JOURNAL OF VIROLOGY, Feb. 1991, p. 922–930
0022-538X/91/020922-09$02.00/0

Vol. 65, No. 2

Mutagenesis of Protease Cleavage Sites in the Human Immunodeficiency Virus Type 1 *gag* Polyprotein

RADONNA J. TRITCH,[1] YIH-SHYUN E. CHENG,[2] FAY H. YIN,[2] AND SUSAN ERICKSON-VIITANEN[2]*

Medical Products[1] and Central Research and Development[2] Departments, E. I. Du Pont de Nemours and Co., Inc., Wilmington, Delaware 19880-0400

Received 26 July 1990/Accepted 15 November 1990

The virally encoded protease of human immunodeficiency virus (HIV) is responsible for specific cleavage events leading to the liberation of the enzymes reverse transcriptase, integrase, ribonuclease H, and the core proteins from the *gag-pol* and *gag* polyprotein precursors. Utilizing *gag* polyprotein synthesized in vitro, we have shown that this substrate is sequentially cleaved by purified HIV protease to yield products that on the basis of their sizes and immunoreactivities correspond to p15, p6, p7, p17, and finally mature p24. We have placed unique restriction sites flanking the p17-p24 domain in order to facilitate replacement of cleavage site sequences by utilizing oligonucleotide cassettes. Replacement of the rapidly cleaved methionine-methionine bond at the p24-p15 junction with tyrosine-proline or replacement of the tyrosine-proline bond at the p17-p24 junction with methionine-methionine results in sites that cannot be efficiently cleaved. A basic amino acid at the p17-p24 scissile bond is not tolerated. Replacement of this cleavage site with an inverted repeat amino acid sequence gives intermediate rates of cleavage. In an attempt to convert the p17-p24 domain into a p24-p15 domain, residues flanking the scissile bond were exchanged in an expanding iterative fashion. When four residues flanking the scissile bond had been replaced, the rate of cleavage relative to that of the native p17-p24 sequence was increased fourfold. The cleavage rate of the native p24-p15 sequence is still some 10-fold greater than that of the p17-p24 sequence, suggesting that more-distant residues significantly affect the cleavage rate.

The protease of the human immunodeficiency virus type 1 (HIV-1) is responsible for specific cleavage events leading to the liberation of the enzymes reverse transcriptase, integrase, and ribonuclease H, and the core proteins from the *gag-pol* and *gag* polyprotein precursors (30). It has been shown through genetic means that this processing by the viral protease is essential for the production of infectious virions (13, 29). The HIV-1 protease has been the object of intensive investigation in the last few years; we and others have expressed active recombinant protease in *Escherichia coli* and other expression systems (2, 3, 5, 10, 11, 20). Biochemical and crystallographic evidence has firmly established that the protease is active as a dimer with structural and functional analogies to the aspartyl acid family of proteases (12, 18, 19, 27, 28, 33, 34). Cocrystals of protease dimer with a synthetic, noncleavable peptide analog suggest multiple contact points between substrate and enzyme in a binding site that includes at least seven residues of the substrate molecule (25). Because of the essentiality of the HIV-1 protease for infectious virion production and the relative wealth of information regarding its mechanism of action, the HIV-1 protease has emerged as a prime target for the development of inhibitors that might prove beneficial in the treatment of AIDS. Recent reports indicate that such inhibitors are antiviral through inhibition of viral polyprotein processing (7, 23, 30).

The rational design of inhibitors of the HIV protease requires understanding not only of the enzyme but also of the substrate specificity. There are four cleavage sites within the *gag* polyprotein precursor, which on cleavage gives rise to the mature core proteins p17, p24, p7, and p6. Sequence comparison does not indicate strict homology between these sites, in terms of either the scissile bond (Met-Met, Tyr-Pro, Leu-Ala, and Phe-Leu) or the immediately surrounding residues. Our goal is to determine those factors responsible for the substrate specificity of the HIV protease. Substrate specificity of the HIV-1 protease has been previously addressed through the use of small peptide mimics of some of the viral cleavage sites (1, 4, 15, 26, 31). As an extension of these studies with small peptides, it would be of interest to determine the role of the context in which the scissile peptide bond is presented in determining how the large polyprotein precursors are processed.

We have previously utilized in vitro translation products of full-length *gag* polyprotein precursor to study the relative processing at the various domains in this polyprotein (8). Our data suggested sequential processing at the amino- and carboxyl-terminal domains of the precursor that was largely independent of the presence of the other domain and which differs in rate by approximately 50-fold. We report here our attempt to alter the relative rates of processing in vitro of the p17-p24 and p24-p15 domains via site-directed mutagenesis. Replacement of proximal residues is poorly tolerated in terms of cleavage. It appears that the cleavage pair must be viewed in the context of the surrounding residues that occur naturally. Our results also suggest that determinants outside the heptapeptide-binding site may play a significant role in determining the overall rate of processing at a particular site.

MATERIALS AND METHODS

Plasmid constructions. DNA manipulations were carried out by using standard procedures (21). Plasmid pDAB72MA encoding full-length *gag* polyprotein in which an internal initiating methionine at position 142 has been mutated to an alanine was previously described (8) and was the starting material for all sequence manipulations. The following mutations in the amino acid sequence were introduced via site-directed mutagenesis by using the method of Kunkel et

* Corresponding author.

al. (16): MMYP, Met-378 to Tyr and Met-379 to Pro; YPKP, Tyr-132 to Lys; and YPFP, Tyr-132 to Phe. In order to facilitate replacement of additional amino acid residues at the p17-p24 junction (variable site 1 [vs1]), unique restriction sites *Xho*I and *Mlu*I flanking the p17-p24 junction were introduced via site-directed mutagenesis. The *Xho*I site was introduced 24 bp upstream of the junction, and the *Mlu*I site was introduced 64 bp downstream of the junction; neither mutation changes the amino acid sequence of the resultant protein. Oligonucleotide cassettes encoding the desired sequence were created from 40-mer to 45-mer oligonucleotide pairs with appropriate overhangs to form an 88-bp cassette with *Xho*I- and *Mlu*I-compatible ends. Plasmids were then constructed by ligation of *Xho*I-*Dra*III and *Mlu*I-*Dra*III fragments together with the oligonucleotide cassette (Fig. 1). Mutations were confirmed by sequencing.

Expression of pr55^{gag} and mutants in vitro. Plasmids were prepared from transformed XL-1 cells by using an alkaline lysis procedure (21). Linear DNA template was prepared for transcription by *Bam*HI digestion. Alternatively, RNA encoding p17 and the first 78 amino acids of p24 was prepared from *Pst*I-digested DNA template. In vitro transcriptions were carried out as previously described (8). In vitro translation was carried out by using a rabbit reticulocyte translation system (Promega Biotec) according to the manufacturer's instructions. Reaction mixtures included either [^{35}S] methionine or [^{35}S]cysteine at 1 μCi/μl (1,100 Ci/mmol). A portion of the total translation mixtures was analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and fluorography.

Protease assays. Translation mixtures corresponding to wild-type or mutant plasmids were adjusted to yield equal concentrations of translation product in the final reaction volume on the basis of the number of methionine residues present and counts per minute incorporated into *gag* polyproteins. Recombinant protease was isolated from inclusion bodies of the plasmid pET3AM as previously described (2) and used at a final concentration of 0.5 to 5 μg/ml. Reactions were carried out in phosphate-buffered saline, pH 7.2, at 30°C. Individual aliquots of proteolytic reactions were quenched with 2× Laemmli SDS-PAGE sample buffer (17), boiled, and frozen until analysis. Peptides were added as indicated to reactions in dimethyl sulfoxide. Cleavage of synthetic peptides was carried out in 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES)–1.0 mM dithiothreitol–0.1 mM EDTA–1.0 M NaCl–10% glycerol (pH 5.5). After the indicated time, samples were quenched with an equal volume of 0.1% trifluoroacetic acid and frozen until analysis. Samples were chromatographed on a Vydak C18 reverse-phase column with acetonitrile as the mobile phase.

Immunoprecipitation of protease-digested proteins. Immunoprecipitations were carried out in a buffer containing 50 mM Tris (pH 7.4), 190 mM NaCl, 6 mM EDTA, 1.0% Triton X-100, and 0.1% SDS. Monoclonal antibodies to p17 and p24 were reconstituted according to the manufacturer's instructions and used at an additional 25-fold dilution. Antibody to *gag* p6 was obtained as a tissue culture supernatant which was concentrated 10-fold by ultrafiltration. To 10 μl of protease reactions, 240 μl of buffer and 10 μl of antibody were added. After 3 h of being shaken at room temperature, 250 μl of Tachysorb-immobilized second antibody (Calbiochem) was added. After an additional hour of incubation, the immobilized complex was collected by centrifugation and washed with additional immunoprecipitation buffer. The pellet was finally resuspended in 2× sample buffer and heated to 100°C prior to electrophoresis.

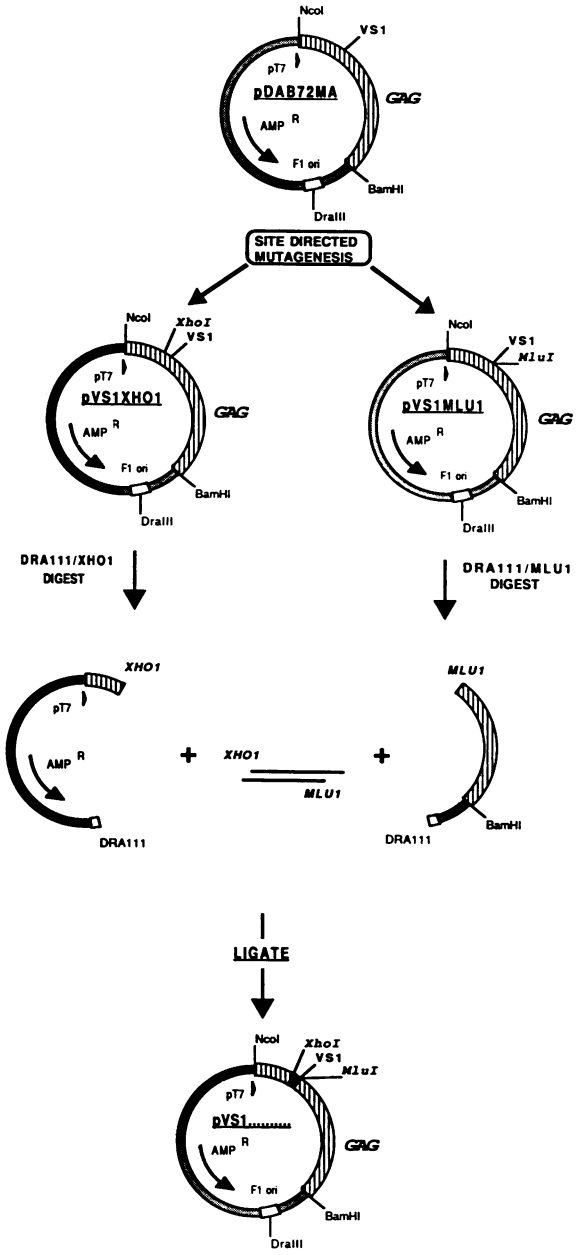

FIG. 1. Creation of altered cleavage sites in HIV-1 *gag* polyprotein. Unique restriction sites (*Xho*I and *Mlu*I) were placed around the p17-p24 domain (designated variable site 1 or VS1) via site-directed mutagenesis. New cleavage sites were then created by ligation of *Xho*I-*Dra*III and *Mlu*I-*Dra*III fragments together with oligonucleotide cassettes (88 bp) encoding the desired sequence. Mutations were verified by sequencing.

Miscellaneous. Samples from protease digestion of translation products were analyzed on 9 to 18% gradient SDS-polyacrylamide gels (0.75 mm) by using the buffer system of Laemmli (17). Gels were fixed and treated with En^{3}Hance for fluorography. Quantitative analysis of protease digestion was obtained with an Ambis Radioanalytic Imaging System (Ambis Systems, Inc.). Reported molecular masses were calculated by using ^{14}C-labeled molecular mass markers. Relative rates were established from comparison of half-life

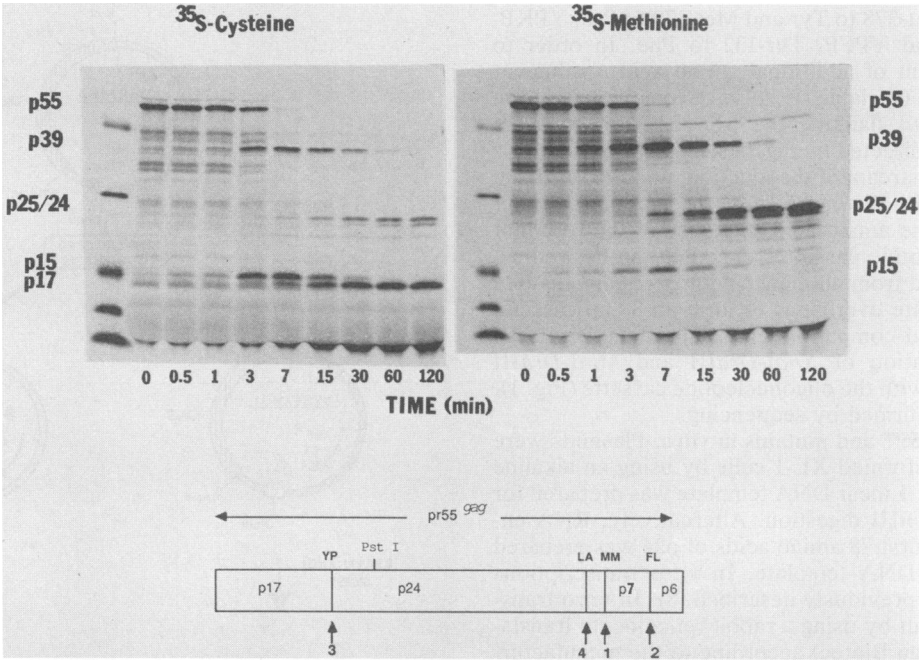

FIG. 2. Sequential cleavage of *gag* polyprotein precursors by the HIV-1 protease. [^{35}S]methionine- or [^{35}S]cysteine-labeled in vitro translation products of mRNA encoding the full-length pr55^{gag} precursor were incubated with 4 μg of HIV-1 protease per ml, and timed aliquots were analyzed by SDS-PAGE and fluorography. Products which correspond in size to full-length p55^{gag}, p39, p25-p24, p15, and p17 are indicated. Shown below is a schematic representation of the *gag* precursor and the sequential processing observed in vitro. The position of the *Pst*I site that gives rise to truncated *gag* precursor is also indicated.

values obtained from linear (first-order) semilog plots of substrate remaining versus time. Peptides were synthesized by using 9-fluorenylmethyloxycarbonyl (FMOC) chemistry via the RaMPS (35) technique. Oligonucleotides were made by using phosphoramidite chemistry on an ABI synthesizer.

RESULTS

Identification of processing intermediates in the cleavage of full-length and truncated pr55^{gag}. The in vitro translation products of mRNA encoding the full-length *gag* polyprotein precursor or a truncated product corresponding to p17 plus the first 78 amino acids of p24 were labeled with [^{35}S] methionine or [^{35}S]cysteine. These products were incubated with HIV-1 protease, and timed aliquots were analyzed by SDS-PAGE (Fig. 2 and 3). Note that in order to distinguish between processing intermediates arising from cleavage of full-length precursor and internally initiated polypeptide, methionine residue 142 has been replaced by alanine as previously described (8). Material migrating at 39 and 18 kDa appears rapidly (1 min [Fig. 2]), whereas a doublet migrating at 24 to 25 kDa appears more slowly. Only with extended incubation times or higher concentrations of protease does the doublet resolve to the 24-kDa species. Mature p17 contains no methionine residues; thus, the band migrating at 18 kDa after 1 or 3 min of digestion likely corresponds to p15 (Fig. 3). Labeling with cysteine reveals a distinct, faster-migrating species whose appearance is coincident with the appearance of p24-p25. The faster-migrating species is also present in protease digests of truncated *gag* precursor (Fig. 3). The identification of the faster and slower 18-kDa species as putative p17 and p15, respectively, was accomplished by immunoprecipitation of protease digests with monoclonal antibodies to p17, p24, and the p6 portion of p15 (Fig. 3). In methionine-labeled digests, anti-p6-reactive species appear in the 5- but not the 30-min samples, consistent with delayed processing of p15 to p7 plus p6 (6). Mature p6 migrates with the dye front in this gel system. There is no cross-reaction of the p39 intermediate with the anti-p6 monoclonal, suggesting that there is no significant cleavage at the p17-p24 junction of the full-length precursor. Monoclonal antibodies to p17 precipitate material running below the 18-kDa size marker from [^{35}S]cysteine-labeled truncated precursor (Fig. 3) or full-length precursor (data not shown). Together, the results presented in Fig. 2 and 3 show that the processing at the four cleavage sites in the *gag* precursor in this in vitro system is a sequential series of events: cleavage first at the amino terminus of p15, cleavage next at the p7-p6 junction closely followed by cleavage at the p17-p24 junction, and finally processing of p25 to p24 (Fig. 2). As reported previously, cleavage at the amino terminus of p15 occurs some 50 times faster than subsequent cleavages at the other cleavage sites (8). Although not proven, the molecular mass, immunoreactivity, and effects of single amino acid substitutions described below suggest that the cleavages are occurring in the in vitro translation products at or very near the sites cleaved by the protease in virally infected cells. The relevant amino acid sequences corresponding to these cleavage events are shown in Table 1. The junction giving rise to the amino terminus of p15 is termed the fast site, whereas the p17-p24 junction is termed a slow site.

Changing a fast cleavage site into a slow site. Site-directed mutagenesis was used to change the methionine-methionine cleavage site corresponding to the amino terminus of p15 to a tyrosine-proline junction (designated MMYP). Compared with a native full-length precursor, the MMYP polyprotein is cleaved more slowly, and new intermediates with apparent molecular masses of 43, 38 to 40, and 33 kDa are observed

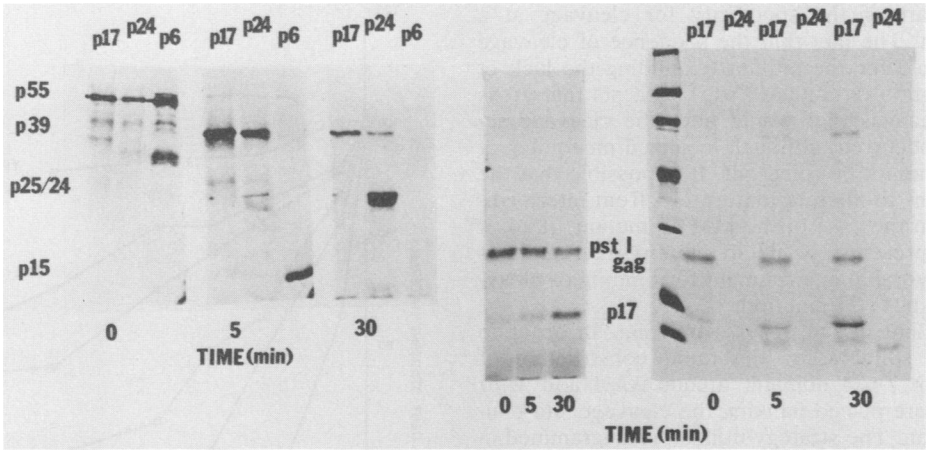

FIG. 3. Immunoprecipitation of protease-digested [^{35}S]methionine- and [^{35}S]cysteine-labeled pr55^{gag} polyproteins. (Left panels) Aliquots corresponding to 5 and 30 min of digestion of [^{35}S]methionine-labeled full-length *gag* polyprotein by HIV-1 protease and samples to which no protease was added (0 min) were immunoprecipitated with antibodies to p17, p24, or p6 and analyzed by SDS-PAGE and fluorography. (Middle panel) Truncated *gag* precursor labeled with [^{35}S]cysteine was digested with HIV-1 protease, and timed aliquots were analyzed by SDS-PAGE and fluorography. (Right panel) Aliquots corresponding to 0, 5, or 30 min of digestion of truncated *gag* precursor with HIV-1 protease were immunoprecipitated with antibodies to p17 or p24 and analyzed by SDS-PAGE and fluorography.

(Fig. 4A). p15 is not observed in the MMYP digestion. Samples corresponding to 2, 20, 40, or 0 min of incubation with HIV-1 protease were immunoprecipitated with monoclonal antibodies to p17, p24, or p6, and the products were analyzed on gels (Fig. 4B). As expected, all three antibodies immunoprecipitate the full-length polyprotein. The material migrating at 43 kDa is immunoprecipitated by antibodies to p17 and p24, whereas the material migrating at 38 kDa is precipitated by antibodies to p24 and p6. The band identified as p33 is immunoprecipitated only by antibodies to p24. These reactivities and molecular masses are consistent with the following assignments: p43 = p17 plus p24 plus p7, p38 = p24 plus p7 plus p6, and p33 = p24 plus p7. p33 can thus arise from intermediate p43 or p38. The band intensities suggest that the digestion pathway involving p43 is preferred. Although the disappearance of p55 is significantly slower in the MMYP digest, the final appearance of p24-p25 is quite similar (Fig. 4A), suggesting that the overall rate-limiting step for p24-p25 production is unchanged.

Close inspection of the autoradiogram of the material migrating at the 24- to 25-kDa region for the MMYP mutant suggests mature p24 has been produced from the intermediates. These results are consistent with the following overall sequence of cleavage events for the MMYP polyprotein: cleavage at the p7-p6 junction, followed closely by cleavage at the p17-p24 junction, followed by cleavage at the carboxy terminus of p24 (Fig. 5). There is no significant cleavage at the MM-to-YP mutated junction. This suggests that the amino acid residues making up the scissile bond per se

TABLE 1. HIV-1 protease cleavage sites within the HIV-1 *gag* polyprotein

Sequence[a]	Location in *gag* polyprotein
Q-V-S-Q-N-**Y-P**-I-V-Q-N-I	p17-p24 junction
H-K-A-R-V-**L-A**-E-A-M-S	p24 COOH terminus
N-S-A-T-I-**M-M**-Q-R-G-N-F	p7 NH_2 terminus
G-R-P-G-N-**F-L**-Q-S-R-P-E	p7-p6 junction

[a] Sequences are given in one-letter code and are taken from reference 4. Boldfaced residues indicate the pair of residues at which cleavage occurs.

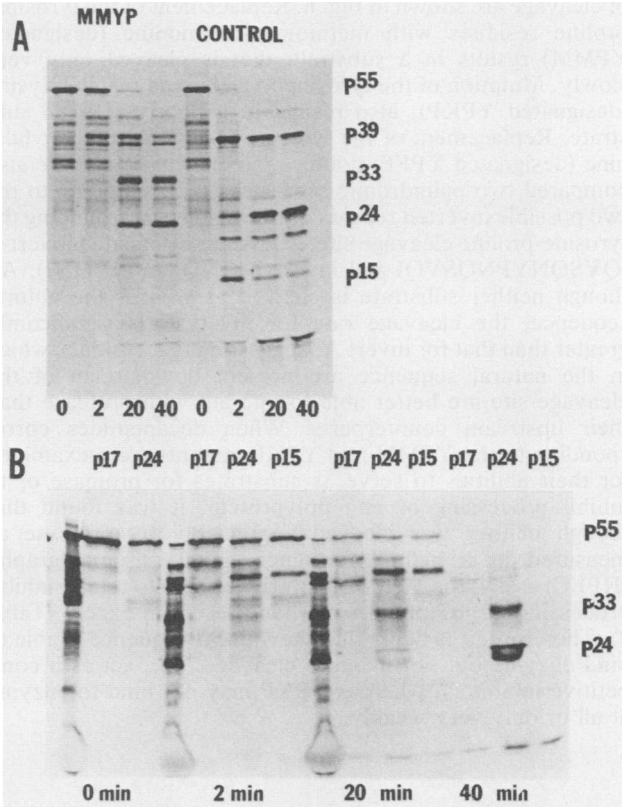

FIG. 4. Replacement of methionine-methionine bond at the p25-p7 junction with tyrosine-proline creates a noncleavable site. (A) Full-length *gag* polyprotein corresponding to wild-type (control) or mutant (MMYP) sequence was digested with HIV-1 protease, and timed (in minutes, shown below lanes) aliquots were analyzed by SDS-PAGE and fluorography. The various digestion products are indicated to the right of the figure. The lower band of the doublet at 38 to 40 kDa is present in all samples and is likely to be a degradation product. (B) Immunoprecipitation of digestion products of MMYP by HIV-1 protease. Samples corresponding to 0, 2, 20, and 40 min were immunoprecipitated with antibodies to p17, p24, or p6 and analyzed by SDS-PAGE and fluorography.

impart a large part of the specificity for cleavage at a particular junction. The fact that the sequence of cleavage events is unchanged for the mutant (excluding the lack of cleavage at the amino terminus of p15) suggests that gross conformational changes that would alter the cleavage sequence have not occurred, although localized misfolding at the mutated site cannot be ruled out. It is possible that the final cleavage event to liberate mature p24 from intermediates is enhanced in the case of the MMYP mutant. If so, it implies that the presence of p7 in mutant intermediates imparts a more favorable environment for binding or cleavage or both at the p24-p25 junction.

Effects of mutations at the p17-p24 junction. In order to provide a vector suitable for the rapid construction of mutations at the p17-p24 domain, unique *Xho*I and *Mlu*I restriction sites were placed flanking the cleavage site residues at this junction. The strategy utilized is diagrammed in Fig. 1. We previously showed that the rate of cleavage at the p17-p24 junction from a truncated *gag* precursor corresponding to all of p17 and the first 78 amino acids of p24 was similar to the cleavage at this site in the full-length polyprotein (8). Therefore, we utilized the truncated precursor in the comparison of mutants. The time course of cleavage of mutant in vitro translation products and their relative rates of cleavage are shown in Fig. 6. Replacement of the tyrosine-proline residues with methionine-methionine (designated YPMM) results in a substrate that is cleaved only very slowly. Mutation of the tyrosine to a charged residue, lysine (designated YPKP), also results in a poorly cleaved substrate. Replacement of the tyrosine residue with phenylalanine (designated YPFP) doubles the cleavage rate. We also compared two palindromic sequences corresponding to the two possible inverted repeats of the residues surrounding the tyrosine-proline cleavage site. These are designated invert A (QVSQNYPNQSVQ) and invert B (INQVIYPIVQNI). Although neither substrate is cleaved as well as the natural sequence, the cleavage rate for invert B is significantly greater than that for invert A, suggesting that residues which in the natural sequence are present downstream of the cleavage site are better able to interact with protease than their upstream counterparts. When decapeptides corresponding to the YPMM and YPKP mutants were examined for their abilities to serve as substrates for protease or to inhibit processing of *gag* polyprotein, it was found that though neither was cleaved appreciably by protease as measured in a high-performance liquid chromatography (HPLC) assay, the YPMM peptide was able to totally inhibit processing of *gag* precursor when present in excess (Table 2). Therefore, it is likely that the YPMM sequence is able to bind enzyme but not undergo cleavage (i.e., act as a competitive inhibitor), whereas YPKP may not bind to enzyme at all or only very weakly.

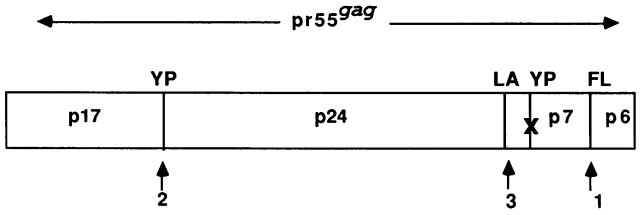

FIG. 5. Schematic representation of cleavage pattern of MMYP mutant *gag* precursor. The order of cleavage events and products indicated by radioimmunoprecipitation are indicated.

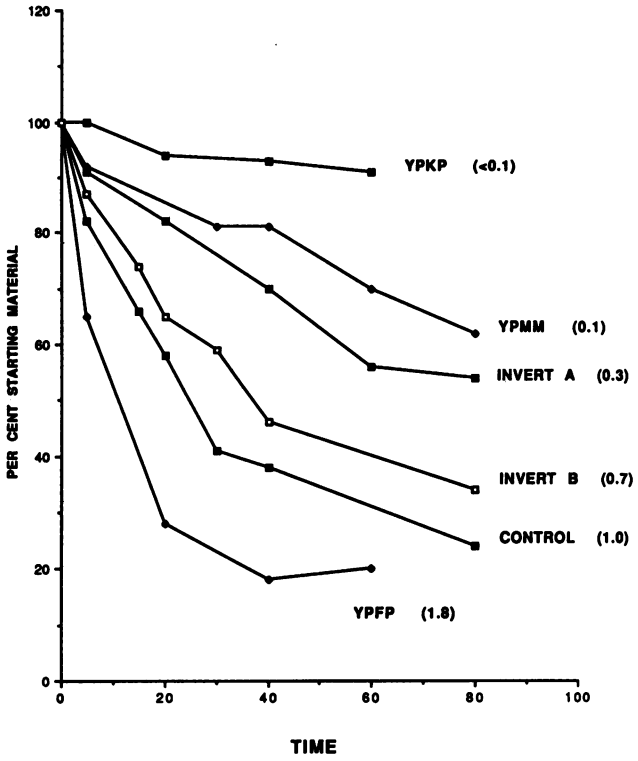

FIG. 6. Protease digestion of *gag* polyprotein mutants. Truncated *gag* precursor containing the indicated sequence at the tyrosine-proline junction was digested with HIV-1 protease, and timed aliquots were analyzed by SDS-PAGE and fluorography. Bands corresponding to substrate were quantified by an Ambis scanner. The relative rates of cleavage (control tyrosine-proline sequence was assigned a value of 1.0) were derived from comparison of half-life values from log substrate versus time plots. Data shown are the averaged curves for two to four experiments for each mutant sequence.

Building a fast cleavage site at the p17-p24 junction. Since it was clear that the pair of methionine residues per se was not sufficient to yield a rapid cleavage site when in the context of the p17-p24 junction, we attempted to achieve the rapid rate of cleavage characteristic of the methionine site by successive replacement of flanking residues with the cognate sequence. Figure 7 illustrates the rate enhancement obtained with the successive addition of proper context residues compared with the control p17-p24 sequence (measured as the cleavage of truncated precursor) and the control p25-p7 sequence (measured as the cleavage of pr55^{gag} precursor). The relative rates of cleavage of these mutants are shown in Table 2 along with the relative rates of cleavage of peptide mimics of the mutations and the abilities of the peptide mimics to inhibit cleavage of the *gag* polyprotein precursor. Even after mutations that create the cognate context for 9 of 10 amino acids surrounding the methionine-methionine junction, the degree of rate enhancement observed is still an order of magnitude short of that observed for the natural methionine-methionine domain. Peptide mimics of the various mutations show differences in rate that are less than that observed for the methionine-methionine site within the polyprotein environment. The inhibition of *gag* polyprotein cleavage observed with mutant sequences that are cleaved is transient, presumably reflecting disappearance of peptide

TABLE 2. Rate of cleavage of p17-p24 mutant peptides and polyproteins

Sequence	Relative cleavage		Inhibition of *gag* polyprotein[c]
	gag polyprotein[a]	Peptide mimic[b]	
VSQNYPIVQN	1.0	1.0	Slight
VSQNMMIVQN	0.1	0.0	100
VSQIMMQVQN	4.0	ND	Transient
VSTIMMQRQN	1.3	1.2	Transient
VATIMMQRGN	2.6	2.0	Transient
SATIMMQRGN	48.0	2.0	Transient
VSQNKPIVQN	0.1	0.1	0

[a] Relative rates were derived from comparison of half-life values from semilog plots of substrate versus time from data such as is shown in Fig. 7. Rates of cleavage were determined by using the truncated *gag* precursor, except for SATIMMQRGN, which represents cleavage of the control p25-p7 sequence within the full-length *gag* precursor.

[b] Reactions contained 0.5 mM peptide and 10 μg of HIV-1 protease per ml in a buffer containing 50 mM MES (pH 5.5), 1 M NaCl, 10% glycerol, 0.1 mM EDTA, and 1 mM dithiothreitol. After 10 or 30 min, reactions were quenched with an equal volume of 0.1% trifluoroacetic acid and chromatographed on a Vydak C18 reverse-phase column. Rates were determined by measuring remaining substrate. ND, Not determined.

[c] Truncated *gag* polyprotein precursor was incubated with HIV-1 protease in the presence of 100 μM of the indicated peptide. After 5, 20, or 40 min or before addition of the *gag* precursor, aliquots were removed for analysis by SDS-PAGE and fluorography. The rate of disappearance of *gag* precursor was compared with that in a reaction in which only peptide vehicle (dimethyl sulfoxide) was included. Transient, Inhibition of substrate cleavage observed after 5 or 20 min of incubation, but cleavage observed between 20 and 40 min of incubation. Values are reported as percentages.

during the experiment. Only the poorly cleaved YPMM mutant peptide is able to inhibit *gag* processing essentially completely throughout the experiment.

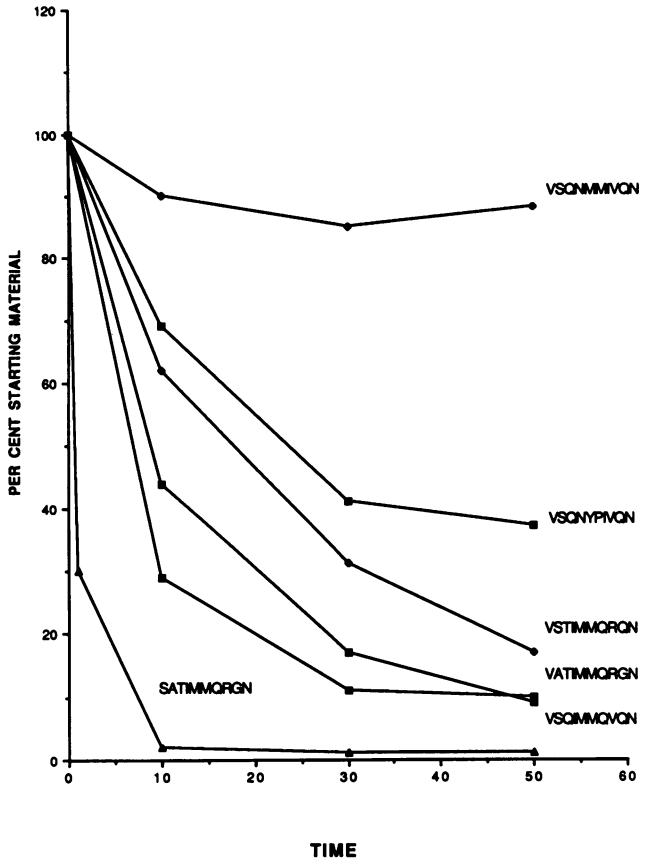

FIG. 7. Effects of context sequence on rate of cleavage of p17-p24 junction with YPMM mutation. Truncated precursor polyproteins corresponding to the native *gag* polyprotein or p17-p24 mutants were digested with HIV-1 protease, and timed aliquots were analyzed by SDS-PAGE and fluorography. Bands corresponding to substrate were quantitated by an Ambis scanner. The curve for the control p25-p7 sequence was derived from quantitating the disappearance of $pr55^{gag}$ in reactions containing the full-length *gag* precursor.

DISCUSSION

The rational design of potent, specific inhibitors of the HIV-1 protease requires understanding not only of the three dimensional structure of the protease but also of the features involved in its complex substrate specificity. From studies with peptide mimics of various cleavage sites as well as crystal structure data, several features of the cleavage site specificity have emerged. (i) Studies with peptide mimics of the seven recognized cleavage sites within the *gag* and *gag-pol* precursor polyproteins have indicated a minimum recognition length of seven residues (4). Seven residues also defines the minimal substrate for the myeloblastosis-associated helper virus protease (32). (ii) A heptapeptide-binding site can be accommodated within the protease dimer as indicated by X-ray crystallography and modeling studies (18, 25, 33, 34). (iii) Comparison of viral cleavage site sequences suggests that hydrophobic residues are required at the scissile bond and immediately adjacent residues, but amino acids at the P3 and P3′ positions and beyond can be acidic, basic, or hydrophobic (for an example, see Table 1). (iv) These requirements are mirrored in the amino acids located at the corresponding positions in the subsites within the protease dimer (18, 25, 33, 34). (v) Several reports attempting to rank the various cleavage sites for their relative efficiencies of cleavage by the HIV-1 protease have appeared (1, 4, 14, 15, 26, 31). With the caveats about the use of different reaction conditions and peptide lengths, it is clear that peptide sequences corresponding to the amino terminus of protease within the *pol* gene, the p17-p24 junction of *gag*, and the amino terminus of p7 of *gag* are cleaved more efficiently than sequences corresponding to the carboxy terminus of p24 in the *gag* polyprotein or the carboxy terminus of protease within the *pol* gene. When K_m and V_{max} values are considered, there is no more than a factor of 5 in efficiency between peptide mimics corresponding to the p17-p24 junction (a tyrosine-proline bond) and the p25-p7 junction (a methionine-methionine bond) of the *gag* polyprotein (4, 15). In agreement with these findings, when noncleavable peptide mimics of these two cleavage sites are compared for their ability to inhibit processing of *gag* polyprotein precursor, apparent K_i values of about 1 μM are observed for both analogs (31a). (vi) Substitutions at the P2′ site of *gag* p17-p24 of hydrophobic residues other than the naturally occurring isoleucine are poorly tolerated (22).

Although much can be learned through the use of small peptide mimics, the context in which the scissile peptide bond is presented may have important consequences in determining how the large viral precursors are processed. By utilizing full-length and truncated versions of the viral *gag* precursor in an in vitro system, we have attempted to explore, via selective mutagenesis, the relative importance

of the scissile peptide bond and the surrounding context of amino acids.

Previously (8), we compared the rate of appearance of intermediates and p24-p25 from full-length and truncated *gag* polyprotein precursors and suggested that the processing of the precursor in vitro was sequential. However, the lack of methionine in mature p17 precluded an unequivocal assignment of p17 and p15 products. We show here that the full-length viral *gag* precursor synthesized in vitro is cleaved in a sequential fashion as demonstrated through differential labeling of in vitro translation products and immunoprecipitation with antibodies to the various domains of the precursor (Fig. 2 and 3). Cleavage at the methionine-methionine bond to liberate p15 occurs first, is followed much more slowly by cleavage at the p7-p6 junction and at the p17-p24 junction, and is then followed finally by cleavage of p25 to yield mature p24. This ordered cleavage is in contrast to the initial cleavage products observed by Gowda et al. (9) by utilizing cells infected with a recombinant vaccinia virus. In that study, the authors observed a small amount of intermediate of approximately 39 kDa that immunoprecipitated with antibodies to p24 and p15, which led the authors to conclude that cleavage of full-length precursor was random, with initial cleavage occurring both at the p17-p24 junction and at the p25-p7 junction. Alternatively, such intermediates could arise from de novo synthesis at a methionine residue located at position 142, 10 amino acids downstream of the p17-p24 junction. We have shown that this methionine residue can serve as an initiation signal in vitro, giving rise to an intermediate of 40 kDa immunoprecipitable by antibodies to p24 but not p17 (8). Such alternative initiation has also been suggested to explain intermediates observed in HIV-infected H9 cells on the basis of the comparison of tryptic digest patterns of full-length and p40 species of *gag* precursors (24).

It is of interest to compare the processing scenario observed by using in vitro translation products with that surmised from steady-state and pulse-labeling measurements of virally infected cells. Processing intermediates corresponding to p17 plus p24 and internally initiated p24 plus p15 have been observed (24). Slow conversion of p25 to p24 has been observed both in HIV-infected H9 cells (24) and in vaccinia virus recombinant-infected cells (9). p6 has been observed in HIV-infected cells only through radiolabeling; Western immunoblot of cell lysates revealed only p15, which has led to the suggestion that p15 processing to p7 and p6 occurs only after virion maturation. Thus, the sequential cleavage pattern observed with the in vitro transcription-translation system mimics that observed in virally infected cells with regard to the early and late steps in the sequence but may differ with regard to the processing of p15. This may reflect a compartmentation of the viral p15 within the infected cell such that it is inaccessible to protease action or some other property of a cellular environment. In the present study, the effects of mutations at the p17-p24 and p25-p7 junctions should be compared with the sequence of events observed in the wild-type precursor in vitro. In an ongoing study, we are attempting to assess the effects of some of these mutations in cells transiently transfected with wild-type and mutant *gag* provirus.

The importance of the pair of amino acid residues that make up the scissile bond per se is demonstrated by the lack of cleavage of mutant *gag* precursors in which the p17-p24 junction was exchanged with the p25-p7 junction (mutants MMYP and YPMM) (Fig. 4 and 6). In the case of MMYP, the remaining sequence of cleavage events appears to occur essentially normally, indicating that the mutation has not created gross misfolding that would render the various cleavage sites inaccessible to the protease (Fig. 4 and 5). These results indicate that the proper context or environment for the scissile bond is an absolute requirement for cleavage (and by inference, recognition) by the HIV-1 protease.

In agreement with the lack of charged residues present at any of the seven recognized cleavage sites within *gag* and *pol* genes, a precursor protein or peptide mimic with a lysine-proline junction cannot serve as a substrate. Phenylalanine, which is present together with proline at two of the recognized cleavage sites, is well tolerated when substituted for tyrosine at the p17-p24 junction. A symmetrical sequence as regards positions P5-P2 and P2′-P5′ when derived from the carboxy-terminal sequence of *gag* p17-p24 is cleaved efficiently (invert B, INQVIYPIVQNI) (Fig. 6). The corresponding amino-terminal symmetrical sequence (QVSQNY PNQSVQ) is a poor cleavage site. Strop et al. (32) compared symmetrical peptide mimics of the myeloblastosis-associated helper virus *pol* p63-p32 cleavage site and observed a relative cleavage rate for the amino-terminal palindromic sequence of 0.6 compared with that of the natural sequence, while the carboxy-terminal palindrome was not cleaved to any extent. Whether these differences reflect a preference for a particular amino acid (e.g., isoleucine versus asparagine) or more complicated reciprocal interactions between subsites as proposed by Strop et al. (32) remains to be established.

The importance of environment is most clearly demonstrated by the relatively slow cleavage rate observed for mutations in which residues at the P4 to P5′ site for the p25-P7 junction were placed into the context of the p17-p24 domain (Fig. 7; Table 2). Mutation of residues corresponding to the P2 and P2′ positions in addition to the YPMM mutation (VSQIMMQVQN) yielded a substrate with a four-fold rate enhancement relative to that of the control tyrosine-proline sequence. Subsequent addition of p25-p7 junction amino acids at positions corresponding to P3 and P3′ (VS TIMMQRQN) actually decreased the cleavage rate to a value near to that of the control sequence. This may reflect poor binding contacts resulting from three adjacent bulky residues (glutamine, arginine, and glutamine) in the P2′-P4′ positions. Subsequent addition of P4 and P4′ residues (VA TIMMQRGN) yields a precursor that is cleaved at an enhanced rate relative to that of the control sequence but that is still far less than that observed for the totally natural sequence present at the p25-p7 junction. Peptide mimics of the various mutations also do not demonstrate large differences in their cleavage rates under the experimental conditions utilized. Therefore, additional conformational features of the natural sequence and/or additional binding determinants outside the actual heptapeptide binding site of the protease must be responsible for the more efficient cleavage of the p25-p7 domain. If such binding determinants exist, they would represent additional targets for the identification of adjunct protease inhibitors that may act to prevent good association of viral substrate with enzyme through action at sites other than the enzyme's active site cleft.

ACKNOWLEDGMENTS

We thank Paul Viitanen for many helpful discussions regarding the mutagenesis approach and caveats about its use for the study of substrate specificity. We thank Denise McCabe, Pamela Cawood, and Violet Daugherty for excellent technical assistance and Bill

Senkowicz and Rose Wilk for peptide synthesis. Monoclonal antibody to p6 was a gift from Fulvia diMarzo Veronese, who is gratefully acknowledged.

REFERENCES

1. **Billich, S., M.-T. Knoop, J. Hansen, P. Strop, J. Sedlack, R. Mertz, and K. Moelling.** 1988. Synthetic peptides as substrates and inhibitors of human immune deficiency virus-1 protease. J. Biol. Chem. **263**:17905–17908.
2. **Cheng, Y. S. E., M. H. Mcgowan, C. A. Kettner, J. V. Schloss, S. Erickson-Viitanen, and F. H. Yin.** 1990. High level synthesis of recombinant HIV-1 protease and the recovery of active enzyme from inclusion bodies. Gene **87**:243–248.
3. **Copeland, P. D., and S. Oroszlan.** 1988. Genetic locus, primary structure and chemical synthesis of HIV protease. Gene Anal. Tech. **5**:109–115.
4. **Darke, P. L., R. F. Nutt, S. F. Brady, V. M. Garsky, T. M. Ciccarone, C. T. Ley, P. K. Lumma, R. M. Fredinger, D. F. Veber, and I. S. Sigal.** 1988. HIV-1 protease specificity of peptide cleavage is sufficient for processing of *gag* and *pol* polyproteins. Biochem. Biophys. Res. Commun. **156**:297–303.
5. **Debouck, D., J. G. Gorniak, J. E. Strickler, T. D. Meek, B. W. Metcalf, and M. Rosenberg.** 1987. Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific maturation of the *gag* precursor. Proc. Natl. Acad. Sci. USA **84**:8903–8906.
6. **diMarzo Veronese, F., R. Rahman, T. D. Copeland, S. Oroszlan, R. C. Gallo, and M. G. Sarnagadharan.** 1987. Immunological and chemical analysis of p6, the carboxyl-terminal fragment of HIV p15. AIDS Res. Hum. Retroviruses **3**:253–265.
7. **Dreyer, G. B., B. W. Metcalf, T. A. Tomaszek, Jr., T. J. Carr, A. C. Chandler III, L. Hyland, S. A. Fakhoury, V. W. Magaard, M. L. Moore, J. E. Strickler, C. Debouck, and T. D. Meek.** 1989. Inhibition of human immunodeficiency virus 1 protease in vitro: rational design of substrate analogue inhibitors. Proc. Natl. Acad. Sci. USA **86**:9752–9756.
8. **Erickson-Viitanen, S., J. Manfredi, P. Viitanen, D. E. Tribe, R. Tritch, C. A. Hutchison III, D. D. Loeb, and R. Swanstrom.** 1989. Cleavage of HIV-1 *gag* polyprotein synthesized in vitro: sequential cleavage by the viral protease. AIDS Res. Hum. Retroviruses **5**:577–591.
9. **Gowda, S. D., B. S. Stein, and E. G. Engleman.** 1989. Identification of protein intermediates in the processing of the p55 HIV-1 *gag* precursor in cells infected with recombinant vaccinia virus. J. Biol. Chem. **264**:8459–8462.
10. **Graves, M. C., J. J. Lim, E. P. Heimer, and R. A. Kramer.** 1988. An 11-kDa form of human immunodeficiency virus protease expressed in *Escherichia coli* is sufficient for enzymatic activity. Proc. Natl. Acad. Sci. USA **85**:2449–2453.
11. **Hostomsky, A., K. Appelt, and R. C. Ogden.** 1989. High level expression of self-processed HIV-1 protease in *Escherichia coli* using a synthetic gene. Biochem. Biophys. Res. Commun. **161**:1056–1063.
12. **Kay, J., and B. M. Dunn.** 1990. Viral proteinases: weakness in strength. Biochim. Biophys. Acta **1048**:1–18.
13. **Kohl, N. E., E. A. Emini, W. A. Schleif, L. J. Davis, J. C. Heimbach, R. A. F. Dixon, E. M. Scolnick, and I. S. Sigal.** 1988. Active human immunodeficiency virus protease is required for viral infectivity. Proc. Natl. Acad. Sci. USA **85**:4686–4690.
14. **Kotler, M., R. A. Katz, W. Danho, J. Leis, and A. M. Skalka.** 1988. Synthetic peptides as substrates and inhibitors of a retroviral protease. Proc. Natl. Acad. Sci. USA **85**:4185–4189.
15. **Krausslich, H.-G., R. H. Ingraham, M. T. Skoog, E. Wimmer, P. V. Pallai, and C. A. Carter.** 1989. Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. Proc. Natl. Acad. Sci. USA **86**:807–811.
16. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
17. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
18. **Lapatto, R., T. Blundell, A. Hemmings, J. Overington, A. Wilderspin, S. Wood, J. R. Merson, P. J. Whittle, D. E. Danley, K. F. Geoghegan, S. J. Hawrylik, S. E. Lee, K. G. Scheld, and P. M. Hobart.** 1989. X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. Nature (London) **342**:299–302.
19. **Loeb, D. D., C. A. Hutchison III, M. H. Edgell, W. G. Farmerie, and R. Swanstrom.** 1989. Mutational analysis of human immunodeficiency virus type 1 protease suggests functional homology with aspartic proteinases. J. Virol. **63**:111–121.
20. **Louis, J. M., E. M. Wondrak, T. D. Copeland, C. A. D. Smith, P. T. Mora, and S. Oroszlan.** 1989. Chemical synthesis and expression of the HIV-1 protease gene in E. coli. Biochem. Biophys. Res. Commun. **159**:87–94.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
22. **Margolin, N., W. Heath, E. Osborne, M. Lei, and C. Vlahos.** 1990. Substitutions at the p2′ site of gag p17-p24 junction affect cleavage efficiency by HIV-1 protease. Biochem. Biophys. Res. Commun. **167**:554–560.
23. **Meek, T. D., D. M. Lambert, G. B. Dreyer, T. J. Carr, T. A. Tomaszek, Jr., M. L. Moore, J. E. Strickler, C. Dobouck, L. J. Hyland, T. J. Matthews, B. W. Metcalf, and S. R. Petteway.** 1990. Inhibition of HIV-1 protease in infected T-lymphocytes by synthetic peptide analogues. Nature (London) **343**:90–92.
24. **Mervis, R. J., N. Ahmad, E. P. Lillehoj, M. G. Raum, F. H. R. Salazar, H. W. Chan, and S. Venkatesan.** 1988. The *gag* gene products of human immunodeficiency virus type 1: alignment within the *gag* open reading frame, identification of posttranslational modifications, and evidence for alternative *gag* precursors. J. Virol. **62**:3993–4002.
25. **Miller, M., J. Schneider, B. K. Sathyanarayana, M. V. Toth, G. R. Marshall, L. Clawson, L. Selk, S. B. H. Kent, and A. Wlodawer.** 1989. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 A resolution. Science **246**:1149–1152.
26. **Moore, M. L., W. M. Bryan, S. A. Fakhoury, V. W. Magaard, W. F. Huffman, B. D. Dayton, T. D. Meek, L. Hyland, G. B. Dreyer, B. W. Metcalf, J. E. Strickler, J. G. Gorniak, and C. Debouck.** 1989. Peptide substrates and inhibitors of the HIV-1 protease. Biochem. Biophys. Res. Commun. **159**:420–426.
27. **Navia, M. A., P. M. D. Fitzgerald, B. M. McKeever, C.-T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer.** 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature (London) **337**:615–620.
28. **Pearl, L. H., and W. R. Taylor.** 1987. A structural model for the retroviral proteases. Nature (London) **329**:351–354.
29. **Peng, C., B. K. Ho, T. W. Chang, and N. T. Chang.** 1989. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. J. Virol. **63**:2550–2556.
30. **Roberts, N. A., J. A. Martin, D. Kinchington, A. V. Broadhurst, J. C. Craig, I. B. Duncan, S. A. Galpin, B. K. Handa, J. Kay, A. Krohn, R. W. Lambert, J. H. Merritt, J. S. Mills, K. E. B. Parkes, S. Redshaw, A. J. Ritchie, D. L. Taylor, G. J. Thomas, and P. J. Machin.** 1990. Rational design of peptide based HIV proteinase inhibitors. Science **248**:358–361.
31. **Schneider, J., and S. B. H. Kent.** 1988. Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. Cell **54**:363–368.

31a. **Shevni, A. B., and S. Erickson-Viitanen.** Unpublished data.

32. **Strop, P., J. Konvalinka, J., L. Pavlickova, I. Blaha, M. Soucek, J. Urban, J. Velek, D Stys, V. Kostka, and J. Sedlacek.** 1989. Specificity and inhibition studies on MAV proteinase, p. 259–267. *In* H.-G. Krausslich, S. Oroszlan, and E. Wimmer (ed.), Viral proteinases as targets for chemotherapy. Cold Spring

Harbor Laboratory, Cold Spring Harbor, New York.
33. **Weber, I. T., M. Miller, M. Jaskolski, J. Leis, A. M. Skalka, and A. Wlodawer.** 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. Science **243**:928–931.
34. **Wlodawer, A., M. Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. H. Kent.** 1989. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science **245**:616–621.
35. **Wolfe, H. R., and R. R. Wilk.** 1989. The RaMPS system: simplified peptide synthesis for life science researchers. Pept. Res. **2**:352–356.