INFECTION AND IMMUNITY, Mar. 1991, p. 1074–1078
0019-9567/91/031074-05$02.00/0

Vol. 59, No. 3

Expression and Deletion Analysis of the *Trypanosoma brucei rhodesiense* Cysteine Protease in *Escherichia coli*

ERIC G. PAMER,[1]†* CHARLES E. DAVIS,[2,3] AND MAGDALENE SO[1]

Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California 92038,[1] and Departments of Pathology and Medicine[2] and Center for Molecular Genetics,[3] University of California at San Diego, San Diego, California 92110

Received 20 September 1990/Accepted 14 December 1990

***Trypanosoma brucei*, the cause of African sleeping sickness, differentiates in the mammalian bloodstream from a long, slender trypanosome into a short, stumpy trypanosome. This event is necessary for infection of the tsetse fly and maintenance of the life cycle. We have previously shown that the stumpy form contains 10- to 15-fold-greater cysteine protease activity than either the slender form or the insect midgut procyclic, and we have isolated a cDNA encoding the protease. In order to determine whether the cDNA encodes the developmentally regulated cysteine protease, we have purified the protease from trypanosomes and have made a polyclonal antiserum against it. The trypanosomal protease gene was then expressed in *Escherichia coli* with three different methionines within the pre- and propeptides acting as initiation sites. In each case, a protein was synthesized that was recognized by an antiserum specific for the developmentally regulated trypanosomal cysteine protease. The protein synthesized from the more upstream initiation site within the propeptide was proteolytically active. The recombinant protease and the trypanosomal enzyme were identical with respect to peptide substrates and protease inhibitors. The protein remained active when synthesized in a truncated form lacking the nine consecutive prolines and carboxy-terminus extension, indicating that the terminal 108 amino acids are not necessary for proteolytic activity.**

Trypanosoma brucei is the protozoan that causes African sleeping sickness and a fatal wasting disease of domesticated animals (17). Following transmission of the infection to the mammalian host by the bite of the tsetse fly, long, slender trypanosomes multiply in the bloodstream. Typically, long, slender trypanosomes differentiate into short, stumpy trypanosomes which do not divide in the mammal. Ingestion of differentiating trypanosomes infects the midgut of the tsetse fly, where bloodstream trypanosomes transform into procyclics and eventually complete the life cycle (16). Strains of *T. brucei* that have lost the ability to differentiate from slender trypanosomes into stumpy trypanosomes are incapable of completing the life cycle in the tsetse fly (1, 2). We have shown previously that differentiation of slender trypanosomes to stumpy trypanosomes is accompanied by a 10- to 15-fold increase in the activity of a trypanosomal cysteine protease (11). Transformation to the procyclic stage results in the loss of almost all cysteine protease activity. Strains of *T. brucei* that remain slender throughout mammalian infection maintain low protease activity, which can, however, be increased 15- to 20-fold by depletion of polyamines with the ornithine decarboxylase inhibitor difluoromethylornithine (11). The regulation of this protease during the *T. brucei* life cycle and the mechanism of induction by polyamine depletion are not understood. The cDNA encoding a trypanosomal cysteine protease has recently been cloned and sequenced (9, 10). Both of the published sequences reveal typical pre- and propeptides but also a highly unusual carboxy-terminal extension.

* Corresponding author.
† Present address: Department of Immunology, University of Washington SL15, Health Sciences Building I-604, Seattle, WA 98195.

Because the cysteine protease of *T. brucei* may be important for the transition from the mammalian to the insect stage of the life cycle and because the protease may play a role in the pathogenesis of mammalian trypanosomiasis (3), we decided to investigate this enzyme in greater detail. We now report the expression of the active trypanosomal cysteine protease in *Escherichia coli*, and we find that it is recognized by antiserum directed against the developmentally regulated trypanosomal protease. The peptidolytic activity of the recombinant protease against synthetic substrates is kinetically indistinguishable from that of the trypanosomal enzyme.

MATERIALS AND METHODS

Materials. Benzyloxycarbonyl-phenylalanine-arginine-7-amino-4-methyl coumarin (z-Phe-Arg-AMC), z-arginine-arginine-AMC (z-Arg-Arg-AMC), z-phenylalanine-proline-arginine-AMC (z-Phe-Pro-Arg-AMC), and z-valine-leucine-arginine-AMC (z-Val-Leu-Arg-AMC) were obtained from Enzyme Systems Laboratories, Livermore, Calif. *Thermus aquaticus* DNA polymerase was obtained from Perkin-Elmer Cetus, Norwalk, Conn. Restriction enzymes were obtained from Boehringer Mannheim Biochemicals, Indianapolis, Ind.

Purification of protease from *T. brucei* and antibody formation. *T. brucei* GUTat 3.1 was grown in mice and harvested as previously described (11). For protease purification, 1.0×10^{10} to 1.5×10^{10} stumpy trypanosomes were used as the starting material. Trypanosomes were lysed by suspension in 5 ml of 50 mM imidazole buffer (pH 6.0) containing 0.1% Triton X-100, followed by four cycles of freeze-thawing. The lysate was cleared by centrifugation at 100,000 × *g*. Protein that was precipitated with 33 to 75% saturated ammonium

sulfate was resuspended in 2 ml of 1 M ammonium sulfate–100 mM imidazole (pH 6.0)–1 mM dithiothreitol (DTT) and applied to a 10-ml phenyl-Sepharose column (Pharmacia, Uppsala, Sweden). This column was washed with decreasing concentrations of ammonium sulfate and then washed sequentially with 50 mM imidazole–1 mM DTT and 25 mM imidazole–1 mM DTT (pH 6.0 for both solutions). The protease was then eluted at a rate of 1 ml/min with distilled water containing 1 mM DTT, resulting in approximately 20-fold purification and a 40% yield. The eluted protease was applied immediately to a DEAE-Sepharose column (Pharmacia) equilibrated in 100 mM imidazole–1 mM DTT. The column was washed with 100 mM imidazole–100 mM NaCl–1 mM DTT until no more protein could be detected in the eluate. The protease was eluted at a rate of 1 ml/min with 200 mM NaCl in buffer, resulting in approximately 8- to 10-fold-further purification and a total yield of 15 to 20%. The active fractions were diluted fourfold with distilled water and applied to a 0.5-ml DEAE column. The column was washed with 3.0 ml of distilled water, and the protease was eluted with 1 M ammonium bicarbonate in a total volume of 1 ml. This material was lyophilized and then was suspended in Laemmli sample buffer (4). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed, and the 29-kDa area, as determined by the location of prestained carbonic anhydrase (30 kDa), was excised, cut into minute particles, and sonicated. The sonic extract was mixed 1:1 with Freund's complete adjuvant (first injection) and Freund's incomplete adjuvant (second and third injections) and was injected subcutaneously into New Zealand White rabbits. Serum was harvested 1 week after the third injection.

Synthesis of the protease in *E. coli*. Oligonucleotide-directed mutagenesis was used to create *Nde*I sites at three methionine codons within the putative pre- and propeptide-encoding region of the trypanosomal cysteine protease cDNA, which was inserted into the Bluescript plasmid (Stratagene, La Jolla, Calif.) to create pTCP (10). We used the polymerase chain reaction with either oligonucleotide S1 (5'TTTGAGGAACATATGGAGCAGGCG3'), oligonucleotide S2 (5'GAGTCATTGCATATGCGTTTTGCT3'), or oligonucleotide S5 (5'GTCTTGCTGCATATGGCAGCGTG C3') and the KS primer of Bluescript (Stratagene) to generate DNA fragments from 0.1 μg of pTCP with an *Nde*I site at the potential start codons. In order to delete the carboxy-terminus extension of the cysteine protease, we introduced a stop codon at the first of the nine consecutive proline residues. This was done by polymerase chain reaction with oligonucleotide S2 and oligonucleotide Stop1 (5'GGGAAT TCTGGTGGTGGTCAAGTGGGACTCC3'), which also added an *Eco*RI restriction site at the 3' end of the product. These DNAs were digested with *Nde*I and *Eco*RI and ligated into *Nde*I-*Eco*RI-digested pRK172 (8), yielding pRK172S1, pRK172S2, pRK172S5, and pRK172S2S1. These plasmids were introduced into *E. coli* HB101 and purified in a CsCl gradient. The purified plasmid was introduced into *E. coli* BL21(DE3) (15), and transformants were selected by plating onto Luria-Bertani carbenicillin plates. Colonies were inoculated into 5 ml of Luria-Bertani carbenicillin broth, grown to mid-log phase, and induced with 0.4 mM isopropyl-D-thiogalactopyranoside (IPTG) for 3 h. Whole-cell preparations for SDS-PAGE were made by resuspending cell pellets obtained from 1 ml of culture in 100 μl of 6 M urea–1% SDS–1% mercaptoethanol–10 mM sodium phosphate, pH 7.2 (cracking buffer). Insoluble preparations were made by suspending the bacterial pellet from 1 ml of culture in 200 μl of TEN (300 mM NaCl–0.5 mM EDTA–50 mM Tris HCl, pH 7.5) and adding lysozyme to a final concentration of 1 mg/ml and Triton X-100 to a final concentration of 1%. After the preparations were incubated overnight at 4°C, we added 1.5 volumes of 1.5 M NaCl–12.0 mM $MgCl_2$ and DNase I to a final concentration of 10 μg/ml. After 2 more hours of incubation at 4°C, the insoluble protein was pelleted in a microcentrifuge, washed twice with TEN, and suspended in 200 μl of cracking buffer.

PAGE and Western blotting (immunoblotting). Electrophoresis was performed on SDS–10% polyacrylamide gels as described by Laemmli (4). Proteins were stained with Coomassie blue. Gelatin substrate gels were run as described previously (11). For antibody probing, proteins were transferred electrophoretically from SDS-polyacrylamide gels to BA 85 nitrocellulose paper (Schleicher & Schuell, Keene, N.H.). Following transfer, the nitrocellulose was blocked with 5% low-fat milk, probed overnight at room temperature with a 1:1,000 dilution of rabbit anti-trypanosomal cysteine protease antiserum, and developed with alkaline phosphatase-linked anti-rabbit immunoglobulin G (Promega, Madison, Wis.) according to the manufacturer's directions.

Protease assays and kinetic determination. We assayed the activity of the *T. brucei* and recombinant cysteine proteases against z-Phe-Arg-AMC, z-Arg-Arg-AMC, z-Val-Leu-Arg-AMC, and z-Phe-Pro-Arg-AMC as described previously (11). In order to determine K_m values for the recombinant and trypanosomal proteases, we partially purified both enzymes. Soluble fractions from 5×10^9 stumpy trypanosomes and 5 ml of mid-log-phase IPTG-induced *E. coli* lysates were precipitated with 33 to 75% saturated ammonium sulfate, dialyzed against 100 mM imidazole (pH 6.0), and applied to a 10-ml DEAE-Sepharose column. The column was washed with 100 mM imidazole (pH 6.0)–100 mM NaCl–1 mM DTT until no further protein could be detected in the eluate. The protease was then eluted with 200 mM NaCl–100 mM imidazole–1 mM DTT. The K_m for each protease was considered to be the substrate concentration with half the maximal rate of substrate hydrolysis.

RESULTS

The trypanosomal cysteine protease was purified by ammonium sulfate precipitation and by hydrophobic-interaction and anion-exchange chromatography. Recovery of the enzyme was hampered by its extreme lability under the conditions required for purification. In the absence of DTT, the protease associated with other cellular proteins. Under reducing conditions, the purified protease was active but underwent autodegradation, with a half-life at 4°C of about 15 min. The purity of the trypanosomal cysteine protease was difficult to assess with certainty since the purified enzyme on SDS-polyacrylamide gels would not stain with either Coomassie blue or silver nitrate. Protease activity in gelatin substrate gels was detected, with the purified enzyme in the 28-kDa region (results not shown). Antiserum from a rabbit immunized with the 28-kDa region of a polyacrylamide gel containing purified protease reacted at a high titer with a Western blot of whole trypanosomal lysate. The antiserum recognized three protein bands (Fig. 1, lane C) that correspond to three areas of proteolysis on a gelatin substrate gel (Fig. 1, lane B). This suggests that these three proteolytic bands are antigenically related and may be processing intermediates of the enzyme. The specificity of

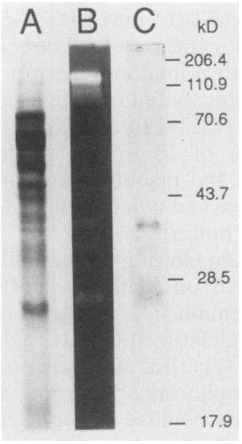

FIG. 1. Specificity of the anti-trypanosomal cysteine protease antiserum. Lanes: A, Coomassie blue-stained 10% polyacrylamide gel of 5×10^6 stumpy trypanosomes; B, gelatin substrate polyacrylamide gel of 2×10^6 stumpy trypanosomes, showing three bands of proteolytic activity at 28, 35, and 38 kDa; C, Western blot of 5×10^6 stumpy trypanosomes probed with a 1:500 dilution of the rabbit anti-trypanosomal cysteine protease serum. Protein molecular masses are indicated on the right.

the antiserum was supported by the finding that it could precipitate 100% of cysteine protease activity from trypanosomal lysates (results not shown).

In order to express the trypanosomal cysteine protease in *E. coli*, we cloned the coding sequence of its cDNA into expression vector pRK172 (8). This plasmid contains a T7 promoter and a ribosome binding sequence upstream of an *Nde*I and *Eco*RI cloning site. In order to insert the *T. brucei* protease sequence into this vector, we converted nucleotides 85 to 90, 145 to 150, and 235 to 240 into *Nde*I sites by oligonucleotide-directed mutagenesis and the polymerase chain reaction. The resulting fragments were cloned into pRK172 and introduced into *E. coli* BL21(DE3), which contains an IPTG-inducible T7 RNA polymerase (15). Figure 2 shows the parts of the open reading frame contained in the inserts and the expected molecular masses of their protein products. Figure 3A shows a polyacrylamide gel containing the recombinant proteins before and after induction by exposure to 0.4 mM IPTG for 3 h. The recombinant proteins had the expected molecular masses of 40 kDa for START1, 44 kDa for START2, and 46 kDa for START5 (Fig. 3). As is typical for many proteins that are overexpressed in *E. coli* (5), the recombinant proteins expressed

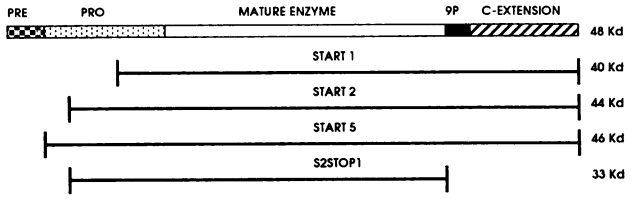

FIG. 2. Diagram illustrating the extents of the protease cDNAs encoded by pRK172S1 (START1), pRK172S2 (START2), pRK 172S5 (START5), and pRK172S2S1 (S2STOP1) and the expected molecular masses of the anticipated protein products. PRE and PRO indicate the putative pre- and propeptides. The segment composed of nine consecutive proline residues is indicated (9P) and is followed by the unusual carboxy-terminus extension (C-EXTENSION).

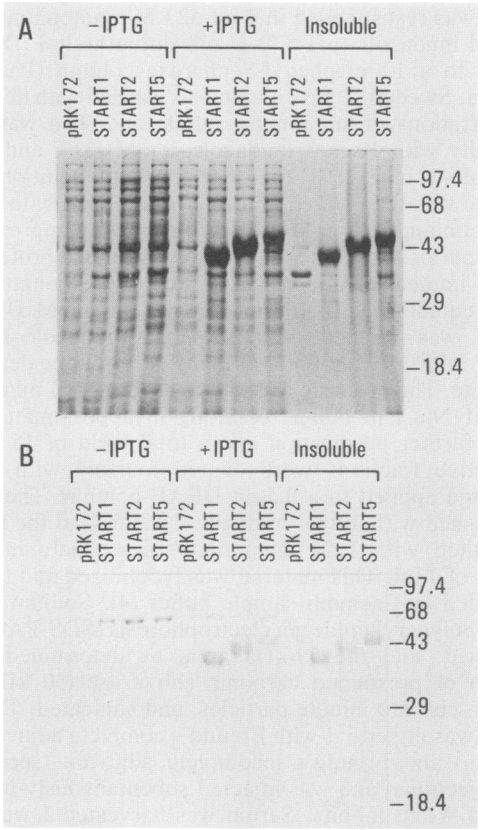

FIG. 3. (A) Expression of the trypanosomal cysteine protease in *E. coli*. Shown are SDS–10% polyacrylamide gels of whole-cell lysates of *E. coli* BL21(DE3) containing pRK172, pRK172S1, pRK172S2, and pRK172S5 before and after induction with 0.4 mM IPTG. The last four lanes represent insoluble protein preparations (see Materials and Methods) of the same *E. coli* strains in the same order. To each lane was added 2 to 10 μl of sample in cracking buffer, prepared as described in Materials and Methods. Protein molecular masses are indicated on the right. (B) Western blot of a polyacrylamide gel (identical to the one in panel A) probed, as described in Materials and Methods, with a 1:1,000 dilution of antibody specific for the *T. brucei* cysteine protease.

from pRK172S1, pRK172S2, and pRK172S5 were highly insoluble. Extraction of the insoluble protein from the induced bacteria greatly enhanced the purity of the recombinant protein (Fig. 3).

In order to determine whether the recombinant proteins represented the developmentally regulated trypanosomal enzyme, they were probed with the cysteine protease-specific antiserum. Figure 3B shows that START1, START2, and START5 are all recognized by the antiserum.

Although the recombinant proteins were found to be insoluble, it was possible that a proportion of the induced protein might be soluble and active. We therefore tested lysates of IPTG-induced BL21(DE3) containing pRK172, pRK172S1, pRK172S2, and pRK172S5 for peptidolytic activity with the cysteine protease substrate z-Phe-Arg-AMC. Only bacteria containing pRK172S2 demonstrated peptidolytic activity. The lysate of a 1-ml mid-log-phase IPTG-induced culture of BL21(DE3) transformed with pRK172S2 contained between 8 and 12 U of activity (1 U of activity produces 1 μmol of AMC per h from z-Phe-Arg-AMC [pH

TABLE 1. Determination of recombinant and *T. brucei* protease K_ms

Protease substrate	K_m (μmol/liter of substrate) of protease		
	T. brucei	Recombinant START2	Recombinant S2STOP1
z-Phe-Arg-AMC	0.85	0.95	1.04
z-Val-Leu-Arg-AMC	0.25	0.29	0.43
z-Phe-Pro-Arg-AMC	10.0	11.0	13.2
z-Arg-Arg-AMC	3.2	3.1	4.2

6.0] at 25°C in the presence of 2 mM DTT). This is equivalent to the amount of activity obtained from 10^6 stumpy trypanosomes. In order to determine whether this activity, which resides in the soluble fraction, represented the recombinant trypanosomal enzyme, the K_m of the protease was determined for four different peptide substrates. In all cases, the K_ms for the trypanosomal and recombinant enzymes were nearly identical (Table 1). This identity, along with the shared activation of the recombinant enzyme by DTT and inhibition by the specific cysteine protease inhibitor E-64 [*trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino) butane], indicates that the peptidolytic activity in the *E. coli* lysates represents the recombinant trypanosomal cysteine protease. Preliminary studies examining the molecular weight of the active recombinant protease reveal that there are at least four discrete size classes of protein with proteolytic activity, suggesting that there is some processing of the enzyme in *E. coli* (results not shown).

The carboxy-terminus extension of the protease was deleted from the coding sequence in order to determine whether it is necessary for the activity of the enzyme. Upon induction with IPTG, lysates of BL21(DE3) transformed with pRK172S2S1, also produced between 8 and 12 U of peptidolytic activity per ml of culture when tested against z-Phe-Arg-AMC. K_m determinations with four peptide substrates again showed great similarity to those for the trypanosomal cysteine protease (Table 1).

DISCUSSION

The proteases of many pathogenic organisms have been implicated in the pathogenesis of the diseases they cause (7). It has also been suggested that the cysteine protease of *T. brucei* is responsible for some of the manifestations of African trypanosomiasis (3). As the first step in an analysis of the role of the protease in pathogenesis, we have cloned and sequenced its cDNA (10). We have expressed the trypanosomal cysteine protease gene in *E. coli* with a T7 promoter fused to three methionine codons within the same open reading frame. When expressed in *E. coli*, the cDNA we have cloned encodes a protein that is immunologically cross-reactive with and kinetically indistinguishable from the developmentally regulated trypanosomal enzyme. All three recombinant proteins were recognized by immune serum directed against purified trypanosomal protease, but the only active recombinant protein contained the longer propeptide sequence without any of the prepeptide. It is believed that the propeptide is necessary for proper folding of proteases, and recent studies have shown that the propeptide can serve this function without covalent linkage to the mature enzyme sequence (13). Our results indicate that amino acids 38 to 68 of the propeptide are necessary for the activity of the trypanosomal enzyme in *E. coli*. Although START5 contains the whole propeptide sequence, it also contains the last two amino acids of the putative prepeptide. It is not clear at this time why this insert does not encode a proteolytically active protein.

The vector system we used provided excellent yields of recombinant protein that could be substantially purified by extracting only the insoluble protein. Recently, Smith and Gottesman (14) reported the expression of active human cathepsin L in *E. coli*, using a vector system very similar to the one used in our studies. They also found that the propeptide was necessary for the production of active recombinant enzyme. In the case of cathepsin L, however, the carboxy terminus appears to be necessary for the development of active enzyme. The trypanosomal cysteine protease has an unusual carboxy-terminus extension that appears to be unnecessary for peptidolytic activity.

We have shown previously (11) that production of the cysteine protease of *T. brucei* is developmentally regulated and inducible by polyamine depletion. Since cysteine proteases can interfere with the proper processing of antigens by antigen-presenting cells (6) and since trypanosomiasis inhibits the priming of trypanosome-specific T cells (12), we are currently investigating the effect of the trypanosomal cysteine protease on cell-mediated responses to trypanosomal infection. Expression of large quantities of active trypanosomal protease in *E. coli* will greatly facilitate these studies, provide material for detailed substrate and inhibitor studies of the enzyme, and make it possible to study the regulation of this enzyme in trypanosomes at several molecular levels.

REFERENCES

1. **Ashcroft, M. T.** 1960. A comparison between a syringe passaged and a Tsetse fly transmitted line of a strain of *Trypanosoma rhodesiense*. Ann. Trop. Med. Parasitol. **54**:44–53.
2. **Hajduk, S. L., and K. Vickerman.** 1981. Absence of detectable alteration in the kinetoplast DNA of a *Trypanosoma brucei* clone following loss of ability to infect the insect vector. Mol. Biochem. Parasitol. **4**:17–28.
3. **Knowles, G., G. Abebe, and S. J. Black.** 1989. Detection of parasite peptidase in the plasma of heifers infected with *Trypanosoma congolense*. Mol. Biochem. Parasitol. **34**:25–34.
4. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **277**:680–682.
5. **Marston, F. A. O.** 1986. The purification of eukaryotic polypeptides synthesized in *E. coli*. Biochem. J. **240**:1–12.
6. **McCoy, K., S. Gal, R. Schwartz, and M. M. Gottesman.** 1988. An acid protease secreted by transformed cells interferes with antigen processing. J. Cell Biol. **106**:1879–1884.
7. **McKerrow, J. H.** 1989. Parasite proteases. Exp. Parasitol. **68**:111–115.
8. **McLeod, M., M. Stein, and D. Beach.** 1987. The product of the *mei3* gene, expressed under control of the mating-type locus, induces meiosis and sporulation in fission yeast. EMBO J. **6**:729–736.
9. **Mottram, J. C., M. J. North, J. D. Barry, and G. H. Coombs.** 1989. A cysteine proteinase cDNA from *Trypanosoma brucei* predicts an enzyme with an unusual C-terminal extension. FEBS Lett. **258**:211–215.
10. **Pamer, E. G., C. E. Davis, A. E. Eakin, and M. So.** 1990. Cloning and sequencing of the *Trypanosoma brucei rhodesiense* cysteine protease. Nucleic Acids Res. **18**:6141.
11. **Pamer, E. G., M. So, and C. E. Davis.** 1989. Identification of a developmentally regulated cysteine protease of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **33**:27–32.
12. **Paulnock, D. M., C. Smith, and J. M. Mansfield.** 1988. Antigen presenting cell function in African trypanosomiasis, p. 135–143.

In L. B. Schook and J. G. Tew (ed.), Antigen presenting cells: diversity, differentiation and regulation. Alan R. Liss, New York.
13. **Silen, J. L., and D. A. Agard.** 1989. The alpha-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. Nature (London) **341**:462–464.
14. **Smith, S. M., and M. M. Gottesman.** 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. J. Biol. Chem. **264**:20487–20495.
15. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
16. **Vickerman, K.** 1965. Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. Nature (London) **208**:762–766.
17. **Vickerman, K.** 1985. Developmental cycles and biology of pathogenic trypanosomes. Br. Med. Bull. **41**:105–114.