Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 10023–10027, December 1990
Genetics

β-Galactosidase containing a human immunodeficiency virus protease cleavage site is cleaved and inactivated by human immunodeficiency virus protease

ELLEN Z. BAUM*, GERALDINE A. BEBERNITZ, AND YAKOV GLUZMAN

Molecular Biology Section, Lederle Laboratories, Medical Research Division, American Cyanamid Company, Pearl River, NY 10965

Communicated by Bernard Roizman, October 1, 1990 (received for review August 12, 1990)

ABSTRACT **A "cleavage cassette" specifying a decapeptide human immunodeficiency virus (HIV) protease cleavage site was introduced into six different locations of β-galactosidase (β-D-galactoside galactohydrolase, EC 3.2.1.23) in *Escherichia coli*. Four of these constructs retained β-galactosidase activity despite the insertion of the cleavage cassette. Of these four constructs, one was cleaved by HIV protease, resulting in the inactivation of β-galactosidase both *in vivo* and *in vitro*. This cleavage was inhibited by pepstatin A, a known inhibitor of HIV protease. Thus, β-galactosidase has been converted into an easily assayed substrate for HIV protease. An analogous construct of β-galactosidase containing a polio protease cleavage site was cleaved likewise by polio protease, suggesting that this system may be generic for monitoring cleavage by a variety of proteases.**

The β-galactosidase enzyme (β-D-galactoside galactohydrolase, EC 3.2.1.23) encoded by the *lacZ* gene of *Escherichia coli* is responsible for the conversion of lactose into glucose and galactose, allowing utilization of lactose as a carbon source (1). Because β-galactosidase will also catalyze the hydrolysis of several other substrates that are converted into colored compounds, β-galactosidase has proved to be an invaluable tool to monitor a wide range of biological processes. For example, β-galactosidase fusions have been used to study transcriptional and translational regulation (ref. 2 and references therein). Fusion of heterologous protein to β-galactosidase can facilitate the purification of that protein (3–6). In addition, interruption of the amino terminus of β-galactosidase is the basis of identifying desired clones when using the pUC plasmid series (7).

We have been studying the protease encoded by human immunodeficiency virus (HIV), the causative agent of AIDS (8). HIV protease is responsible for generating mature viral capsid proteins and enzymes from polyprotein precursors translated from gag and gag/pol mRNA (9). The protease has been shown to be required for generating infectious virus (10), and inhibitors of the protease prevent viral replication in cultured cells (11–13).

Certain methods for assaying HIV protease utilize polyprotein or oligopeptide substrates, and they require gel electrophoresis or HPLC to ascertain the presence of cleavage products (14–19). Thus, these methods are not easily adapted for assaying protease activity in large numbers of samples. Other methods recently developed for multiple samples include resonance-energy transfer (20), release of radioactive products (21), and spectrophotometry of chromogenic peptide substrates (22, 23). As a means of monitoring activity of HIV protease, we have inserted one of its recognition sites into the β-galactosidase gene such that β-galactosidase activity was retained. The altered β-galactosidase was cleaved and inactivated by HIV protease both *in vivo* and *in vitro*. This cleavage was prevented by pepstatin A, a known inhibitor of HIV protease (24). Thus, β-galactosidase has been converted into an easily assayed substrate for HIV protease. We have also inserted a polio 3C protease cleavage recognition sequence (25) into β-galactosidase and found that this site was cleaved by polio protease. These β-galactosidases bearing protease "cleavage cassettes" are potentially useful for monitoring the activity of a wide variety of proteases and perhaps for identifying potential protease inhibitors.

METHODS AND MATERIALS

Strains and Plasmids. *E. coli* strain MC1061 (*araD139*, Δ(*ara, leu*)*7697*, Δ*lacX74*, *galU*, *galK*, *hsdR*, *strA*) was obtained from Pharmacia. pZM1, containing the *E. coli* β-galactosidase gene, was derived from pCH110 (Pharmacia). The smaller *Pst* I/*Bam*HI fragment from pCH110 containing the β-lactamase gene was replaced by the corresponding *Pst* I/*Bam*HI fragment from a variant of pBR322 lacking the *Eco*RI site normally present on that fragment. Next, the *Eco*RI/*Bam*HI fragment containing the carboxyl terminus of the β-galactosidase gene was replaced with an oligonucleotide that recreates the carboxyl terminus of β-galactosidase and places two stop codons at the end of the gene.

Plasmids p1+IQ and p15−IQ, which respectively encode wild-type and mutant HIV protease under *tac* promoter control, were constructed as follows. The protease genes from pMae2 (wild type) or p48Mae6 (mutant, Asp-29 → Gly) (8) were excised as *Nde* I/*Bam*HI fragments, made blunt by Klenow enzyme, and ligated into the *Sma* I site of pKK223-3 (Pharmacia), creating p1+ and p15−. The *Sph* I/*Alw*NI fragment from p1+ and p15−, containing the HIV protease gene, was then ligated to the *Pvu* I/*Alw*NI fragment containing the *lacI*^{q} gene from pGEX2T (Pharmacia), creating p1+IQ and p15−IQ (Fig. 1*A*). To delete the *Bam*HI restriction site in the polylinker of p1+IQ and p15−IQ, these plasmids were partially digested with *Bam*HI, made blunt with Klenow enzyme, and ligated. The resultant plasmids contain a unique *Bam*HI site directly upstream of the *tac* promoter.

Cleavage Cassette Cloning. The decapeptide Val-Ser-Phe-Asn-Phe-Pro-Gln-Ile-Thr-Leu corresponds to the p6/PR cleavage site of the HIV gag/pol polyprotein and is cleaved efficiently *in vitro* by HIV protease (18). Six pairs of oligonucleotides (cleavage cassettes) encoding this decapeptide were synthesized. The DNA sequence shown below containing a *Hin*dIII restriction site (underlined) encodes the decapeptide in each oligonucleotide pair.

```
5'-GTA AGC TTT AAC TTC CCT CAG ATC ACT CTG-3'
3'-CAT TCG AAA TTG AAG GGA GTC TAG TGA GAC-5'
```

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; ONPG, *o*-nitrophenyl β-D-galactoside; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside; IPTG, isopropyl β-D-thioglactopyranoside.

*To whom reprint requests should be addressed.

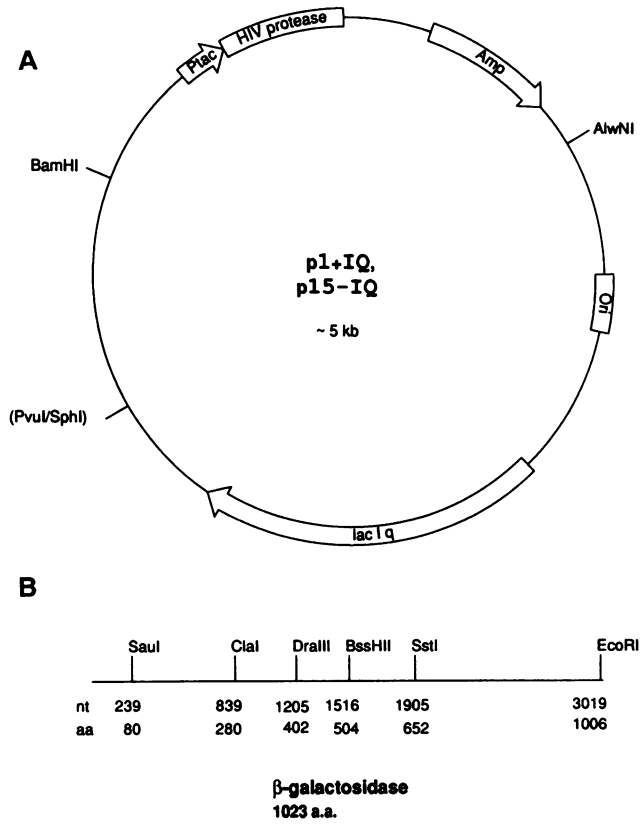

FIG. 1. (*A*) HIV protease plasmids. p1+IQ and p15−IQ encode wild-type and mutant HIV protease, respectively, under *tac* promoter control. *E. coli* strain MC1061 harboring these plasmids was used as a source of HIV protease for *in vitro* cleavage of β-galactosidase. For monitoring cleavage *in vivo*, the β-galactosidase/cleavage cassette genes were excised from the pZM plasmid series with *Stu* I and *Bam*HI and were ligated into the *Bam*HI site of p1+IQ or p15−IQ by using *Bam*HI linkers. (*B*) Insertion sites of the HIV protease cleavage cassette into β-galactosidase. Oligonucleotides encoding the HIV protease cleavage site were inserted into the indicated restriction sites of the β-galactosidase gene. The nucleotide (nt) and amino acid (aa) positions within β-galactosidase are indicated. kb, Kilobase.

Each pair contained additional nucleotides specifying different cohesive ends designed for insertion in frame into the unique *Bss*H2, *Cla* I, *Dra* III, *Eco*RI, *Sau* I, or *Sst* I restriction sites of the *E. coli* β-galactosidase gene of pZM1 (Fig. 1*B*). To simplify the cloning procedure, only the cohesive ends of each oligonucleotide pair resemble the target restriction site, so that successful ligation of the oligonucleotide into the β-galactosidase gene destroys that restriction site. Therefore, after ligation of the oligonucleotide into pZM1, the ligation mixture was digested with the target restriction enzyme to linearize any residual parent plasmid pZM1. After precipitation with ethanol, the digested ligation mix was used to transform strain MC1061. Correct clones were identified by the presence of a new *Hin*dIII site. The resultant plasmids are designated pZM-Bss, pZM-ClaA, pZM-Dra, pZM-Eco, pZM-Sau, and pZM-Sst, respectively. pZM-ClaB contains the cleavage cassette in the wrong orientation, encoding the unrelated decapeptide Glu-Ser-Asp-Leu-Arg-Glu-Val-Lys-Ala-Tyr, which should not be recognized by HIV protease. In each case, insertion of the cleavage cassette within the β-galactosidase gene was confirmed by DNA sequencing.

β-Galactosidase Assays. β-Galactosidase was quantitated by using *o*-nitrophenyl β-D-galactoside (ONPG) (26). Plates contained 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal) at 40 μg/ml.

***In Vitro* Protease Cleavage Assays.** *E. coli* strain MC1061 harboring protease plasmids p1+IQ or p15−IQ was grown to 0.5 absorbance unit at 595 nm, and protease synthesis was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to 1 mM. After 1 hr, cells were pelleted, resuspended in 0.66 vol of M buffer [50 mM 2-(*N*-morpholino)ethanesulfonic acid (Mes), pH 6/1 mM dithiothreitol/25 mM NaCl] containing 1 mM EDTA and 1 mg of lysozyme per ml and incubated on ice for 10 min. Cells were pelleted and resuspended in 0.01 vol of M buffer and were lysed by three freeze/thaw cycles (freezing rapidly in dry ice/ethanol and thawing slowly in ice water). After the final thawing, the extract was centrifuged, and the clear supernatant was used as a source of protease. The same method was used to prepare extracts containing β-galactosidase from MC1061 cells harboring the pZM plasmid series, except that IPTG was omitted and 0.1 vol of M buffer was used for the freeze/thaw cycle. For cleavage reactions, 20 μl of β-galactosidase extract was typically incubated with 5 μl of protease extract at 37°C for 1 hr, followed by electrophoresis of incubation reactions on SDS/polyacrylamide gels. β-Galactosidase was visualized by immunoblot (Western blot) analysis using mouse anti-β-galactosidase polyclonal antiserum (Sigma) and the ImmunoSelect kit (Bethesda Research Laboratories). [^{35}S]Methionine-labeled control substrate corresponding to the HIV gag polyprotein (nucleotides 109–1417) (9) was produced from the T7 promoter in *E. coli* strain BL21/pLysS (27, 28). Polio 3C protease was expressed in BL21/pLysS in similar fashion as Nicklin *et al.* (29).

RESULTS

Insertion of a HIV Protease Cleavage Site into β-Galactosidase. Studies from several laboratories using synthetic peptides that span the different cleavage sites of the gag and gag/pol polyproteins have determined the minimal amino acid sequence requirements for cleavage by HIV protease (14, 15, 18, 19). The p6/PR site is cleaved most efficiently *in vitro* (18), and as few as seven amino acids are sufficient for cleavage (14, 15). Accordingly, we designed a series of oligonucleotides (cleavage cassettes) corresponding to the 10 amino acids that span the p6/PR junction and containing the appropriate cohesive ends for insertion in frame into six unique sites of the β-galactosidase gene of plasmid pZM1 (Fig. 1*B*).

To determine the effect of cleavage cassette insertion on β-galactosidase activity, *E. coli* strain MC1061 harboring each construct was plated onto X-Gal plates (Table 1). In this assay, functional β-galactosidase results in the formation of blue colonies, but lack of β-galactosidase activity produces white colonies. Insertion of the cleavage cassette at the *Sau* I, *Cla* I, *Bss*H2, or *Eco*RI sites resulted in blue colonies, indicating that the enzymatic activity of β-galactosidase was retained. Insertion of the cleavage cassette in the opposite orientation in the *Cla* I site such that it encodes a decapeptide unrelated to the HIV protease cleavage site (Table 1, pZM-ClaB) also resulted in the formation of blue colonies. These data indicate that the β-galactosidase protein can tolerate insertions at several locations and still retain detectable enzymatic activity. Conversely, insertion at either *Dra* III or *Sst* I destroyed β-galactosidase activity, since colonies harboring these constructs were white.

The intensity of blue color in strains harboring the pZM plasmids -Bss, -ClaA, -ClaB, -Eco, and -Sau exhibited considerable variation, and in each case was less intense than that conferred by the parent plasmid pZM1 (data not shown). This variation in β-galactosidase activity was quantitated by using hydrolysis of ONPG (Table 1). As expected, white clones pZM-Dra and pZM-Sst did not produce detectable levels of β-galactosidase activity in this assay. The blue clones exhibit significantly less β-galactosidase activity than

Table 1. β-Galactosidase activity of cleavage cassette constructs and susceptibility to cleavage by HIV protease

Plasmid	β-Galactosidase assays: Colonies on X-Gal*	β-Galactosidase assays: Activity†, units	Cleavage by HIV protease‡: *In vitro*	Cleavage by HIV protease‡: *In vivo*
pZM-Bss	Blue	5	No	No
pZM-ClaA§	Blue	10	No	No
pZM-ClaB§	Blue	163	No	No
pZM-Dra	White	0	No	ND
pZM-Eco	Blue	2	No	No
pZM-Sau	Blue	63	Yes	Yes
pZM-Sst	White	0	No	ND
pZM1	Blue	1179	No	ND
None	White	0	—	—

Plasmids are described in *Materials and Methods* and in Fig. 1 and were transformed into *E. coli* host strain MC1061. ND, not done.
*Colony color on X-Gal plates.
†β-Galactosidase activity was determined by hydrolysis of ONPG and is expressed in units per ml of culture. Note that the amount of β-galactosidase/cleavage cassette protein per ml is approximately 1/10th to 1/20th the wild type level encoded by pZM1 (see text).
‡Cleavage of β-galactosidase by HIV protease was determined by Western analysis (Figs. 2 and 4).
§pZM-ClaA contains the HIV protease cleavage cassette in the proper orientation; pZM-ClaB contains this oligonucleotide in the opposite orientation.

the parent plasmid pZM1, ranging from about 1/10th (pZM-ClaB) to about 1/500th (pZM-Eco). The reduced β-galactosidase activity observed upon insertion of the cleavage cassette could result from a change in the structure of this protein attenuating its enzymatic function or from β-galactosidase/cleavage cassette protein displaying wild-type enzymatic activity but present in reduced amounts, or from a combination of these two factors.

To further examine the basis for reduced β-galactosidase activity in the β-galactosidase/cleavage cassette clones, the amount of β-galactosidase was quantitated by Western analysis with anti-β-galactosidase antibody. We found that the pZM-Sau construct produced significantly less β-galactosidase protein compared with the parent pZM1 (Fig. 2*B*, compare the intensity of the β-galactosidase band in lane 4

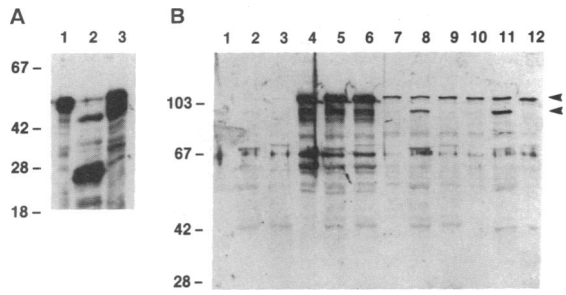

FIG. 2. (*A*) *In vitro* cleavage of gag polyprotein by HIV protease. Gag polyprotein was incubated with buffer (lane 1) or with wild-type (lane 2) or mutant (lane 3) HIV protease, followed by electrophoresis and autoradiography. (*B*) *In vitro* cleavage of β-galactosidase/cleavage cassette proteins by HIV protease. Extracts of MC1061 cells harboring HIV protease plasmid were incubated with extracts harboring β-galactosidase, followed by Western blotting with anti-β-galactosidase. Lanes: 1, MC1061 (no plasmid); 2, wild-type HIV protease extract; 3, mutant HIV protease extract; 4–6, MC1061/pZM1 extract (lane 4) incubated with wild-type (lane 5) or mutant (lane 6) HIV protease; 7–12, MC1061/pZM-Sau extract from two independent clones (lanes 7 and 10) incubated with wild-type (lanes 8 and 11) or mutant (lanes 9 and 12) HIV protease. The positions of molecular mass markers are shown to the left of each panel. Arrowheads show positions of β-galactosidase (upper) and the cleavage product from pZM-Sau (lower).

with that in lanes 7 and 10). This difference was quantitated more accurately by Western analysis of different dilutions of cells harboring pZM1 and pZM-Sau, and it was determined that the pZM-Sau clone produces 1/10th to 1/20th of the β-galactosidase protein as the parent clone pZM1 (data not shown). Thus, for pZM-Sau, a reduced amount of β-galactosidase protein with approximately wild-type levels of enzymatic activity could account for the reduction of β-galactosidase activity observed on X-Gal plates and in the ONPG assay (Table 1). Western analysis on the other cleavage cassette clones demonstrated similarly reduced levels of β-galactosidase protein (data not shown). Constructs such as pZM-Bss, pZM-Cla A, and pZM-Eco, whose β-galactosidase activities are substantially less than 1/10th to 1/20th when compared with wild type (e.g., 1/100th to 1/500th) probably encode attenuated β-galactosidase. Since the native β-galactosidase encoded by pZM1 and the altered β-galactosidase/cleavage cassette proteins are expressed from identical vectors that differ only by the insertion of the cleavage cassette in the β-galactosidase gene, it appears that the β-galactosidase/cleavage cassette proteins are less stable than the native protein.

Cleavage of β-Galactosidase/Cleavage Cassette by HIV Protease *in Vitro*. The β-galactosidase/cleavage cassette proteins were tested for susceptibility to cleavage by HIV protease *in vitro*. Extracts were prepared from *E. coli* harboring either wild-type or mutant HIV protease (p1+IQ or p15−IQ, respectively). Extract containing wild-type protease but not mutant protease is able to cleave a gag polyprotein, which is a natural substrate for HIV protease (Fig. 2*A*). Extracts of *E. coli* harboring β-galactosidase/cleavage cassette constructs were incubated with either wild-type or mutant HIV protease extract, and cleavage of β-galactosidase was monitored by Western analysis using anti-β-galactosidase antibodies (Fig. 2*B*). Native β-galactosidase encoded by pZM1 is unaffected by these incubations (Fig. 2*B*, lanes 4–6). Of the six β-galactosidase/cleavage cassette proteins examined, five of these (pZM-Bss, -ClaA, -Dra, -Eco, and -Sst) were likewise undigested by HIV protease (Table 1 and data not shown), as was the pZM-ClaB clone containing a reversed decapeptide insertion used as a control. In contrast, the β-galactosidase/cleavage cassette protein encoded by two independent pZM-Sau clones was cleaved by wild-type HIV protease (Fig. 2*B*, lanes 8 and 11), producing a cleavage product of ≈100 kDa, consistent with cleavage at the *Sau* I insertion site located 80 amino acids from the amino terminus. The other expected cleavage product of ≈8 kDa was not detected and may be unstable. Incubation of the pZM-Sau β-galactosidase/cleavage cassette with mutant HIV protease (Fig. 2*B*, lanes 9 and 12) failed to produce cleavage, indicating that the cleavage event is mediated by HIV protease and not by an *E. coli* protease.

Since pepstatin A is a known inhibitor of HIV protease (3), we tested whether cleavage of the β-galactosidase/cleavage cassette protein is inhibited by this compound. Pepstatin A at 1 μg/ml completely inhibited the cleavage of the altered β-galactosidase, further indicating that the cleavage is due to HIV protease (Fig. 3).

***In Vivo* Cleavage of β-Galactosidase/Cleavage Cassette by HIV Protease.** To test for cleavage of the β-galactosidase/cleavage cassette encoded by pZM-Sau *in vivo*, this altered β-galactosidase gene was subcloned into the plasmids p1+IQ and p15−IQ, which encode wild-type and mutant HIV protease, respectively. Total protein from cells harboring these resultant plasmids p1+IQSau and p15−IQSau was subjected to SDS/polyacrylamide gel electrophoresis, and cleavage of β-galactosidase was monitored by Western analysis (Fig. 4). This β-galactosidase/cleavage cassette was cleaved *in vivo* by wild-type HIV protease (Fig. 4*A*, lane 1, and Fig. 4*B*, lane 2) but not by mutant protease (Fig. 4*B*, lane 3), generating a

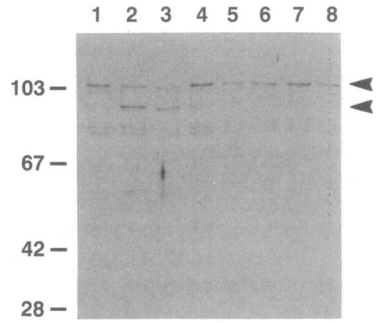

FIG. 3. Inhibition of cleavage of β-galactosidase by pepstatin A. β-Galactosidase/cleavage cassette protein from MC1061 cells harboring pZM-Sau was incubated with HIV protease in the presence or absence of pepstatin A, followed by Western blotting with anti-β-galactosidase. Lanes: 1, no protease; 2, with protease; 3–8, addition of pepstatin A dissolved in dimethyl sulfoxide (10% final concentration) to the incubation at 0, 0.2, 1, 5, 25, and 50 μg/ml. The positions of molecular mass markers are shown at left. Arrowheads indicate positions of β-galactosidase (upper) and the cleavage product from pZM-Sau (lower).

cleavage product of the same size as that observed *in vitro* (Fig. 2, lanes 8 and 11). We estimate that ≈50–75% of the β-galactosidase/cleavage cassette protein was cleaved. These cleavage assays were done on midlogarithmic-phase liquid cultures. In stationary-phase liquid cultures (Fig. 4*C*, lane 2) or in colonies scraped from plates (data not shown), we detect no uncleaved β-galactosidase/cleavage cassette protein in *E. coli* MC1061 harboring p1+IQSau. Also, as logarithmic-phase cells were incubated with IPTG for progressively longer times, β-galactosidase/cleavage cassette protein disappeared in the presence of wild-type but not mutant HIV protease (Fig. 4*C*, lanes 4–6, and data not shown). In contrast, stationary-phase *E. coli* MC1061 harboring p15−IQSau contained significant amounts of intact β-galactosidase/cleavage cassette protein (Fig. 4*C*, lane 7). Consistent with this is the observation that colonies on plates (whose physiological state is more similar to stationary-phase than to logarithmic-phase liquid culture) that harbor p1+IQSau were white in the presence of X-Gal, whereas p15−IQSau conferred blue color. These data further confirm that β-galactosidase/cleavage cassette protein is digested by wild-type but not by mutant HIV protease.

Our results *in vitro* had demonstrated that, with the exception of the pZM-Sau construct, the β-galactosidase/cleavage cassette proteins were not cleaved by HIV protease. The cleavage cassette in the protease-resistant proteins might be buried in the interior of the proteins and so be inaccessible to the protease in an *in vitro* assay. For this reason, it was useful to introduce HIV protease into the same cell as these β-galactosidase/cleavage cassette proteins, to allow for cotranslational cleavage *in vivo*. Therefore, the β-galactosidase/cleavage cassette regions of pZM-Bss, pZM-ClaA, and pZM-Eco were also subcloned into p1+IQ. The β-galactosidase produced by pZM-ClaB was used as a control. Since pZM-Dra and pZM-Sst produce nonfunctional β-galactosidase, these will not be considered further. HIV protease did not cleave any of these other β-galactosidase/cleavage cassette proteins *in vivo* that were protease-resistant *in vitro* (Fig. 4*A*, lanes 2–5, and Table 1). As expected, cells harboring p1+IQBss, p1+IQClaA, p1+IQClaB, or p1+IQEco are blue on X-Gal plates, consistent with lack of cleavage by HIV protease (data not shown).

Applicability of the β-Galactosidase/Cleavage Cassette System to Other Proteases. To examine whether other protease cleavage sites can be used in this system, an oligonucleotide encoding the polio 3C protease cleavage site Met-Glu-Ala-Leu-Phe-Gln-Gly-Pro-Leu-Gln-Tyr-Lys-Asp (25) was inserted in frame into the *Sau* I site of β-galactosidase, creating pZM3CSau. Cells harboring this plasmid were blue. These data further indicate that the *Sau* I site is an appropriate site for insertion of heterologous peptide sequences without destroying β-galactosidase activity. This hybrid protein was cleaved *in vitro* by polio 3C protease (Fig. 5, lane 5) but not by HIV protease (Fig. 5, lane 6), as expected.

DISCUSSION

We have inserted a cleavage site for HIV protease into the *Sau* I site of the β-galactosidase gene of *E. coli*. The resultant

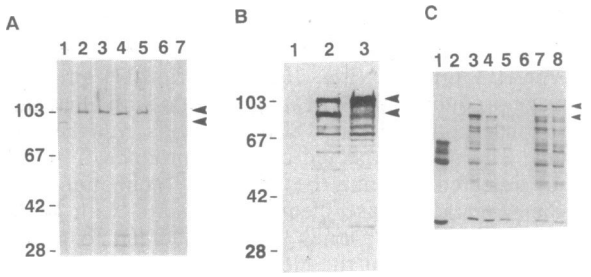

FIG. 4. *In vivo* cleavage of β-galactosidase/cleavage cassette proteins by HIV protease. (*A*) MC1061 cells harboring p1+IQSau (lane 1), p1+IQClaA (lane 2), p1+IQClaB (lane 3), p1+IQBss (lane 4), p1+IQEco (lane 5), p1+IQ (lane 6), or without plasmid (lane 7) were grown to midlogarithmic phase, and IPTG was added to 1 mM. After 30 min, cells were pelleted and resuspended in sample buffer, followed by electrophoresis and Western analysis with anti-β-galactosidase. (*B*) MC1061 cells without plasmid (lane 1) or harboring p1+IQSau (lane 2) or p15−IQSau (lane 3) were treated as in *A*. The positions of molecular mass markers are shown to the left. (*C*) Time course of cleavage by HIV protease. MC1061 cells without plasmid (lane 1) or harboring p1+IQSau (lanes 2–6) or p15−IQSau (lanes 7 and 8) were grown for the indicated times, and cleavage of β-galactosidase was monitored by Western analysis as described above. Lanes: 1, 2, and 7, stationary-phase cells (16 hr); 3 and 8, midlogarithmic-phase cells (A_{600} = 0.5); 4, 5, and 6, midlogarithmic-phase cells incubated with 1 mM IPTG for 0.5, 2, and 20 hr, respectively. The amount of cells equivalent to 0.1 ml of 0.5 A_{600} units was loaded in each lane. Arrowheads indicate positions of β-galactosidase (upper) and the cleavage product from p1+IQSau (lower).

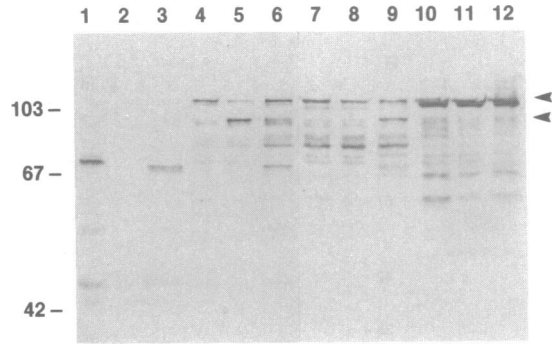

FIG. 5. Cleavage of polio protease cleavage cassette by polio protease *in vitro*. Extracts of β-galactosidase containing either the polio protease cleavage site (in MC1061/pZM3CSau) or the HIV protease cleavage site (in MC1061/pZM-Sau) were incubated with either polio or HIV protease, followed by Western analysis with anti-β-galactosidase. Lanes: 1, MC1061 without plasmid; 2, polio 3C protease; 3, HIV protease; 4–6, MC1061/pZM3CSau unincubated (lane 4) or incubated with polio (lane 5) or HIV (lane 6) protease; 7–9, MC1061/pZM-Sau unincubated (lane 7) or incubated with polio (lane 8) or HIV (lane 9) protease; 10–12, MC1061/pZM1 unincubated (lane 10) or incubated with polio (lane 11) or HIV (lane 12) protease. The positions of molecular mass markers are shown to the left. Arrowheads indicate positions of β-galactosidase (upper) and the cleavage product from the Sau constructs (lower). Lower molecular mass bands probably represent β-galactosidase degradation products and/or nonspecific reaction with the antiserum.

fusion protein retains β-galactosidase activity. Upon cleavage by HIV protease, this altered β-galactosidase in inactivated. This cleavage reaction is inhibited by pepstatin A, a known inhibitor of HIV protease. We also have demonstrated that a cleavage cassette for polio 3C protease can be inserted into β-galactosidase at the *Sau* I site. This fusion protein also retains β-galactosidase activity and is cleaved by 3C protease. We expect that this system can be adapted for other proteases as well. These data expand the use of β-galactosidase as a reporter gene to study proteolysis and represent a powerful technique for monitoring protease activity.

The observation that β-galactosidase retains enzymatic activity despite insertion of foreign protein sequences at four of six target sites used in this study indicates that this enzyme is extremely flexible. β-Galactosidase is a large protein (1023 amino acids) that forms tetramers (30); perhaps because of its size, it may have several regions that are not essential. Most missense mutations have little or no effect on β-galactosidase activity (31, 32). Previous studies have shown also that the amino terminus of β-galactosidase can be replaced by heterologous sequences (33). The phenomenon of "α complementation" is particularly interesting, in which the amino terminus of β-galactosidase (the α fragment) is unattached to the remainder of the molecule (the ω fragment) and is functional in trans (34, 35). Cleavage of our β-galactosidase/cleavage cassette inserted at the *Sau* I site removes 80 residues from the amino terminus, which is significantly larger than the M15 deletion of the ω fragment (residues 11–41) (34, 35). Since β-galactosidase activity was not observed, the cleavage products may be inherently incapable of complementation. Alternatively, the amino-terminal fragment may be too unstable to complement, since it was not detected by Western analysis.

The insertion of the decapeptide cleavage cassette at various loci within β-galactosidase is a useful tool for probing the structure of this enzyme. For example, the observation that insertion of the cleavage cassette at the *Sau* I, *Cla* I, *Bss*H2, or *Eco*RI sites does not destroy enzymatic activity may indicate that these sites are within regions of the gene that encode loops and that the decapeptide has simply enlarged the loop. Insertion at the *Dra* III or *Sst* I sites destroys enzymatic activity, suggesting that these regions of the protein are critical for β-galactosidase function and cannot tolerate alteration. Furthermore, the protease cleavage site inserted at the *Sau* I site is likely to be on the exterior of β-galactosidase, since it is accessible to proteases. Secondary structure analysis (University of Wisconsin Genetics Computer Group) of β-galactosidase indicates that all four regions that tolerated insertion are potential α-helices; turns signifying potential loops were not consistently predicted in these regions. Although the *Sau* region is predicted to be somewhat hydrophilic and therefore on the exterior of the protein, other regions that were more hydrophilic were not cleaved by protease. Secondary structure predictions are not always accurate, and confirmation must await the crystallization of β-galactosidase.

Fusion of heterologous protein to β-galactosidase is a widely used technique; however, most fusions have been to either the amino terminus (3, 33) or to the carboxyl terminus (4–6). We demonstrated that foreign protein sequences can be inserted into specific regions of β-galactosidase, resulting in both the retention of β-galactosidase activity and the accessibility of the foreign sequences to proteases. The fact that β-galactosidase can now be cleaved and inactivated by the cognate protease provides a simple means to monitor protease activity and to search for inhibitors of these important enzymes.

We thank I. Kovesdi, J. Morin, B. O'Hara, S. Plotch, and B. Rasmussen for comments on the manuscript. We thank Z. Misulovin for construction of pZM1 and R. Nilakantan for computer analysis.

1. Jacob, F. & Monod, J. (1961) *J. Mol. Biol.* **3,** 318–356.
2. Silhavy, T. J., Berron, M. L. & Enquist, L. W. (1984) *Experiments with Gene Fusions* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
3. Koenen, M., Ruther, U. & Muller-Hill, B. (1982) *EMBO J.* **1,** 509–512.
4. Ruther, U. & Muller-Hill, B. (1983) *EMBO J.* **2,** 1791–1794.
5. Germino, J. & Bastia, D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4692–4696.
6. Scholtissek, S. & Grosse, F. (1988) *Gene* **62,** 55–64.
7. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
8. Baum, E., Bebernitz, G. & Gluzman, Y. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5573–5577.
9. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., 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.
10. 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.
11. McQuade, T. J., Tomasselli, A. G., Liu, L., Karacostas, V., Moss, B., Sawyer, T. K., Heinrikson, R. L. & Tarpley, W. G. (1990) *Science* **247,** 454–456.
12. Meek, T. D., Lambert, D. M., Dreyer, G. B., Carr, T. J., Tomaszek, T. A., Jr., Moore, M. L., Strickler, J. E., Debouck, C., Hyland, L. J., Matthews, T. J., Metcalf, B. W. & Petteway, S. R. (1990) *Nature (London)* **343,** 90–92.
13. Roberts, N. A., Martin, J. A., Kinchington, D., Broadhurst, A. V., Craig, J. C., Duncan, I. B., Galpin, S. A., Handa, B. K., Kay, J., Krohn, A., Lambert, R. W., Merrett, J. H., Mills, J. S., Parkes, K. E. B., Redshaw, S., Ritchie, A. J., Taylor, D. L., Thomas, G. J. & Machin, P. J. (1990) *Science* **248,** 358–361.
14. Billich, S., Knoop, M. T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R. & Moelling, K. (1988) *J. Biol. Chem.* **263,** 17905–17908.
15. Moore, M. L., Bryan, W. M., Fakhoury, S. A., Magaard, V. W., Huffman, W. F., Dayton, B. D., Meek, T. D., Hyland, L., Dreyer, G. B., Metcalf, B. W., Strickler, J. E., Gorniak, J. G. & Debouck, C. (1989) *Biochem. Biophys. Res. Commun.* **159,** 420–425.
16. Blumenstein, J. J., Copeland, T. D., Oroszlan, S. & Michejda, C. J. (1989) *Biochem. Biophys. Res. Commun.* **163,** 980–987.
17. Dreyer, G. B., Metcalf, B. W., Tomaszek, T. A., Jr., Carr, T. J., Chandler, A. C., III, Hyland, L., Fakhoury, S. A., Magaard, V. W., Moore, M. L., Strickler, J. E., Debouck, C. & Meek, T. D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 9752–9756.
18. Krausslich, H. G., Ingraham, R. H., Skoog, M. T., Wimmer, E., Pallai, P. V. & Carter, C. A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 807–811.
19. Darke, P. L., Nutt, R. F., Brady, S. F., Garsky, V. M., Ciccarone, T. M., Leu, C. T., Lumma, P. K., Freidinger, R. M., Veber, D. F. & Sigal, I. S. (1988) *Biochem. Biophys. Res. Commun.* **156,** 297–303.
20. Matayoshi, E. D., Wang, G. T., Krafft, G. A. & Erickson, J. (1990) *Science* **247,** 954–958.
21. Billich, A., Hammerschmid, F. & Winkler, G. (1990) *Biol. Chem., Hoppe-Seyler* **371,** 265–272.
22. Nashed, N. T., Louis, J. M., Sayer, J. M., Wondrak, E. M., Mora, P. T., Oroszlan, S. & Jerina, D. M. (1989) *Biochem. Biophys. Res. Commun.* **163,** 1079–1085.
23. Richards, A. D., Phylip, L. H., Farmerie, W. G., Scarborough, P. E., Alvarez, A., Dunn, B. M., Hirel, P. H., Konvalinka, J., Strop, P., Pavlickova, L., Kostka, V. & Kay, J. (1990) *J. Biol. Chem.* **265,** 7733–7736.
24. Seelmeier, S., Schmidt, H., Turk, V. & von der Helm, K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6612–6616.
25. Pallai, P. V., Burkhardt, F., Skoog, M., Schreiner, K., Bax, P., Cohen, K. A., Hansen, G., Palladino, D. E. H., Harris, K. S., Nicklin, M. J. & Wimmer, E. (1989) *J. Biol. Chem.* **264,** 9738–9741.
26. Rothstein, D. M., Pahel, G., Tyler, B. & Magasanik, B. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 7372–7376.
27. Moffatt, B. A. & Studier, F. W. (1987) *Cell* **49,** 221–227.
28. Studier, F. W. & Moffat, B. A. (1986) *J. Mol. Biol.* **189,** 113–130.
29. Nicklin, M. J. H., Harris, K. S., Pallai, P. V. & Wimmer, E. (1988) *J. Virol.* **62,** 4586–4593.
30. Miller, J. H. & Reznikoff, W. S. (1978) *The Operon* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
31. Welply, J. K., Fowler, A. V. & Zabin, I. (1981) *J. Biol. Chem.* **256,** 6811–6816.
32. Cupples, C. G. & Miller, J. H. (1988) *Gentics* **120,** 637–644.
33. Brickman, E., Silhavy, T. J., Bassford, P. J., Jr., Shuman, H. A. & Beckwith, J. R. (1979) *J. Bacteriol.* **139,** 13–18.
34. Langley, K. E., Villarejo, M. R., Fowler, A. V., Zamenhof, P. J. & Zabin, I. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1254–1257.
35. Zabin, I. (1982) *Mol. Cell. Biochem.* **49,** 87–96.