected alkylated cysteine peptide fragment; however, an additional product of lower molecular weight (3684.7 Da) corresponding to the unmodified peptide was also detected. The ratio of the modified and unmodified products varied from experiment to experiment, with the unalkylated peptide often dominating the spectra, despite apparent complete alkylation of the enzyme before BrCN treatment (based on direct ESI-MS). On reflection, it became apparent that cyanogen bromide could activate the alkylated cysteine in a manner similar to the activation of methionine (Fig. 5(B)). Subsequent cyclization and cleavage of the inhibitor would remove the dipeptide, resulting in the regeneration of the unmodified fragment seen in the LC/ESI-MS analysis. The rate of cleavage of the inhibitor from the protein is presumably dependent on the substrate dipeptide and, as such, limits the use of this approach to qualitative confirmation of major adducts.

To confirm and extend the above results, trypsinolysis was used in place of BrCN for fragmentation before LC/ESI-MS. After inactivation of the HAV 3C by either the four or eight component mixture used for the BrCN studies above, samples were dialyzed into digestion buffer and treated with trypsin for 18 hours. LC/ESI-MS analysis of the adducts formed with the four component mixture showed the presence of three

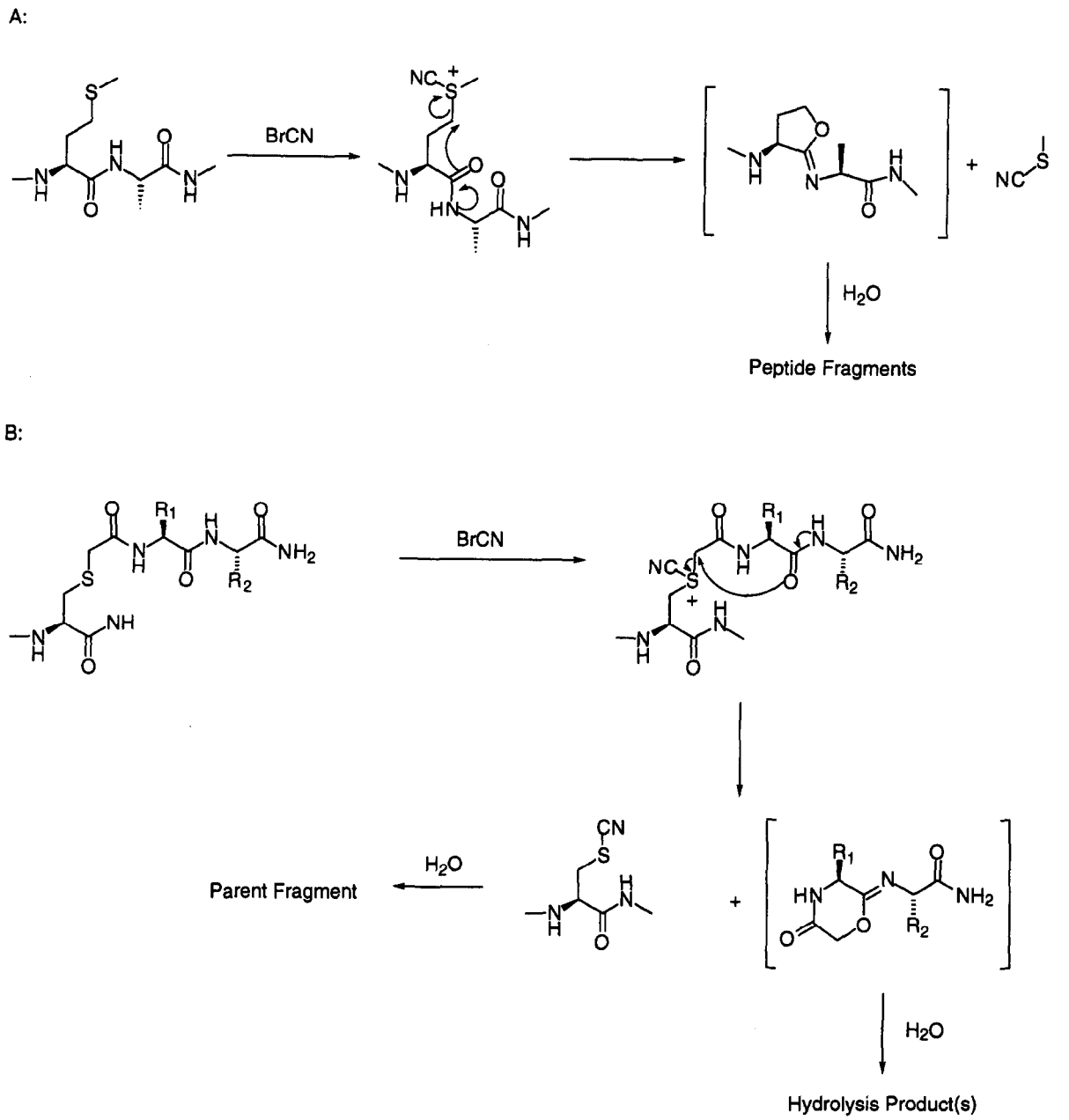

Fig. 5. (A) Mechanism of cyanogen bromide cleavage of peptides after methionine residues. (B) Mechanism of cyanogen bromide cleavage responsible for the removal of the dipeptide inactivator from the alkylated cysteine residue, resulting in the appearance of the native enzyme cleavage products.

distinct species (Fig. 7). The major peak (4112.8 Da) corresponded to HAV 3C fragment 163–202 (an incomplete trypsin digestion product), plus acetyl-Ser-Phe-NH_2. The second peak (minor) was found to be the expected complete digestion product of HAV 3C (residues 165–202) inactivated with acetyl-Ser-Phe-NH_2 (3928.8 Da). The last peak, 4192.4 Da, appeared to correspond to residues 163–202 with the

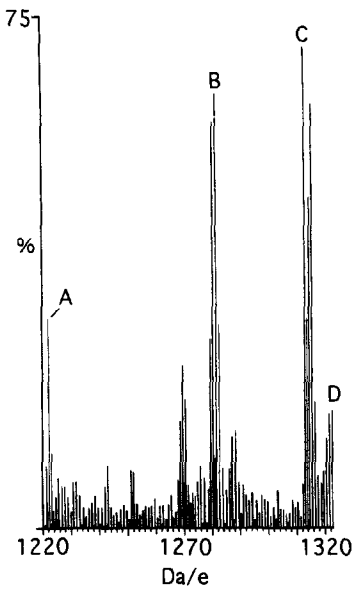

Fig. 6. ESI-MS spectrum of BrCN digestion of Ser-Phe/Ser-Cha inhibited HAV 3C. Peak A corresponds to uninhibited HAV 3C (1225 Da), B is an incomplete digestion product (residues 138–208), C corresponds to HAV 3C inhibited with *N*-iodoacetyl-Ser-Phe-NH_2 (1317 Da) and D corresponds to HAV 3C inhibited with *N*-iodoacetyl-Ser-Cha-NH_2 (1320 Da).

addition of acetyl-Tyr-Phe-NH_2 (4188.9 Da). The study using the eight component mixture, produced comparable results (major product, acetyl-Ser-Phe-NH_2, 4110.0 Da; minor product, acetyl-Tyr-Phe-NH_2, 4190.8 Da).

3. Conclusions

ESI-MS analysis of adducts formed between limiting amounts of enzyme and a mixture of irreversible inhibitors made possible the rapid identification of the optimal activated dipeptide inactivators of HAV 3C protease (*N*-iodoacetyl-Ser-Phe-NH_2). Fragmentation of complexes before LC/ESI-MS analysis also allowed less potent inhibitors to be identified and mass resolution to be increased. This approach should have great utility in the initial phases of inhibitor identification and optimization where large numbers of compounds, with disparate activities, need to be quickly screened. The major limitation of direct

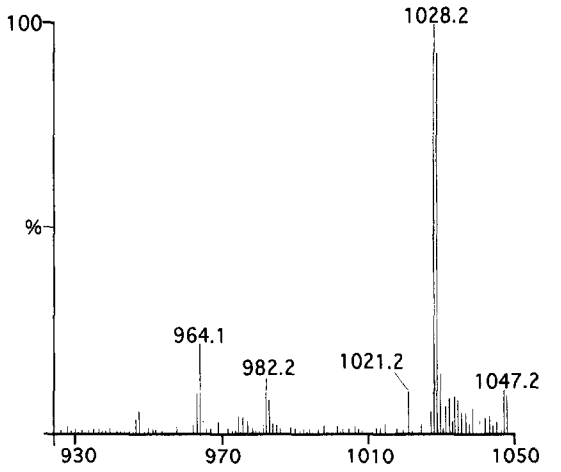

Fig. 7. Portion of the raw data for HAV 3C inhibited by the four component mixture and digested using trypsin. The area shown corresponds to the +4 region of the spectrum showing all modified peptides isolated from the digestion experiment.

ESI-MS analysis appears to be the mass averaging inherent in the technology, which somewhat limits mass resolution; however, judicious design of mixtures should circumvent or minimize this limitation. Alternatively, fragmentation of complexes followed by LC/ESI-MS can increase mass resolution, allowing inhibitors of closer molecular weight to be screened in the same mixture. These approaches should be widely applicable to irreversible inhibitors of all enzymes that use a covalent catalysis mechanism (e.g. lipases, glycosylases, and proteinases).

4. Appendix

4.1. General experimental procedures

4.1.1. Synthesis of iodoacetyldipeptides

Peptide amides were chemically synthesized using FMOC strategy on 4-(2′,4′-dimethoxyphenyl-*N*-Fmoc-aminomethyl)-phenoxy "RINK resin" [16] (Advanced Chemtech, Louisville, KY) and were reacted with freshly prepared iodoacetic acid active ester. In a typical procedure:

Iodoacetylserinylphenylalaninamide. A solution of iodoacetic acid (185 mg, 1 mmol), 1-hydroxybenzo-

triazole (132 mg, 1 mmol), and dicyclohexylcarbodiimide (226 mg, 1 mmol) in dimethyl formamide (DMF) (3 mL) was stirred for 90 minutes at RT and then filtered, washing the DCU with a further 1 mL of DMF. To this solution was added, with stirring, Ser-Phe-RINK resin (synthesized from 0.2 mmol RINK resin) in small portions and the mixture was stirred for 2 h at room temperature (RT). The resin was filtered, washed with DMF, EtOH, and CH_2Cl_2, and dried *in vacuo* for 30 min. The peptide was cleaved from the resin by treatment with 95% aqueous TFA (10 mL) for 1 h. The resin was filtered and the filtrate concentrated *in vacuo*. The crude residue was dissolved in toluene (15 mL) and evaporated *in vacuo*, to remove excess TFA. The resulting solid was washed with Et_2O and purified by reverse phase high performance liquid chromatography (HPLC) (Resolve, 10 Å, 0–30% MeCN/H_2O+0.1% TFA) to give *N*-iodoacetyl-Ser-Phe-NH_2 (39 mg, 33%); mp 190°C [decomposed (dec)]; IR (CH_3CN cast) 3430, 3412, 3285, 3120, 1678, 1665, 1629, 1548, 1495, 1173, 702 cm^{-1}; ^1H NMR (300 MHz, CD_3OD/CD_3CN 4:1) 7.25–7.12 (m, 5H, ArC<u>H</u>), 4.60 (dd, 1H, *J* 9.0, 5.1 Hz, NHC<u>H</u> Phe), 4.30 (t, 1H, *J* 6.0 Hz, NHC<u>H</u> Ser), 3.77–3.60 (m, 4H, $IC\underline{H}_2$ and $C\underline{H}_2$ Ser), 3.21 (dd, 1H, *J* 14.0, 5.1 Hz, $C\underline{H}_2$ Phe), 2.94 (dd, 1H, *J* 14.0, 9.0 Hz, $C\underline{H}_2$ Phe); ^{13}C NMR (75 MHz, CD_3OD/CD_3CN 4:1) 175.7, 171.9, and 171.2 (qC), 138.6, 130.5, 129.7, and 128.0 (ArC), 62.7 (CH_αSer), 57.1 (CH_αPhe), 55.8 (CH_2 Ser), 38.5 (CH_2 Phe), −1.5 (ICH_2); MS (EI, 70 eV) 419.0342 (1.6) [M^+], 402.0063 (1.0) [M^+-H_3N], 389.0217 (1.1) [M^+-CH_2O], 375.0193 (5.8) [M^+-CH_2ON], 147.0684 (51) [C_9H_9ON], 120.0813 (100) [$C_8H_{10}N$], 91.0547 (29) [C_7H_7], 60.0448 (47) [C_2H_6ON]; high resolution mass spectrometry (HRMS) *m/z* 419.03420 (M^+) (calculated for $C_{14}H_{18}O_4N_3I$ 419.03421).

Iodoacetylglycylvalinamide. IR (microscope) 3275, 3198, 1676, 1649, 1625, 1551 cm^{-1}; ^1H NMR (360 MHz, CD_3OD) 4.24 (d, 1H, *J* 6.3 Hz NHC<u>H</u> Val), 3.90 (s, 2H, $IC\underline{H}_2$), 3.78 (s, 2H, $NHC\underline{H}_2$ Gly), 2.12 (m, 1H, C<u>H</u> Val), 0.98 (m, 6H, 2x$C\underline{H}_3$ Val); MS *m/z* 297.0100 (7.6) [M^+-CH_2ON], 72.0814 (100) [$C_4H_{10}N$]; 72.0449 [C_3H_6ON]; HRMS *m/z* 297.0100 (M^+-CH_2ON) (calculated for $C_9H_{16}O_3N_3I$-CH_2ON 297.0103).

Iodoacetylglycylprolinamide. IR (microscope) 3320 (br), 1640 (C=O), 1166 cm^{-1}; ^1H NMR (400 MHz, CD_3OD, major rotamer) 4.41 (dd, 1H, *J* 9.0 3.3 Hz, NHC<u>H</u> Pro), 4.04 (q_{AB}, 2H, *J* 16.8, $IC\underline{H}_2$), 3.78 (s, 2H, $NHC\underline{H}_2$ Gly), 3.71–3.53 (m, 2H, $C\underline{H}_2$ Pro), 2.19 (m, 1H, $C\underline{H}_2$ Pro), 2.00 (m, 3H, CH_2 Pro); MS *m/z* 339.0064 (0.15) [M^+], 212.1032 (0.24) [M^+-I], 195.0770 (1.9) [M^+-H_3N-I], 169.0972 (2.1) [M^+-CHON], 70.0649 (100) [C_4H_8N]; HRMS *m/z* 339.0064 (M^+) (calculated for $C_9H_{14}O_3N_3I$ 339.0080).

Iodoacetylglycylthreoninamide. mp 104°C; IR (microscope) 3430, 3399, 3324, 3256, 1656, 1652, 1639, 1581, 1563, 1151 cm^{-1}; ^1H NMR (360 MHz, CD_3OD) 4.29 (m, 1H, NHC<u>H</u> Thr), 4.24 (m, 1H, C<u>H</u> OH Thr), 3.94 (q_{AB}, 2H, $IC\underline{H}_2$), 3.78 (s, 2H, $NHC\underline{H}_2$ Gly), 1.17 (d, 3H, *J* 5.3 Hz, $C\underline{H}_3$ Thr); ^{13}C NMR (100 MHz, CD_3OD) 175.2, 172.1, 171.6 (qC), 68.0 (CHOH Thr), 59.7 (C_α Thr), 44.3 (C_α Gly), 20.2 (CH_3 Thr), −2.7 (ICH_2); MS *m/z* [electrospray ionization (EI), 70 eV] 299.9860 (20) [$M^+-C_2H_3O$], 282.9569 (16) [M^+-H_3N], 225.9365 (40) [Iac-Gly], 197.9416 (49) [$IAcNHCH_2$], 57.0575 (99) [C_3H_7N], 44.9977 (100) [CHO_2]; HRMS *m/z* 299.9860 (M+−C_2H_3O) (calculated for $C_8H_{14}O_4N_3I-C_2H_3O$: 299.9845).

Iodoacetylglycyltyrosinamide. IR (microscope) 3378, 3307, 3207, 1672, 1656, 1634, 1610, 1594, 1548, 1513, 835 cm^{-1}; ^1H NMR (360 MHz, D_2O) 7.18 (d, 2H, *J* 8.4 Hz, ArC<u>H</u>), 6.88 (d, 2H, *J* 8.4 Hz, ArC<u>H</u>), 4.58 (dd, 1H, *J* 8.4, 6.0 Hz, NHC<u>H</u> Tyr), 3.88 (q_{AB}, 2H, $IC\underline{H}_2$), 3.81 (s, 2H, $NHC\underline{H}_2$ Gly), 3.13 (dd, 1H, *J* 14.1, 6.0 Hz $C\underline{H}_2$ Tyr), 2.96 (dd, 1H, *J* 14.1, 8.6 Hz $C\underline{H}_2$ Tyr); ^{13}C NMR (100 MHz, CD_3OD/DMSO-d_6 4:1) 175.2, 170.7, and 170.1 (qC), 156.8, 130.9, 128.7, and 115.8 (ArC), 55.4 (C_α Tyr), 43.5 (C_α Gly), 37.6 (CH_2 Tyr), −2.0 (ICH_2); MS *m/z* (FAB^+, Cleland) 405.2 (2.9) [MH^+].

Iodoacetylglycylleucylserineamide. mp 189°C (dec); IR (microscope) 3431, 3296, 3273, 3208, 1675, 1655, 1625, 1549, 1173 cm^{-1}; ^1H NMR (360 MHz, DMSO-d_6) 8.49 (t, 1H, NH), 8.05 (d, 1H, *J* 7.3 Hz, NH), 7.82 (d, 1H, *J* 7.9 Hz, NH), 7.12 and 7.04 (2xbrs, 2H, NH_2), 4.83 (t, 1H, *J* 5.4 Hz, OH), 4.32 (dd, 1H, NHCH Leu), 4.15 (t, 1H, *J* 5.4 Hz, NHCH Ser), 3.73 (2xs, 2H, $NHCH_2$ Gly), 3.56 (m, 2H, *J* 5.4 Hz, CH_2 Ser), 1.59 (m, 1H, CH Leu), 1.46 (m, 2H, CH_2 Leu), 0.87 and 0.83 (2d, 6H, 2xCH_3 Leu); ^{13}C NMR (100 MHz, CD_3OD/DMSO-d_6 4,1) 173.8, 173.7, 170.7 (qC), 62.4 (CH_2 Ser), 56.3 (C_α Leu), 52.8 (C_α Ser), 43.7 (C_α Gly), 41.0 (CH_2 Leu), 25.2 (CH Leu), 23.3 and 21.5 (CH_3 Leu), −1.7 (ICH_2); MS *m/z* (FAB, Cleland) 464.9 (4.3) [MNa^+], 442.9 (2.2) [MH^+]; HRMS *m/z* 412.0605 (M^+-CH_2O) (calculated for $C_8H_{14}O_4N_3I-CH_2O$ 412.0608).

Iodoacetylphenylalanylphenylalaninamide. IR (microscope) 3371, 3297, 3272, 3209, 1657, 1636, 1546, 753, 702 cm^{-1}; ^1H NMR (360 MHz, DMSO-d_6) 8.41 (d, 1H, *J* 8.3 Hz, NH), 8.12 (d, 1H, *J* 8.3 Hz, NH), 7.28–7.15 (m, 10H, ArH), 7.08 (br s, 1 H), 4.50–4.39 (m, 2H, 2xNHCH Phe), 3.60 (q_{AB}, 2H, *J* 9.9 Hz, ICH_2), 2.97 (2xdd, 2H, *J* 13.8, 5.0, 13.9, 4.5 Hz, CH_2 Phe), 2.81 (dd, 1H, *J* 13.9, 8.8 Hz, CH_2 Phe), 2.69 (dd, 1H, *J* 13.9, 9.4 Hz, CH_2 Phe); ^{13}C NMR (100 MHz, DMSO-d_6) 172.52, 170.4, and 167.3 (qC), 137.7, 137.4, 129.18, 129.14, 128.06, 128.00, and 126.2 (ArC), 54.0 and 53.8 (C_α Phe), 37.5 and 37.4 (CH_2 Phe), 0.5 (ICH_2); MS *m/z* (EI, 70 eV) 479.0697 (1.0) [M^+], 462.0440 (2.0) [M^+-H_3N], 435.0556 (1.4) [M^+-CH_2ON], 147.0684 (19) [C_9H_9ON], 120.0813 (100) [$C_8H_{10}N$]; HRMS *m/z* 479.0697 (M^+) (calculated for $C_{20}H_{22}O_4N_3I$ 479.0706).

Iodoacetylserinylphenylglycinamide. IR (CH_3CN cast) 3436, 3420, 3320, 3202, 1676, 1656, 1628, 1556 cm^{-1}; ^1H NMR (500 MHz, CD_3OD) 7.43 (m, 2H, ArCH) 7.36–7.28 (m, 3H, ArCH), 5.41 (s, 1H, NHCH Phg), 4.40 (t, 1H, *J* 5.4 Hz, NHCH Ser), 3.81–3.71 (m, 3H, *J* 5.4 Hz, CH_2 Ser and ICH_2), 3.73 (dd, 1H, *J* 5.4 Hz, CH_2 Ser); ^{13}C NMR (125 MHz, from {^1H, ^{13}C}–HMQC ^{13}C dimension) 129.2 (ArC), 128.0 (ArC), 61.9 (CH_2 Ser), 58.0 (C_α Phg), 56.7 (C_α Ser); MS *m/z* (Fab$^+$, Cleland) 405.9 (2.9) [MH^+], 428.0 (3.7) [MNa^+]; HRMS *m/z* 361.0024 (M^+) (calculated for $C_{13}H_{16}O_4N_3I-CH_2ON$: 361.0049).

Iodoacetylserinylcyclohexylalaninamide. mp 173°C (dec); IR (CH_3CN cast) 3370, 3296, 3210, 3065, 1636, 1585, 1545 cm^{-1}; ^1H NMR (360 MHz, CD_3 OD) 4.45–4.38 (m, 2H, NHCH Ser and Cha), 3.83–3.75 (m, 3H, ICH_2 and CH_2 Ser), 3.72 (dd, 1H, *J* 10.6, 6.7 Hz, CH_2 Ser), 1.80 (brd, 1H, *J* 12.8 Hz, CH Cha), 1.74–1.52 (m, 6H, CH_2 Cha), 1.48–1.10 (m, 4H, CH_2 Cha), 1.05–0.82 (m, 2H, CH_2 Cha); ^{13}C NMR (100 MHz, CD_3OD) 177.8, 172.4, and 171.5 (qC), 62.8 (CH_2 Ser), 57.0 (C_α Ser), 52.3 (C_α Cha), 40.2, 35.3, 35.0, 33.0, 27.6, 27.4, and 27.1 (CH_2 and CH Cha), −2.2 (ICH_2); MS *m/z* (EI, 70 eV) 389.0355 (1.1) [M^+-CH_2O], 381.0661 (8.0) [M^+-CH_2ON], 126.1281 (100) [$C_8H_{16}N$], 60.0454 (27) [C_2H_6ON]; HRMS *m/z* 395.0710 (M^+-CH_2O) (calculated for $C_{14}H_{24}O_4N_3I-CH_2O$: 395.0706).

Iodoacetylhomoserinylphenylalaninamide. mp 164°C (dec); IR (CH_3CN cast) 3417, 3289, 1695, 1549, 1540, 1169, 752, 696 cm^{-1}; ^1H NMR (400 MHz, CD_3OD) 7.29–7.15 (m, 5H, ArCH), 4.60 (m, 1H, NHCH HSe), 4.35 (m, 1H, NHCH Phe), 3.70 (q_{AB}, 2H, *J* 10.0 Hz, ICH_2), 3.56 (m, 2H, CH_2 HSe), 3.18 (dd, 1H, *J* 14.0, 5.6 Hz, CH_2 Phe), 2.93 (dd, 1H, *J* 14.0, 8.8 Hz, CH_2 Phe), 1.93 (m, 1H, CH_2 HSe), 1.74 (m, 1H, CH_2 Hse); ^{13}C NMR (100 MHz, CD_3OD) 175.9, 173.5, and 171.5 (qC), 138.4, 130.3, 129.5, and 127.8 (ArC), 59.2 (C_α HSe), 55.6 (C_α Phe), 52.9 (CH_2 HSe), 38.6 (CH_2 Phe), 35.4 (CH_2 HSe), −2.3 (ICH_2); MS *m/z* (EI, 70 eV) 389.0355 (4.5) [M^+-CH_2ON], 371.0258 (4.1) [$M^+-CH_2ON-H_2O$], 261.1241 (2.2) [M^+-CH_2ON-HI], 147.0684 (26) [C_9H_9ON], 120.0811 (100) [$C_8H_{10}N$], 91.0549 (31) [C_7H_7], 74.0604 (50) [C_3H_8ON]; HRMS *m/z* 389.03550 (M^+-CH_2ON) (calculated for $C_{15}H_{20}O_4N_3I-CH_2ON$ 389.03622).

Iodoacetylvalinylphenylalaninamide. mp 219°C (dec); IR (microscope) 3396, 3281, 3202, 1675, 1641, 1632, 700 cm^{-1}; ^1H NMR (300 MHz, CD_3OD) 7.28–7.15 (m, 5H, ArH), 4.61 (dd, 1H, *J* 8.7, 5.8 Hz,

NHCH Phe), 4.06 (d, 1H, *J* 7.2 Hz, NHCH Val), 3.79 and 3.68 (2xd, 1H, *J* 9.9 Hz, ICH_2), 3.14 (dd, 1H, *J* 13.9, 5.8 Hz, CH_2 Phe), 2.91 (dd, 1H, *J* 13.9, 8.8 Hz, CH_2 Phe), 2.02 (apparent sextet, 1H, *J* 7.2, 6.7 Hz, CH_2 Val), 0.86 (dd, 6 H, *J* 6.7 Hz, 2xCH_3 Val); ^{13}C NMR (100 MHz, CD_3OD/DMSO-d_6 4:1) 175.0, 172.3, and 170.6 (qC), 138.4, 130.1, 129.2, and 127.4 (ArC), 60.2 (C_α Phe), 55.1 (C_α Val), 38.5 (CH_2 Phe), 31.7 (CH_2 Val), 19.5 and 18.3 (CH_3 Val), −1.3 (ICH_2); MS (EI, 70 eV) *m/z* 431.0707 (1.7) [M^+], 414.0448 (2.6) [M^+-H_3N], 387.0568 (5.9) [M^+-CH_2ON], 147.0684 (34) [C_9H_9ON], 120.0813 (100) [$C_8H_{10}N$], 91.0547 (21) [C_7H_7], 72.0814 (75) [$C_4H_{10}N$]; HRMS *m/z* 431.0707 (M+) (calculated for $C_{22}H_{22}O_3N_3I$ 431.0706).

Iodoacetylhistidylphenylalaninamide. IR (microscope) 3276 (br), 2803 (br), 1671, 1640, 1545, 1192, 1132 cm^{-1}; ^1H NMR (360MHz, CD_3OD) 8.75 (d, 0.5H, *J* 1.3 aromatic His H), 7.29–7.17 (m, 6H, ArH), 4.66–4.59 (m, 2H, NHCH His and NHCH Phe), 3.69–3.62 (2d, 2H, *J* 9.9 Hz, ICH_2), 3.21–3.12 (m, 2H, CH_2 Phe and CH_2 His), 3.06–3.00 (dd, 1H, *J* 15.4, 7.6 Hz, CH_2 His), 2.96–2.90 (dd, 1H, *J* 14.0, 9.0 Hz, CH_2 Phe); ^{13}C NMR (150 Mhz, CD_3OD, from {^1H, ^{13}C}−HMQC ^{13}C dimension) 135 (His ArC), 130 (ArC), 128 (ArC), 118 (ArC), 56 and 54 (C_α Phe and C_α His), 49 (CH_2 Phe), 26 (CH_2 His), −2 (ICH_2); MS *m/z* (ESI-MS) 492 (4) [MNa^+], 470.0 (87) [MH^+].

Iodoacetylglycylphenylalaninamide. IR (microscope) 3404, 3298, 3192, 1673, 1645, 1619, 1546 cm^{-1}; ^1H NMR (300 MHz, CD_3OD) 7.30–7.15 (m, 5H, ArH), 4.63–4.58 (dd, 1H, *J* 8.9, 5.4 Hz, NHCH Phe), 3.86–3.69 (m, 4H, $NHCH_2$ Gly and ICH_2), 3.21–3.13 (dd, 1H, *J* 13.9, 5.5 Hz, CH_2 Phe), 2.95–2.80 (dd, 1H, *J* 13.9, 8.9 Hz, CH_2 Phe); ^{13}C NMR (100 MHz, CD_3OD) 176.0, 171.9, and 170.5 (qC), 138.4, 130.3, 129.5, and 127.7 (ArC), 55.9 (C_α Phe), 44.0 (C_α Gly), 38.8 (CH_2 Phe), −2.6 (ICH_2); MS *m/z* (ESI-MS) 412.0 (100) [MNa^+].

Iodoacetyltyrosylphenylalaninamide. IR (microscope) 3337, 3277, 3189, 1652, 1631, 1538, 1515, 1438 cm^{-1}; ^1H NMR (360 MHz, CD_3OD) 7.28–7.16 (m, 5H, ArCH Phe), 7.01–6.99 (dd, 2H, *J* 6.6, 2.0 Hz, ArCH Tyr), 6.69–6.65 (dd, 2H, *J* 6.6, 2.1 Hz, ArCH Tyr), 4.57–4.53 (dd, 1H, *J* 8.5, 5.7 Hz, NHCH Tyr) 4.47–4.43 (dd, 1H, *J* 8.5, 6.2 Hz, NHCH Phe) 3.67–3.59 (2d, 2H, *J* 10.0 Hz, ICH_2), 3.15–3.10 (dd, 1H, *J* 13.9, 5.65 Hz, CH_2 Phe), 2.95–2.86 (m, 2H, CH_2 Tyr and CH_2 Phe), 2.75–2.69 (dd, 1H, *J* 14.0, 8.5 Hz, CH_2 Tyr), ^{13}C NMR (75 MHz, CD_3OD+DMF-d_7) 175.6, 172.2, and 170.0 (qC), 138.7, 131.4, 130.4, 129.4, 128.7, 127.6, and 116.2 (ArC), 56.6 and 55.5 (C_α Phe and Tyr), 38.7 and 38.8 (CH_2 Tyr and Phe), −1.9 (ICH_2); MS *m/z* (ESI-MS) 518 (100) [MNa^+].

Iodoacetylglutaminylphenylalaninamide. IR (microscope) 3396, 3263, 1671, 1639, 1540 cm^{-1}; ^1H NMR (360 MHz, CD_3OD) 7.28–7.15 (m, 5H, ArH), 4.61–4.57 (dd, 1H, *J* 8.8, 5.6 Hz, NHCH Phe), 4.25–4.21 (dd, 1H, *J* 8.1, 6.2 Hz, NHCH Gln), 3.74–3.65 (2xd, 2H, *J* 10.0 Hz, ICH_2), 3.19–3.13 (dd, 1H, *J* 13.8, 5.6 Hz, CH_2 Phe), 2.96–2.90 (dd, 1H, *J* 13.8, 8.8 Hz, CH_2 Phe), 2.23–2.18 (m, 2H, CH_2 Gln), 2.01–1.81 (m, 2H, CH_2 Gln); ^{13}C NMR (125 MHz, CD_3OD/DMF-d_7) 176.6, 175.1, 172.5, and 170.2 (qC), 139.9, 130.4, 129.4, and 127.6 (ArC), 55.5 and 54.7 (C_α Phe and Gln), 38.7 (CH_2 Phe), 32.4 (CH_2 Gln), 28.9 (CH_2 Gln), −0.6 (ICH_2); MS *m/z* (ESI-MS) 483.0 (100) [MNa^+].

Iodoacetylarininylphenylalaninamide. IR (microscope) 3276, 1630, 1540, 1427 cm^{-1}; ^1H NMR (600 MHz, CD_3OD) 7.30–7.18 (m, 5H, ArCH Phe), 4.66–4.60 (dd, 1H, *J* 7.8, 2.2 Hz, NHCH Phe) 4.29–4.26 (dd, 1H, *J* 7.8, 6.0 Hz, NHCH Arg), 3.76–3.68 (2xd, 2H, *J* 9.88 Hz, ICH_2), 3.18–3.12 (m, 3H, CH_2 Phe and CH_2 Arg), 2.95–2.89 (dd, 1H, *J* 14.1, 8.9 Hz, CH_2 Phe), 1.61–1.53 (m, 1H, CH_2 Arg), 1.40–1.20 (m, 1H, CH_2 Arg), 1.17–1.00 (m, 2H, CH_2 Arg); ^{13}C NMR (150 Mhz, CD_3OD, from {^1H, ^{13}C}−HMQC ^{13}C dimension) 130.0 and 128.0 (ArC), 56.0 (C_α Phe), 55.0 (C_α Arg), 41.8 (CH_2 Arg), 38.1 (CH_2 Phe), 30.0 (CH_2 Arg), 25.9 (CH_2 Arg), −2.2 (ICH_2); MS *m/z* (ESI-MS) 489.0 (100) [MH^+].

Iodoacetylasparginylphenylalaninamide. IR (microscope) 3147, 3284, 1656, 1547, 1408 cm^{-1}; ^1H NMR (360 MHz, DMSO-d_6) 8.41 (d, 0.6H, *J* 7.9 Hz, NH), 8.13 (d, 0.6H, *J* 8.2 Hz, NH), 7.42 (brs, 1.5H, NH), 7.26–7.13 (m, 4H, ArH), 6.97 (brs, 0.5H, NH), 4.50–4.41 (m, 1H, NHCH), 4.35–4.28 (m, 1H, NHCH), 3.68–3.62 (Abq, 2H, *J* 9.9 Hz, ICH_2), 3.15–3.08 (dd, 1H, *J*, CH_2), 2.82–2.75 (dd, 1H, *J*, CH_2), 2.50–2.47 (dd, 1H, *J*, CH_2), 2.37–2.30 (dd, 1H, *J*, CH_2); ^{13}C NMR (125 MHz, DMSO-d_6) 172.7, 171.5, 170.2, and 167.3 (qC), 138.2, 128.9, 128.1, and 126.1 (ArC), 54.0 (C_α Phe), 49.9 (C_α Asn), 36.7 (CH_2 Phe), 30.6 (CH_2 Asn), 0.5 (ICH_2); MS *m/z* (ESI-MS) 469.0 (100) [MNa^+].

Mixtures were synthesized analogously, using equimolar mixtures of resins bearing different peptides for the iodoacetylation step.

4.1.2. Assay of protease activity

The activity of inhibitors was measured by monitoring the rate changes of HAV 3C proteinase activity in the presence of inhibitor as previously described [10, 12].

4.1.3. Preparation of HAV 3C-inactivator complexes

HAV 3C proteinase [16] was dialyzed against 50 mM potassium phosphate (pH 7.5), 1 mM EDTA to remove DTT using Amicon (Boston, MA) centricon ultrafiltration units. Dialyzed HAV 3C proteinase (4 nmol, 192 μg/mL) was incubated with mixtures of *N*-iodoacetyldipeptides (10 equivalents of (each) inhibitor, 8 mM in dimethylformamide) at 25°C for 1 h with mixing. The HAV 3C/acetyldipeptide amide complexes were then precipitated with 1.2 volumes of cold acetone at −20°C for 1 hour and recovered by centrifugation (15 min at 14,000 *g*). The residue was dried *in vacuo* and redissolved in a small amount of 25% aqueous acetic acid. Alternatively, the complexes were dialyzed with 100 mM ammonium acetate buffer (pH 8.0) to a volume of approximately 150 μL, using an Amicon centricon ultrafiltration unit, for use in enzymatic digestion experiments.

4.1.4. Mass spectrometry

An aliquot of the aqueous acetic acid solution was subjected to ESI-MS on a Fisons VG Quattro instrument (Micromass, Beverly, MA) with sample introduction as a 5% solution in 1% aqueous CF_3CO_2H. Raw spectra were smoothed for mass-to-change ratios ranging from 500 to 1500 and then deconvoluted. Native HAV 3C C24S samples were found to give an accuracy of 23,877 ± 2.5 Da. The value of 23,877 Da, which is the same as the mass calculated form the amino acid sequence for HAV 3C C24S, was taken as such for calculation of the Dm value. A tolerance of ±2.5 Da of the adduct mass was then assumed for the determination of the best inhibitor.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the U.S. National Institutes of Health. The authors thank Manjinder Lall, Robert S. Parker, and Yunlong Pu for experimental assistance.

References

[1] For reviews see (a) S. H. Dewitt, A. W. Czarnik, Acc. Chem. Res. 29 (1996) 114–122. (b) J. A. Ellman, Acc. Chem. Res. 29 (1996) 132–143. (c) E. M. Gordon, M. A. Gallop, D. V. Patel, Acc. Chem. Res. 29 (1996) 144–154. (d) L. A. Thompson, J. A. Ellman, Chem. Rev. 96 (1996) 555–600. (e) J. N. Abelson, (Ed.) Methods Enzymol. 267 (1996) 3–465. (f) N. K. Terrett, M. Gardner, D. W. Gordon, R. J. Kobylecki, J. Steele, Tetrahedron 30 (1995) 8135–8173.
[2] J. S. Früchtel, G. Jung, Angew. Chem. Int. Ed. 35, 17–42.
[3] For a recent example see S. Cheng, D. Comer, J. P. Williams, P. L. Myers, D. L. Boger, J. Am. Chem. Soc. 118 (1996) 2567–2573.
[4] R. F. Service, Science 272 (1996) 1266–1268.
[5] K. D. Janda, Proc. Natl. Acad. Sci. USA 91 (1994) 10779–10785.
[6] (a) C. L. Brummel, J. C. Vickerman, S. A. Carr, M. E. Hemling, G. D. Roberts, W. Johnson, J. Weinstock, D. Gaitanopoulos, S. J. Benkovic, N. Winograd, Anal. Chem. 68 (1996) 237–242. (b) Y.-H. Chu, Y. M. Dunayevskiy, D. P. Kirby, P. Vouros, B. L. Karger, J. Am. Chem. Soc. 118 (1996) 7827–7835.
[7] X. Cheng, R. Chen, J. E. Bruce, B. L. Schwartz, G. A. Anderson, S. A. Hofstadtler, D. C. Gale, R. D. Smith, J. Gao, G. B. Sigal, M. Mammen, G. M. Whitesides, J. Am. Chem. Soc. 117 (1995) 8859–8860.
[8] For reviews see (a) B. A. Malcolm, Protein Sci. 4 (1995)

1439–1445. (b) W. G. Dougherty, B. L. Semler, Microbiol. Rev. 57 (1993) 781–822. (c) H. G. Kräusslich, E. Wimmer, Annu. Rev. Biochem. 57 (1988) 701–754.

[9] S. W. Kaldor, M. Hammond, B. A. Dressman, J. M. Labus, F. W. Chadwell, A. D. Kline, B. A. Hunt, Bioorg. Med. Chem. Lett. 5 (1995) 2021–2026.

[10] B. A. Malcolm, S. M. Chin, D. A. Jewell, J. R. Stratton-Thomas, K. B. Thudium, R. Ralston, S. Rosenberg, Biochemistry 31 (1992) 3358–3363.

[11] B. A. Heinz, J. Tang, J. M. Labus, F. W. Chadwell, S. W. Kaldor, M. Hammond, Antimicrob. Agents Chemother. 40 (1996) 267–270.

[12] T. S. Morris, S. Frormann, S. Shechosky, C. Lowe, M. S. Lall, V. Gauss-Müller, R. H. Purcell, S. U. Emerson, J. C. Vederas, B. A. Malcolm, Bioorg. Med. Chem. 5 (1997) 797–807.

[13] D. A. Jewell, W. Swietnicki, B. M. Dunn, B. A. Malcolm, Biochemistry 31 (1992) 7862–7868.

[14] J. R. Petithory, F. R. Masiarz, J. F. Kirsch, D. V. Santi, B. A. Malcolm, Proc. Natl. Acad. Sci. USA 88 (1991) 11510–11515.

[15] (a) E. M. Bergmann, M. N. G. James, personal communication, 1997. (b) For crystal structure details for HAV 3C see E. M. Bergmann, S. C. Mosimann, M. M. Chernaia, B. A. Malcolm, M. N. G. James, J. Virol. 71 (1997) 2436–2448.

[16] H. Rink, Tetrahedron Lett. 28 (1987) 3787–3890.