Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7129–7133, October 1988
Biochemistry

Chemical synthesis and enzymatic activity of a 99-residue peptide with a sequence proposed for the human immunodeficiency virus protease

(retrovirus/aspartic protease/protein synthesis/solid-phase synthesis)

RUTH F. NUTT, STEPHEN F. BRADY, PAUL L. DARKE, TERRENCE M. CICCARONE, C. DYLION COLTON, ELKA M. NUTT, JOHN A. RODKEY, CARL D. BENNETT, LLOYD H. WAXMAN, IRVING S. SIGAL, PAUL S. ANDERSON, AND DANIEL F. VEBER

Departments of Medicinal Chemistry, Biochemistry and Molecular Biology, Merck Sharp & Dohme Research Laboratories, West Point, PA 19486

Communicated by Samuel J. Danishefsky, June 13, 1988

ABSTRACT **Retroviral proteins, including those from the human immunodeficiency virus (HIV), are synthesized as polyprotein precursors that require proteolytic cleavage to yield the mature viral proteins. A 99-residue polypeptide, encoded by the 5′ end of the *pol* gene, has been proposed as the processing protease of HIV. The chemical synthesis of the 99-residue peptide was carried out by the solid-phase method, and the isolated product was found to exhibit specific proteolytic activity upon folding under reducing conditions. Upon size-exclusion chromatography, enzymatic activity was eluted at a point consistent with a dimeric molecular size. Specificity was demonstrated by the cleavage of the natural substrate HIV gag p55 into gag p24 and gag p17, as well as cleavage of small peptide substrates representing processing sites of HIV fusion proteins. The proteolytic action of the synthetic product could be inhibited by pepstatin, an aspartic protease inhibitor.**

A key step in the maturation of retroviruses is the posttranslational cleavage of polyprotein fusions into their constituent functional proteins (1). Proteases responsible for this processing step have been characterized for several retroviruses and are thought to belong to the aspartic protease group of enzymes (2, 3). For the human immunodeficiency virus (HIV), a peptide sequence of an analogous protease is encoded at the N-terminal portion of the pol precursor polypeptide (4). The N- and C-termini of this protease may be surmised from the appearance of two sites at positions 69 and 167 of the pol reading frame that share sequence similarities with known retroviral protease cleavage sites.

Studies with murine leukemia virus have implicated an essential role for the protease of that virus in maintaining its infectivity (5). In addition, it was recently observed that a mutation in the HIV genome, representing a replacement of the proposed active site Asp-25 with Asn, eliminated the infectivity of the virus (6). These findings substantiate the current interest in HIV protease inhibitors as potential therapeutic agents in the treatment of acquired immunodeficiency syndrome (7, 8). Exploration of this potential approach to therapy necessitates the availability of enzyme in amounts sufficient to establish screening assays for inhibitors and to fully characterize the enzymology. In view of the health hazards associated with handling live virus, isolation of adequate amounts of enzyme from native sources is precluded, and alternative approaches such as microbial expression or chemical synthesis must be considered. The biogenetic expression of the HIV protease has been described recently (9, 10), but the isolation and characterization of the enzyme have not been reported. Total chemical synthesis can be envisioned as an alternative approach to rapidly obtaining this protease in useful quantities. A specific advantage of the synthetic route would be that any proteolytic activity would be intrinsic and could not be attributable to isolation artifacts.

Few reports exist for the successful chemical synthesis of natural enzymes. The first enzyme syntheses were ribonucleases A and S (11, 12). These syntheses established the viability and limitations of both the solution and solid-phase methods. It was not until 1980 that solution methods were reported to give a totally characterized, crystalline enzyme (13). The synthesis of analogs of large peptides has generally been precluded because the methods of synthesis have not been considered sufficiently reliable to produce products that can be characterized by means other than direct comparison with the natural products. In spite of these concerns, recent advances in the speed and fidelity of solid-phase peptide synthesis, improvements in the HF deblocking procedure, and development of powerful purification methods emboldened us to attack the synthesis of the proposed protease sequence in Fig. 1, independent of the availability of the natural product. We report here the total chemical synthesis of the postulated 99-amino-acid HIV protease and characterization of its enzymatic properties.

MATERIALS AND METHODS

t-Butoxycarbonyl (Boc)-Phe-*O*-phenylacetamidomethyl-resin, Boc-protected amino acids in preloaded cartridges, and all solid-phase synthesis reagents were supplied by Applied Biosystems (Foster City, CA). Boc-protected amino acids of side-chain-protected aspartic acid, glutamic acid, cysteine (acetamidomethyl), and histidine were obtained from Bachem Fine Chemicals (Torrance, CA), and the solvents dimethylformamide and CH_2Cl_2 were purchased from Burdick and Jackson (Muskegon, MI).

Peptide-Resin Synthesis. Assembly of the polypeptide chain was carried out by stepwise techniques on solid support by using an Applied Biosystems 430A automated peptide synthesizer (14) starting with 0.50 mmol of Boc-Phe-*O*-phenylacetamidomethyl-resin (0.675 g, substituted at 0.74 mmol of phenylalanine per gram of resin). The following side-chain protection was used: tosyl for arginine, cyclohexyl for aspartic acid and glutamic acid (15), *p*-chlorocarbobenzoxy for lysine, 2-bromocarbobenzoxy for tyrosine, N^{π}-benzyloxymethyl for histidine (16), formyl for tryptophan (17), and benzyl for serine and threonine. Boc removal was accomplished by using one 2-min wash and one 20-min treatment with 65% trifluoroacetic acid in CH_2Cl_2. Neutralization was

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HIV, human immunodeficiency virus; Boc, *t*-butoxycarbonyl.

carried out with three successive 3-min treatments of 10% diisopropylethylamine in dimethylformamide. Boc-protected amino acids were activated by using dicyclohexylcarbodiimide and were introduced either as hydroxybenzotriazole esters [arginine(tosyl), asparagine, glutamine, cysteine(acetamidomethyl), and histidine(benzyloxymethyl)] or as symmetrical anhydrides, preformed in CH_2Cl_2 followed by solvent exchange with dimethylformamide. Amino acids were introduced by using a minimum of two couplings per residue (see Fig. 1). Recouplings were conducted after a neutralization step with diisopropylethylamine in dimethylformamide. Each residue incorporation was followed by a "capping" cycle using acetic anhydride (18) (ca. 5% in dimethylformamide) for a period of 10 min. The resin was treated with diisopropylethylamine/dimethylformamide before the capping step and was washed with CH_2Cl_2 afterwards. A double acetylation step was used after Lys-42 and Leu-38. To retain efficient mixing of the reaction mixture throughout the synthesis, approximately one-third of the peptide-resin was removed at the 74–99 and 17–99 stages of the synthesis. As much as possible, the assembly was continuous, and the time at break points was minimized. The final weight of the fully assembled peptide-resin was 1.76 g. The efficiency of couplings was monitored at selected steps by quantitative ninhydrin assay (19) and by carrying out preview analysis (20) on the peptide-resin by using an Applied Biosystems automated protein sequenator model 470A (Table 1).

HF Cleavage and Isolation of Reduced Product. Removal of the Boc group from Boc-99-residue peptide-resin (318 mg) was effected by treatment with trifluoroacetic acid/CH_2Cl_2, 65:35 (vol/vol) for 2 min and 20 min; the product was washed six times with CH_2Cl_2 and was dried *in vacuo* for 15 min to give 382 mg of peptide-resin. Deblocked peptide-resin was suspended in a mixture of 1.0 ml of *p*-cresol/*p*-thiocresol, 50:50 (vol/vol), 0.5 ml of ethanedithiol, and 7.0 ml of dimethyl sulfide and chilled to −70°C in an HF apparatus (Peninsula Laboratories, San Carlos, CA). The total volume was adjusted to 11.5 ml with condensed HF. After 2.5 hr of stirring at 0–5°C, the HF and dimethyl sulfide were removed by evaporation *in vacuo* with a liquid N_2 trap. The residue was triturated with ether, filtered, and dried *in vacuo* for 10 min to give 322 mg of peptide-resin. The product was then mixed with 0.5 ml of *p*-cresol and 0.5 ml of 1, 4-butanedithiol. The mixture was cooled to −70°C, and 9–10 ml of HF was condensed into the vessel. The reaction mixture was stirred at 0–5°C for 1.5 hr, the HF was evaporated *in vacuo*, and the residue was triturated with ether. After filtration of the precipitate and drying *in vacuo* for 10 min, the solids were triturated with 50% aqueous HOAc for 20 min. The mixture was filtered, and the filtrate was applied to a 5- × 100-cm column of Sephadex G-50 fine and was eluted with 50% HOAc. Elution of the product was monitored at 254 and 280 nm. Fractions showing a major component on reverse-phase HPLC were combined and concentrated to a volume of about 20 ml, which was then applied to a 5- × 100-cm column of Sephadex G-75 fine, which was eluted with 50% HOAc by gravity feed at a flow rate of 12–15 ml/hr (Fig. 2). Fractions showing both a main component at 15 min by reverse-phase HPLC (Vydac C_4, 30–50% CH_3CN over 30 min) and a major band of molecular mass of ≈10 kDa by sodium dodecyl sulfate (SDS)/PAGE (21) were combined to give 20 mg [by *A* at 280 nm by using $\varepsilon^{Trp+Tyr} = 12,500$ (22)] of product. Product was analyzed for amino acid ratios after hydrolysis with 6 M HCl (Table 2) and for amino acid sequence before and after treatment with cyanogen bromide (Table 1).

Folding to Active Enzyme. Unfolded protein (100 μl of 50% HOAc containing 30 μg of protein as measured by UV absorbance at 280 nm) was introduced into a 12-mm diameter cellulose dialysis bag (molecular weight cutoff of 1000; Spectrum Medical Industries, Los Angeles) that contained 3 mg of bovine serum albumin (Sigma no. A 7638) in 200 μl of dialysis buffer (23). The buffer consisted of 0.05 M NaOAc, 10^{-3} M dithiothreitol, 10^{-4} M Na_2EDTA (Fisher), 10% glycerol, and 5% ethylene glycol adjusted to pH 5.5 with HOAc. Dialysis was carried out in 50 ml of buffer at 0–5°C for 2 hr with a change in buffer three times to retain a constant pH of 5.5. Samples were analyzed for enzymatic activity by adding 10 μl (1 μg of enzyme) of reaction solution to 10 μl of solution containing the synthetic substrate Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val (2 mg/ml in 0.05 M NaOAc buffer at pH 5.5). After incubation at 30°C for 30 min, the reaction was quenched with 80 μl of 12% HOAc. The final solution was analyzed by HPLC for starting peptide (2.4 min) and cleavage products (N-terminal pentapeptide and C-terminal tripeptide at 1.5 and 1.9 min, respectively) on a 5-cm Vydac (Hesperia, CA) C_{18} column eluted at a flow rate of 5 ml/min with the following gradient using 0.1% trifluoroacetic acid in H_2O (solvent A) and 0.1% trifluoroacetic acid in CH_3CN (solvent B): 0 min, 0% solvent B; 2.6 min, 20% solvent B. Products were quantified by comparing integrated peak areas with synthetic standards. Specific activities were expressed as nmol of products formed per min per mg of enzyme under the assay conditions described.

Gel Filtration Chromatography of Folded Protease. A 12-ml sample containing 1 mg (as measured by UV absorbance at 280 nm) of folded protease (as described above but containing 0.1% bovine serum albumin) with a specific activity of 230 nmol/min per mg was chromatographed at 4°C on a 2.5- × 75-cm column packed with Sephadex G-75 medium (40–120 μm beads). The eluting buffer contained 50 mM Mes (sodium salt) (pH 5.5), 1 mM dithiothreitol, 1 mM EDTA, 10% glycerol, and 5% ethylene glycol. Activity assays of gel filtration fractions contained 0.1% bovine serum albumin in addition to what was in the sample assayed. Total applied activity was recovered in the eluate. A peak specific activity of 640 nmol/min per mg was measured in tube 8 (Fig. 3).

Digestion of Isolated p55 by the HIV Protease. HIV gag p55 was expressed in yeast and was purified to homogeneity (M. Polokoff and G. Vlasuk, personal communication). Purified soluble protein (20 μg) was incubated with 0.2–0.5 μg of folded synthetic protease or the purified protease from the avian myeloblastosis virus in a volume of 50 μl, which contained 1 mM dithiothreitol, 0.1 mM EDTA, and 50 mM sodium acetate at pH 5.5. The digestion was carried out for 3–8 hr at room temperature (25% completion at 8 hr). Samples were prepared for SDS/PAGE by adding SDS to a final concentration of 1% and 2-mercaptoethanol to 50 mM and heating the mixture at 90°C for 5 min. Electrophoresis was carried out on 12.5% or 15% acrylamide gels (21), and proteins were electroblotted onto Immobilon-P transfer membranes (Millipore) (24). The protein bands were located by staining the membranes with Coomassie blue and were cut out for sequence analysis.

RESULTS AND DISCUSSION

Initial synthetic studies focused on the preparation of the 99-residue peptide as the disulfide-linked monomer. The general method chosen was the solid-phase method of Merrifield as modified to use the chemically more stable phenylacetamidomethyl-resin (for reviews, see refs. 14, 18, and 25). HF-labile side-chain protection was employed except for acetamidomethylcysteine (26), which was chosen to facilitate purification at an intermediate stage of the synthesis. Early in the peptide-resin synthesis, incomplete couplings were observed by ninhydrin analysis. In particular, Arg-87 could only be incorporated to the extent of 91%, even after four couplings. To minimize generation of deletion sequences arising from incomplete couplings, capping of N termini with acetic anhydride was included at the end of each amino acid

incorporation (18). Monitoring by the ninhydrin method substantiated termination of unreactive amino groups by acetylation. Comparison of the cleaved 26-residue peptide prepared by either the noncapping or the capping protocol indicated higher purity and more facile isolation for material synthesized by the latter route. The peptide-resin synthesis using the capping procedure was also evaluated at the 52- and 67-residue stages after cleavage with HF. The crude products in each case contained a major component that could be isolated by HPLC to yield material with the correct amino acid composition. Although the final 99-residue product was cleaved from the resin by the S_N2/S_N1 method developed by Tam *et al.* (27), incomplete removal of the formyl group from tryptophan was observed as indicated by sequence analysis and UV spectroscopy (28). A further problem was seen in the exceptionally slow removal of the acetamidomethyl group with iodine. Under conditions required to achieve complete acetamidomethyl removal, partial oxidation of methionine occurred, which proved difficult to reverse (29). Despite the synthetic problems encountered, these initial studies led to enzymatically active product after reduction.

As a result of the synthetic problems identified in the first synthesis, the modified approach described in detail in *Materials and Methods* was devised. It incorporated the HF removable *p*-methylbenzyl group (30) as cysteine protection, affording final product in fully reduced form. In the peptide-resin assembly, extra couplings were introduced at previously recognized difficult points (Fig. 1). In almost every case, capping with acetic anhydride was shown to cover more than 99.4% of the available amino groups detectable by ninhydrin. Sequence analysis of the 99-residue peptide-resin after removal of the terminal Boc-protecting group showed cumulative preview of 7% within the N-terminal 47 residues. Sequence analysis is a sensitive method for detection of deletions in solid-phase synthesis because it is additive at each step. It therefore reflects the total amount of impurity from failures of coupling and deprotection (20). An additional 8% preview was detected at residue 53 (Table 1). The cause of this extra preview is not clear, but it may be related to the peptide-resin having been stored for a time at the 52–99 stage of the synthesis (14). Sequence analysis carried out for 82 cycles also confirmed the accuracy of synthesis in the N-terminal region.

To ensure complete formyl group removal from tryptophan in the two-step HF reaction, 1, 2-ethanedithiol was added in the S_N2 step, and thiocresol was replaced with 1, 4-butanedithiol in the S_N1 step. These modifications avoided potential side reactions arising from treatment of tryptophan(formyl) peptides with thiocresol under S_N1-type HF reaction conditions (27). While no effort was made to determine at which stage the complete removal of tryptophan protection occurred, successful deprotection with 1, 4-butanedithiol under S_N1 conditions has been described previously (28). The crude product was immediately purified by gel filtration with 50% aqueous acetic acid as eluant to prevent protein aggregation, preclude absorption effects, and minimize oxidation of thiol (Fig. 2). The product was characterized at this point for structure and purity. Amino acid analysis after a 110-hr acid hydrolysis showed ratios that were within 6% of expected values (Table 2). Sequence analysis carried out for 65 cycles showed the product to be of correct structure in the N-terminal region and also showed no evidence of other identifiable by-products. Cumulative preview for 35 cycles was 7.3%, which is indicative of small amounts of deletion sequences arising from incomplete couplings in the N-terminal region (Table 1). At cycle 6, the phenyl thiohydantoin of tryptophan eluted at 21.99 min, with no evidence of tryptophan(formyl), which gives a peak at 22.91 min. The UV spectrum was also consistent with unprotected tryptophan (28). Sequence analysis after cyanogen bromide cleavage resulted in the expected three fragments starting with Pro-1,

Table 1. Sequence preview analysis of synthetic products before and after cleavage from the solid support

Peptide-resin			Free peptide*		
Cycle	Amino acid	Preview, %	Cycle	Amino acid	Preview, %
4	Leu	4.5	2	Ile	0
9	Leu	4.4	4	Leu	0
10	Val	6.4	9	Leu	0.5
12	Ile	4.2	10	Val	4.0
21	Ala	6.0	13	Lys	4.0
27	Ala	7.4	18	Leu†	12.9
32	Leu	7.0	21	Ala	4.6
46	Ile	7.4	31	Val	6.6
52	Phe	15.3	32	Leu	4.3
55	Val	15.5	35	Met	7.3

*After gel chromatography.
†Incomplete separation from phenylthiohydantoin of lysine.

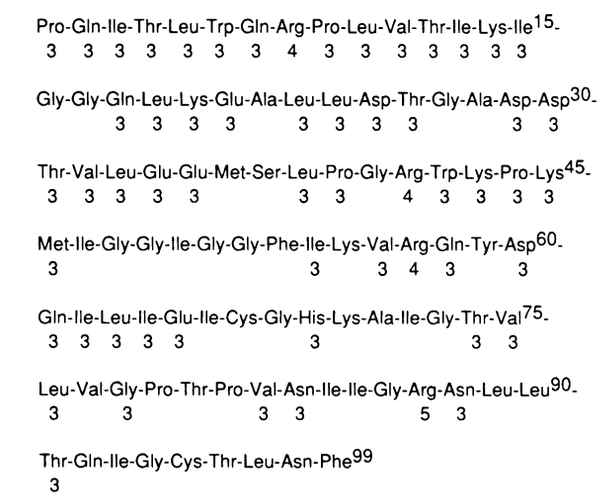

FIG. 1. Proposed amino acid sequence of the HIV protease as synthesized (sequence from ref. 4). Numbers below amino acid residues denote number of couplings used for residue incorporation into peptide-resin. Two couplings were used for residues with no numbers.

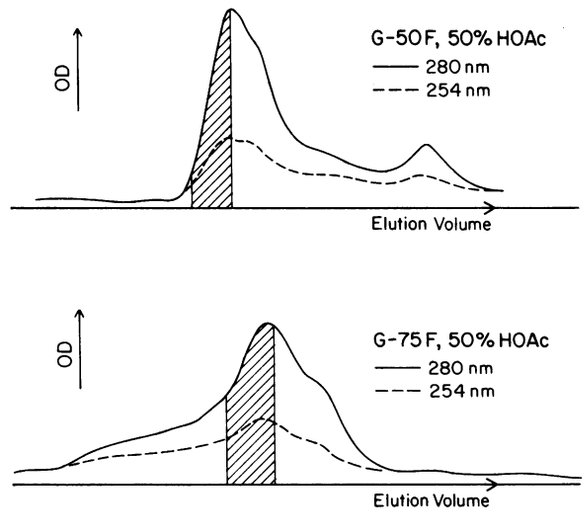

FIG. 2. Gel-filtration chromatography of the synthetic enzyme. Fractions from the hatched area from the Sephadex G-50 fine column (*Upper*) were combined and purified on a Sephadex G-75 fine column (*Lower*). Fractions from the hatched area from the Sephadex G-75 fine column were used for characterization and folding experiments.

Table 2. Amino acid analysis of synthetic HIV protease

Amino acid	Number		Amino acid	Number	
Asp	7.33	(7)	Leu	12.10	(12)
Thr*	7.67	(8)	Tyr	1.03	(1)
Ser*	1.06	(1)	Phe	2.02	(2)
Glu	10.13	(10)	His	1.02	(1)
Gly	13.41	(13)	Lys	5.99	(6)
Ala	3.07	(3)	Arg	3.88	(4)
Val	6.09	(6)	Pro	5.64	(6)
Met†	1.23	(2)	Cys‡	2.34	(2)
Ile + allo-Ile	12.56	(13)			

Amino acid analysis was determined after hydrolysis with 6 M HCl at 100°C for 110 hr. The theoretical number of amino acids is in parentheses.
*Corrected for decomposition during hydrolysis.
†Uncorrected.
‡Determined as cysteic acid after performic acid oxidation.

Ser-37, and Ile-47. Edman degradation, however, indicated a continuation of sequence past Met-46 to the extent of about 10%, which could be interpreted as evidence for the presence of as much as 10% methionine oxide (not subject to CNBr cleavage) in the final product. A small amount of a peptide fragment starting with Lys-43 was also observed. This fragment may have been generated by chemical cleavage at Trp-42 during the strongly acidic treatment of the protein in the presence of sulfoxide (i.e., methionine oxide) (31). HPLC analysis was consistent with a major component or group of closely related peptides. Analysis of product by SDS/PAGE showed a main protein band at 10 kDa with trace amounts of by-products of lower molecular size.

Initial attempts to fold this linear sequence into a native and enzymatically active form showed an increase in enzymatic activity upon folding in the pH range of 4.5–6.5 and in the presence of a thiol (glutathione or dithiothreitol). EDTA was added to the folding medium to prevent metal-catalyzed oxidation of thiols. Highest yields were obtained in a protein concentration range of 100–300 μg/ml. Highest enzymatic activity was obtained in about 1 hr, and loss of activity occurred upon further dialysis over a period of 24 hr. As shown in Table 3, the active product was significantly stabilized by the addition of ethylene glycol and glycerol to the folding medium. Optimal yields and stability were achieved when the solution was also made either 1% or 0.1% in bovine serum albumin.

When a protein sample, folded under optimal conditions, was applied to a Sephadex G-75 medium column under non-denaturing conditions (Fig. 3), all of the enzymatically active product was eluted at a point consistent with a molecular mass of about 20 kDa, which suggests a dimeric structure. Inactive protein was eluted later at a volume consistent with a monomer. Gel electrophoresis under denaturing conditions (SDS) showed the protein in the active fractions to migrate as a monomer of molecular mass 10 kDa. This observation may support proposals relating this enzyme to much higher molecular mass acid proteases through dimerization (3, 8). Folding and gel chromatography of the reduced polypeptide were performed in the presence of 1 mM dithiothreitol, making it likely that dimerization does not involve disulfide linkage, but this possibility has not been ruled out.

Two lines of evidence have been used to establish the specificity and activity of the synthetically produced protein. First, it catalyzes the hydrolysis of the octapeptide Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val, which corresponds to the cleavage site between p17 and p24 in the HIV polyprotein gag (residues 128–135). The cleavage is specific between the tyrosine and proline, giving only a penta- and a tripeptide, both of which were identified by independent synthesis (P.L.D., R.F.N, S.F.B, V. M. Garsky, T.M.C., C.-T. Leu, P. K. Lumma, R. M. Freidinger, D.F.V., and I.S.S., unpublished observations). A pH optimum range for cleavage of this substrate was found to be 4.5–5.5, and a K_m of 2.5 mM has been determined. A minimum substrate of seven residues has been found in this sequence (P.L.D. *et al.*, unpublished observations). In addition, the synthetic enzyme specifically cleaves synthetic peptides representing seven of the nine known *in vivo* proteolytic processing sites within the precursor polyproteins. These HIV protease cleavage sites occur after residues 132, 363, 377, and 448 of gag and after residues 68, 167, and 727 of pol and account for all of the cleavages for the gag and pol precursors (P.L.D. *et al.*, unpublished observations). Secondly, studies of cleavage of a natural substrate, gag p55, expressed in yeast, have also shown the synthetic

Table 3. Effect of time and medium composition on yield of active enzyme

Dialysis solution composition*			Specific activity†			
EG, %	Glycerol, %	BSA, %	1 hr	2 hr	4 hr	24 hr
0	0	0	300	310	181	58
3	10	0	276	210	190	140
0	0	1	351	363	330	296
0	0	0.1	323	343	362	337
3	10	1	530	550	680	570

BSA, bovine serum albumin; EG, ethylene glycol.
*The dialysis solution also contained 10^{-4} M EDTA and 10^{-3} M dithiothreitol at pH 5.5 (0–5°C). The protein concentration was 100 μg/ml.
†The specific activity is expressed in terms of the nmol of Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val cleaved per min per mg of enzyme.

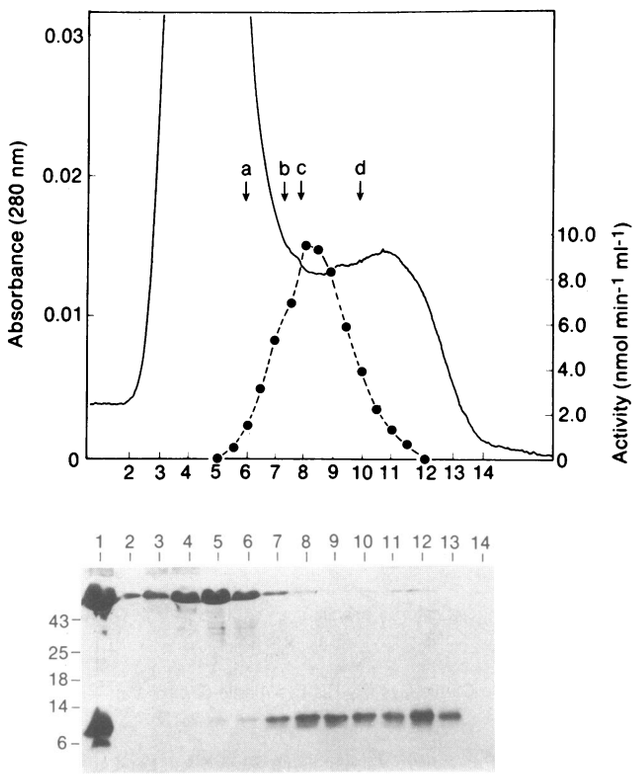

FIG. 3. (*Upper*) Gel chromatography profile (Sephadex G-75 medium) of folded synthetic protease. ——, Protein absorbance at 280 nm; ●---●, enzymatic activity as measured by cleavage of HIV gag residues 128–135. Arrows denote peak elution volumes of molecular size standards: a, ovalbumin (43 kDa); b, carbonic anhydrase (29 kDa); c, α-chymotrypsinogen (25 kDa); d, lysozyme (14 kDa). (*Lower*) SDS/PAGE (16%) on fractions from above gel column stained with silver stain. Molecular size standards (in kDa) are indicated at left.

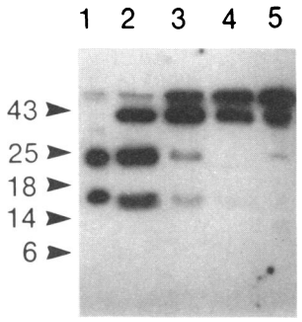

FIG. 4. Cleavage of HIV gag p55 (4 μl of yeast-expressed crude isolate at 1 mg/ml) by synthetic protease after incubation for 60 min at 30°C. Products were analyzed by SDS/PAGE (16%), electroblotted onto nitrocellulose membranes, and incubated with murine monoclonal antibodies (DuPont) against HIV gag p24 and p17. Immunoreactive proteins were detected with ^{125}I-labeled goat anti-mouse antisera (Amersham) followed by autoradiography. Lanes: 1, viral lysate (DuPont); 2, p55 plus 10 μg of enzyme; 3, p55 plus 2 μg of enzyme; 4, p55 plus 0.5 μg of enzyme; 5, p55. An immunoreactive polypeptide at an apparent molecular mass of 43 kDa was observed with different preparations of gag p55 and is probably a breakdown product of the gag p55 generated by endogenous yeast proteases. Sizes (in kDa) are indicated at left.

protein to be effective and specific for the expected cleavage sites. As shown in Fig. 4, crude p55 (expressed in yeast) is cleaved by synthetic enzyme to give a 24-kDa protein and 17-kDa protein, both of which were identified by monoclonal antibodies to the proteins from viral lysates. In addition, when the expressed gag p55 was purified (M. Polokoff and G. Vlasuk, personal communication), it was also cleaved by the synthetic enzyme. After digestion, two major product bands with apparent molecular masses of 24 and 17 kDa could be detected with Coomassie blue. After isolation and sequence analysis, the 24-kDa band had the N-terminal sequence expected for p24 (i.e., Pro-Ile-Val-Gln-Asn-Leu-Gln). The band at 17 kDa had an N-terminal sequence that corresponded to p15 (i.e., Ala-Glu-Ala-Met-Ser-Gln-Val-Thr-Asn-Pro-Ala-Thr-Ile-Met-Ile-Gln). Cleavage is believed also to occur at a Met-Met sequence near the N terminus of p15. However, in the clone used in these experiments, this sequence is Met-Ile and appears not to be hydrolyzed by the protease. A sequence for p17, the N-terminal fragment of p55, was not found. This result is consistent with a blocked N terminus as described for p17 from HIV lysates (32). In contrast to the specific action of the synthetic protease with the HIV sequence, the protease purified from the avian myeloblastosis virus degraded p55 in a nonspecific manner, whereas trypsin completely hydrolyzed p55 to peptides less than 10 kDa.

An additional characteristic property of the synthetic enzyme is the inhibition of cleavage of the octapeptide substrate by pepstatin. The inhibition appeared to be mostly competitive in nature. From Dixon plots of 1/velocity versus inhibitor concentration at different substrate concentrations, a K_i of 1.4 μM was obtained. Inhibition by pepstatin is normally viewed as a characteristic of the aspartic proteases.

In conclusion, it has proven possible to prepare a 99-residue sequence that can be chemically characterized as having the intended sequence and composition. Even though the purity of the synthetic product has not been rigorously established, this material can be folded to a form that catalyzes the proteolysis of synthetic and natural peptide substrates in a specific manner that relates it to the protease from HIV. The availability of synthetic product has allowed characterization of structural properties of the enzyme and aspects of substrate specificity. The diversity of cleavages observed preceding both proline, a secondary amino acid, and several primary amino acids could easily have been attributed to a mixture of diverse enzymes had the source been *in vivo* synthesis. The revelation of an unexpected importance of this enzyme to viral protein processing strengthens the potential importance of inhibitors as therapeutic agents. Toward this goal, the successful chemical synthesis of the protease has served as a practical route to the establishment of screening for the discovery of specific and more potent inhibitors.

We would like to acknowledge Drs. G. Vlasuk and M. Polokoff for supplies of purified p55, Ms. Jill C. Heimbach for conducting experiments on the yeast-expressed gag p55 substrate, Drs. V. Garsky and R. Freidinger for supplies of the synthetic octapeptide substrate, Mrs. L. Wassel for amino acid analyses, and Mrs. V. Finley for typing of the manuscript.

1. Coffin, J. (1982) in *RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 261–368.
2. Toh, H., Ono, M., Saigo, K. & Miyata, T. (1985) *Nature (London)* **315**, 691.
3. Pearl, L. H. & Taylor, W. R. (1987) *Nature (London)* **329**, 351–354.
4. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, G., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313**, 277–284.
5. Crawford, S. & Goff, S. P. (1985) *J. Virol.* **53**, 899–907.
6. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A. F., Scolnick, E. M. & Sigal, I. S. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4686–4690.
7. Kramer, R. A., Schaber, M. D., Skalka, A. M., Ganguly, K., Wong-Staal, F. & Reddy, E. P. (1986) *Science* **231**, 1580–1584.
8. Katoh, I., Yasunaga, T., Ikawa, Y. & Yoshinaka, Y. (1987) *Nature (London)* **329**, 654–656.
9. Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W. & Rosenberg, M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8903–8906.
10. Mous, J., Heimer, E. P. & LeGrice, S. F. J. (1988) *J. Virol.* **62**, 1433–1436.
11. Gutte, B. & Merrifield, R. B. (1969) *J. Am. Chem. Soc.* **91**, 501.
12. Hirschmann, R., Nutt, R. F., Veber, D. F., Vitali, R. A., Varga, S. L., Jacobs, T. A., Holly, F. W. & Denkewalter, R. G. (1969) *J. Am. Chem. Soc.* **91**, 506.
13. Yajima, H. & Fujii, N. (1981) *Biopolymers* **20**, 1859–1867.
14. Kent, S. & Clark-Lewis, I. (1985) in *Synthetic Peptides in Biology and Medicine*, eds. Alitalo, K., Partanen, P. & Vaheri, A. (Elsevier, the Netherlands), pp. 29–57.
15. Tam, J. P., Wong, T. W., Riemen, M. W., Tjoeng, F. S. & Merrifield, R. B. (1979) *Tetrahedron Lett.* 4033–4036.
16. Brown, T., Jones, J. H. & Richards, J. D. (1982) *J. Chem. Soc. Perkin Trans. 1*, 1553–1561.
17. Previero, A., Coletti-Previero, M. A. & Cavadore, J. (1967) *Biochim. Biophys. Acta* **147**, 453–461.
18. Barany, G. & Merrifield, R. B. (1980) in *The Peptides*, eds. Gross, E. & Meienhofer, J. (Academic, New York), Vol. 2, 1–284.
19. Sarin, V. K., Kent, S. B. H., Tam, J. P. & Merrifield, R. B. (1981) *Anal. Biochem.* **117**, 147–157.
20. Niall, H. D., Tregear, G. W. & Jacobs, J. (1972) in *Chemistry and Biology of Peptides*, ed. Meienhofer, J. (Ann Arbor Science Publishers, Ann Arbor, MI), pp. 695–699.
21. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
22. Greenstein, J. P. & Winitz, M. (1961) in *Chemistry of the Amino Acids* (Wiley, New York), Vol. 2, p. 1689.
23. Sabatier, J. M., Darbon, H., Fourquet, P., Rochat, H. & Van Rietschoten, J. (1987) *Int. J. Pept. Protein Res.* **30**, 125–134.
24. Matsudaira, P. (1987) *J. Biol. Chem.* **262**, 10035–10038.
25. Barany, G., Kneib-Cordonier, N. & Mullen, D. G. (1987) *Int. J. Pept. Protein Res.* **30**, 705–739.
26. Veber, D. F., Milkowski, J. D., Varga, S. L., Denkewalter, R. G. & Hirschmann, R. (1972) *J. Am. Chem. Soc.* **94**, 5456–5461.
27. Tam, J. P., Heath, W. F. & Merrifield, R. B. (1983) *J. Am. Chem. Soc.* **105**, 6442–6455.
28. Matsueda, G. R. (1982) *Int. J. Pept. Protein Res.* **20**, 26–34.
29. Fujii, N., Otaka, A., Funakoshi, S., Bessho, K. & Yajima, H. (1987) *J. Chem. Soc. Chem. Commun.*, 163–164.
30. Erickson, B. W. & Merrifield, R. B. (1973) *J. Am. Chem. Soc.* **95**, 3750–3756.
31. Savige, W. E. & Fontana, A. (1977) *Methods Enzymol.* **47**, 459–469.
32. Paul, A. V., Schultz, A., Pincus, S. E., Oroszlan, S. & Wimmer, E. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7827–7831.