THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 12, Issue of March 21, pp. 10458–10464, 2003

Printed in U.S.A.

Sequential Autolytic Processing Activates the Zymogen of Arg-gingipain*

Received for publication, October 15, 2002, and in revised form, January 15, 2003
Published, JBC Papers in Press, January 17, 2003, DOI 10.1074/jbc.M210564200

Jowita Mikolajczyk‡, Kelly M. Boatright‡, Henning R. Stennicke§, Tamim Nazif¶, Jan Potempa‖, Matthew Bogyo¶, and Guy S. Salvesen‡**

From ‡The Burnham Institute, La Jolla, California 92037, the §Department of Protein Design, Novo Nordisk, DK-2880 Bagsvaerd, Denmark, the ¶Department of Biochemistry and Biophysics, University of California, San Francisco, California 94043, and the ‖Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602 and the Faculty of Biotechnology, Department of Microbiology, Jagiellonian University, Krakow 30-060, Poland

Most proteases are synthesized as inactive precursors to protect the synthetic machinery of the cell and allow timing of activation. The mechanisms used to render latency are varied but tend to be conserved within protease families. Proteases belonging to the caspase family have a unique mechanism mediated by transitions of two surface loops, and on the basis of conservation of mechanism one would expect this to be preserved by caspase relatives. We have been able to express the full-length precursor of the Arg-specific caspase relative from the bacterium *Porphyromonas gingivalis*, Arg-gingipain-B, and we show that it contains N- and C-terminal extensions that render a low amount of latency, meaning that the zymogen is substantially active. Three sequential autolytic processing steps at the N and C terminus are required for full activity, and the N-propeptide may serve as an intramolecular chaperone rather than an inhibitory peptide. Each step in activation requires the previous step, and an affinity probe reveals that incremental activity enhancements are achieved in a stepwise manner.

Proteases of the gingipain family are virulence factors of the periodontal pathogenic bacterium *Porphyromonas gingivalis* (1–3). This group contains two genes that encode Arg-specific proteases (*rgpA* and *rgpB*) and one gene encoding a Lys-specific protease (*kgp*). Gene ablation studies have shown that RgpA and RgpB are required for the activation of Kgp, placing the Arg-specific proteases at the top of a proteolytic pathway required for bacterial growth (4, 5). The protein encoded by *rgpB* is predicted to consist of three distinct segments (see Fig. 1), but only two are found in the mature product isolated from bacterial cultures, the catalytic unit and an Ig domain. Consequently the 205-amino acid N-terminal segment may constitute an activation peptide that restrains the activity of the protease until it reaches its site of action.

Structural analysis of the catalytic unit of RgpB (6) demonstrates that it shares its evolutionary origin with a common ancestor of caspases, proteases involved in apoptosis and cytokine activation (7). Moreover, homology mapping suggests that the clan encompassing gingipains and caspases also may contain bacterial clostripain, plant and animal legumains (processing proteases) (8), and separase (required for sister chromatid separation during anaphase) (9). This clan is known as protease clan CD (10) or the caspase-hemoglobinase fold (11).

The majority of proteases are synthesized as zymogens that await activation at a suitable time to protect the biosynthetic machinery of the cell against activation and to act as a timing event in biological function (12). Thus, one of the key events in any proteolytic pathway is the conversion of the zymogen to the active enzyme. Different protease clans utilize distinct strategies for zymogen maintenance and activation, but within clans there seems to be conservation of a particular strategy. On the basis of conservation of mechanism one would imagine that protease clan CD would embrace a similar activation pathway, meaning that the zymogens of caspases and RgpB should be stabilized by homologous molecular interactions. The molecular determinants of caspase activation have been elucidated (13–15), and this seems not to involve the removal of N-terminal segments common for other protease clans (12).

Consequently the understanding that the precursor of RgpB (pro-RgpB) may require truncation at its N terminus (or even C terminus) for its activation serves as a good model to test the conservation hypothesis for protease zymogen activation since these observations would seem to contrast with the caspase activation mechanism. This study presents a detailed investigation of the autocatalytic processing of recombinant pro-RgpB, including the characterization of intermediates on the activation pathway, to clarify the mechanism of pro-RgpB maturation.

MATERIALS AND METHODS

Strains and Media—Saccharomyces cerevisiae strain YG227 (Mat**a**,Δ*alg6*::*HIS3*, *ade2–101*, *his3Δ200*, *ura3–52*, *lys2–801*) was kindly provided by Markus Aebi (Institute of Microbiology, ETH Zentrum, Zurich). *Escherichia coli* DH5α was used as the host for construction and propagation of all plasmids. Standard and synthetic media were prepared and supplemented with nutrients appropriate for selection and maintenance of plasmids as described previously (16). Yeast cells were grown in 2% glucose as carbon source and 2% galactose and 1% raffinose to induce protein expression from *GAL1* promoter.

Plasmid Construction—The pRS316Gal(Δ*Acc*65I) expression vector was produced for internal laboratory purposes as follows, although the deletion of *Acc*65I was not specifically necessary for the cloning strategy used in this report. The pRS316(Δ*Acc*65I) was first generated by digestion of pRS316 (17) with *Acc*65I and blunt ending of the overhang by T4 polymerase and religation in the presence of *Acc*65I. Clones lacking the *KpnI*/*Acc*65I site were identified by restriction enzyme digest and used to generate the expression vector. The sequence encoding the *Gal1* promoter was amplified from genomic yeast DNA using primers Gal1f

* This work was supported by National Institutes of Health Grant CA69381 (to G. S. S.), California Breast Cancer Research Program Fellowship 8GB-0137 (to K. M. B.), and National Institutes of Health Grant DE009761 and Committee of Scientific Research, Poland Grant KBN-6 P04A 047 17 (to J. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed: The Burnham Inst., 10901 North Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-646-3114; Fax: 858-713-6274; E-mail: gsalvesen@burnham.org.

(5-ttggagctcacatggcattaccaccatatacatatc-3) and Gal1r (5-agagcggccgccggtaccgttttttctccttgacgttaaagt-3). This fragment was introduced intro pRS316(Δ*Acc*65I) as a *Sac*I-*Not*I insert resulting in pRS316Gal plasmid. The full-length pro form of RgpB was amplified from genomic DNA isolated from *P. gingivalis* strain HG66. The 24-residue signal peptide of gingipain was replaced by the signal peptide of yeast carboxypeptidase Y amplified from pRA21 plasmid (18). The insert was then ligated into the yeast expression plasmid pRS316Gal, and insertion of a FLAG epitope sequence (DYKDDDDK) at the C terminus resulted in the pro-RgpB FL[1]-WT construct (Fig. 1). The *Xho*I restriction site encoding two amino acids (Leu and Glu) separates the FLAG tag sequence from the authentic C-terminal sequence of pro-RgpB. Specific mutants (Fig. 1) were constructed using an overlap polymerase chain reaction technique and the pro-RgpB FL-WT constructs as a template. The double mutant[2] $Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$ was generated in the same way but with pro-RgpB FL $Arg^{-103} \rightarrow Ala$ as a template. All constructs were sequenced completely to confirm that no undesired mutations were present.

Production of Recombinant Progingipain in Yeast Cells—Yeast cells were transformed by the lithium acetate method, and transformants were selected on SC plates without uracil (SC−uracil) for auxotrophic selection (19). Large scale expression was performed as follows. Single colony transformants were used to inoculate 20 ml of SC−uracil medium, and cultures were grown at 30 °C for 24 h followed by inoculation to 1.0 liter of medium, and the culture was allowed to grow for 12 more h. Next cells were washed in water and inoculated into induction medium (SC−uracil, 2% galactose, 1% raffinose). Cells were harvested by centrifugation 6 h after induction, washed in water, and resuspended in 20 mM Bis-Tris, 10 mM NaCl, 1 mM $CaCl_2$, pH 6.5 containing the protease inhibitors 1 mM phenylmethanesulfonic acid and 100 μM 3,4-dichloroisocoumarin (20). A half-volume of glass beads was added, and the cells were mechanically broken in a bead beater. Cell debris were removed by centrifugation at 18,000 × *g* for 30 min followed by filtration, and the supernatant fluid was used for further purification.

Purification of Recombinant Progingipain—Yeast extract was applied to a Sepharose Q (HiTrapQ HP5, Amersham Biosciences) column equilibrated with 20 mM Bis-Tris, pH 6.5 and washed with 5 column volumes of the same buffer following which bound protein was eluted with a two-step gradient (0–300 mM NaCl, 20 column volumes; and 300–500 mM NaCl, 5 column volumes). Fractions were assayed for activity, and Western blot analysis was performed using an enhanced chemiluminescence detection system with anti-FLAG monoclonal antibodies (Sigma).

Fractions containing recombinant protein were pooled and further purified by immunoaffinity chromatography by binding to the to M-2 anti-FLAG agarose gel (Sigma) slurry overnight at 4 °C. Next the gel was washed on the column with TBS buffer (20 mM Tris, 137 mM NaCl, pH 7.6). Bound protein was eluted with FLAG-peptide (100 μg/ml). 1-ml fractions were collected and assayed for active enzyme or by Western blot with anti-FLAG antibodies. Fractions containing the recombinant protein were pooled and store at −70 °C. Identification of the recombinant protein was achieved by using both N-terminal sequencing and mass spectrometry (MALDI-TOF).

Processing of Zymogen—Zymogen processing was analyzed in two ways. First, purified pro-RgpB was subjected to self-processing by incubating the recombinant protein in assay buffer (200 mM Tris, pH 7.6, 100 mM NaCl, 5 mM $CaCl_2$, 10 mM cysteine) for 2 h at 37 °C. Part of the reaction mixture was used for affinity labeling and activity assay; the other part of the reaction was stopped by adding 100 μM leupeptin, and cleavage products were analyzed by Western blot with anti-FLAG antibodies or antiserum to the mature RgpB (21). Second, purified pro-RgpB was tested for self-processing at different concentrations to test for inter- *versus* intramolecular activation. The RgpB antiserum was raised against a peptide corresponding to the N-terminal 35 residues of mature RgpB and preferentially recognizes denatured protein. It does not require a free N terminus for reactivity as demonstrated under "Results."

Synthesis of Ac-biotinyl-Lys-Tyr-6-aminohexanoic-Arg-acyloxymethyl Ketone (BiRK)—All chemicals used in the synthesis of the acyloxymethyl ketone were purchased from Advanced Chemtech and Sigma. The biotinylated inhibitor BiRK was synthesized by a solid-phase method from an arginine chloromethyl ketone according to the procedure described in Ref. 22 with minor modifications. The chloromethyl ketone was synthesized essentially as described using Fmoc-Arg(Pbf)-OH where Fmoc is *N*-(9-fluorenyl)methoxycarbonyl and Pbf is 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (23). Following cleavage from the matrix and deprotection, BiRK was purified on a Waters C-18 reverse phase high pressure liquid chromatography column and verified by mass spectrometry.

Determination of Inhibition Constants of RgpB by BiRK—The association rate of BiRK was determined by titration against RgpB of known activity purified from *P. gingivalis* (24). Enzyme was used at a concentration of 0.25 nM. Enzymatic activity was measured by cleavage of Boc-FPR-AMC as described above. Inhibition of RgpB with BiRK was determined by progress curve analysis as follows. Enzyme was pre-equilibrated in assay buffer for 10 min at 37 °C and then added to a pre-equilibrated reaction mix containing 0–10 nM BiRK and 2 mM Boc-FPR-AMC in assay buffer to a final volume of 100 μl. Kinetic constants were obtained by a nonlinear least-squares fit of the data to the equation: $y = (v_s t - (v_s - v_0)(1 - \exp^{-kt})/k_{obs}) - A$. The slope of the plot of k_{obs} *versus* inhibitor concentration was used to determine the second order rate constant $k_a = 3.5 \times 10^5\ M^{-1}s^{-1}$.

Active Site Labeling—Active site labeling was performed with a trace (non-saturated) and a high (saturated) biotinylated affinity probe concentration. The time and concentration required to obtain trace and saturation probe binding were estimated according to the equation $t_{½} = \ln 2/k_a\ [I]$ where I is the concentration of inhibitor required for free enzyme to decrease by 50% (half-life $t_{½}$). A 30-μl portion of purified recombinant proenzyme (126 ng) subjected to autoprocessing was activated in assay buffer for 10 min at 37 °C. BiRK was added to a final concentration of 10 nM for 5 min at room temperature (trace probe condition) and 5 μM for 45 min at room temperature (saturating probe condition). Active site labeling was terminated by adding leupeptin (100 μM final concentration) for 5 min followed by boiling samples in SDS-PAGE sample buffer.

Enzyme Activity Assay—Routinely the activity of recombinant forms of gingipain was determined by recording the release of 7-amino-4-methylcoumarin (AMC) generated by cleavage of Boc-QGR-AMC (100 μM) at 37 °C by measuring the increase in fluorescence at excitation/emission 380/490 nm using an *f*max fluorescence microplate reader (Amersham Biosciences) operating in the kinetic mode. Assays were performed in 100 μl of assay buffer. Specific activity was defined as the amount of AMC released/min/μg of purified recombinant protein used for assay. The catalytic parameters K_m and k_{cat} were calculated by a non-linear regression fit to the Michaelis-Menten equation using substrate at a concentration ranging from 2.3 to 300 μM. Active enzyme concentration of fully processed recombinant wild-type RgpB was based on active site titration with leupeptin. Because of the low concentration of double $Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$ pro-RgpB mutant and its relatively poor activity, we were not able to accurately evaluate the enzyme concentration by active site titration. Therefore we estimated the enzyme concentration based on protein concentration determined by a modified Bradford assay (Pierce) with adjustment for the active pro-RgpB using incorporation of BiRK.

RESULTS

Expression of Recombinant Pro-RgpB Mutants—Heterologous expression of soluble and active full-length or mutant forms of RgpB in *E. coli* was unsuccessful. Therefore we adapted the constructs to a *Saccharomyces* expression system that is engineered to drive synthesis and secretion. We added a FLAG epitope tag at the C terminus of all constructs for ease of purification and identification. Recombinant proteins were purified by a two-step protocol using anion exchange (Q-Sepharose) followed by anti-FLAG immunoaffinity chromatography and analyzed initially by Coomassie protein staining (Fig. 2*A*) and Western blot using a FLAG antiserum (Fig. 2*B*). Yields were low and ranged from 10 to 20 μg of protein/liter of yeast culture. Overexpression of the full-length pro-RgpB resulted predominantly in the expected 82-kDa protein but also in partial processing demonstrated by bands of lower molecular mass (68 and 56 kDa). The N-terminal sequence could not be ob-

[1] The abbreviations used are: FL, full-length; WT, wild-type; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; BiRK, Ac-biotinyl-Lys-Tyr-6-aminohexanoic-Arg-acyloxymethyl ketone; Boc, *t*-butoxycarbonyl; AMC, 7-amino-4-methylcoumarin; SC, synthetic complete.

[2] The minus sign before the number indicates residues in the RgpB prodomain counting backward using the first residue of the mature protein as the origin.

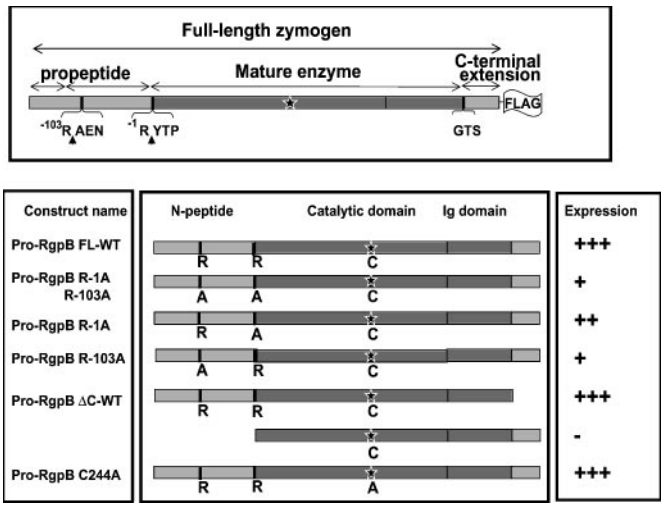

FIG. 1. **RgpB constructs expressed in this study.** The initial transcript of the *rgpB* gene consists of an N-terminal propeptide and a C-terminal extension that flank the mature enzyme. The propeptide cleavage sites are denoted by *arrowheads*, and the catalytic Cys residue is denoted by a *star*. For purification and identification purposes all constructs were tagged at the C terminus with the FLAG sequence (DYKDDDDK). The numbering system used throughout this report places residue 1 at the N terminus of the mature enzyme from the crystal structure (6) with residues upstream of this in the precursor preceded by a negative sign (44). The *lower panel* illustrates the limits and relative expression levels of the various constructs. *R-1A*, $Arg^{-1} \rightarrow$ Ala; *R-103A*, $Arg^{-103} \rightarrow$ Ala.

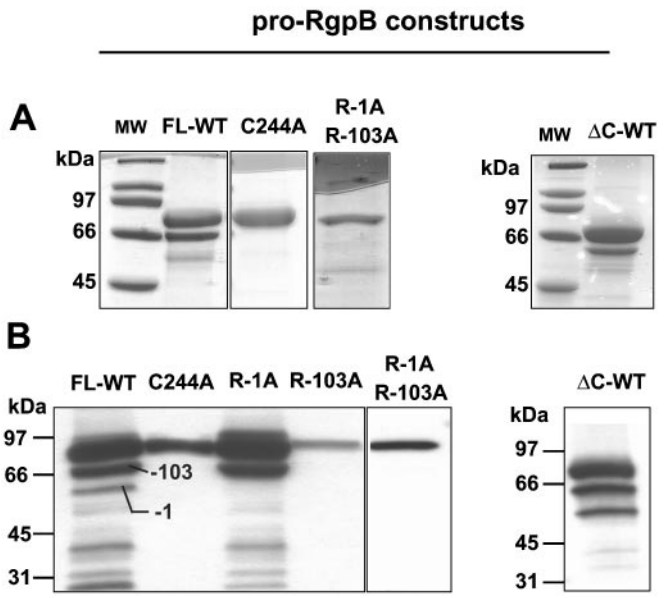

FIG. 2. **Recombinant pro-RgpB expression in yeast.** Recombinant proteins were purified from yeast cells extracts. *A*, Coomassie protein staining of full-length wild-type zymogen (*FL-WT*), C-terminally truncated zymogen (Δ*C-WT*), and mutants described in Fig. 1 based on a full-length background. *B*, Western blot probed with anti-FLAG antiserum of equivalent samples to those in *panel A* and including individual $Arg^{-1} \rightarrow$ Ala (*R-1A*) and $Arg^{-103} \rightarrow$ Ala (*R-103A*) mutants. The intermediate band denoted *−103* represents the truncation at the Arg^{-103}, and *−1* denotes the protein with the same N terminus as the native RgpB as identified by N-terminal sequencing. *MW*, molecular mass markers.

tained for this 82-kDa protein, probably due to a blocked N terminus, but based on the molecular mass in SDS-PAGE and assignment of tryptic peptides by MALDI-TOF (80% sequence coverage), the 82-kDa form corresponds to full-length zymogen with signal peptide attached. This was unexpected since the expression system was designed for secretion with expected signal peptide removal. The observation may explain why the recombinant product was not released from the cells, necessitating the extraction of the protein from the cell pellet. N-terminal sequence analysis revealed that the 68-kDa protein began at Ala^{-104}, while the 56-kDa protein began at Tyr^{1} (see Fig. 1 for a description of the numbering system). Since both of these residues follow Arg in the coding sequence, it is likely that they are generated by a self-processing mechanism. To clarify whether the conversion occurs autocatalytically or whether it is due to a yeast-encoded protease we expressed a catalytic mutant, $Cys^{244} \rightarrow$ Ala. Substitution of Ala for the catalytic Cys completely abolished the smaller products, confirming an autoprocessing mechanism (Fig. 2, *A* and *B*).

To assess the importance of autoprocessing we constructed individual Arg → Ala substitutions at the determined cleavage sites and a double mutant containing both substitutions. Both pro-RgpB $Arg^{-103} \rightarrow$ Ala and the double mutant were found to be unprocessed, whereas an intermediate 68-kDa form corresponding to pro-RgpB processed at Arg^{-103} was detected in the $Arg^{-1} \rightarrow$ Ala mutant (Fig. 2*B*). This indicates that pro-RgpB undergoes sequential two-step processing in which cleavage at Arg^{-103} is required for subsequent processing at Arg^{-1}.

Native mature RgpB obtained from *P. gingivalis* is truncated at the C terminus (6), and consequently we were unable to purify this protein because the C-terminal FLAG tag would be removed. To obtain this derivative we generated the construct pro-RgpB-ΔC-WT that encodes the pro form lacking the 72-residue C-terminal sequence (Fig. 1) Expression and purification of this protein gave a band pattern similar to that of a full-length protein, but as expected, each derivative was slightly smaller (Fig. 2*B*).

Autocatalytic Processing—Recombinant pro-RgpB proteins were incubated at 37 °C for 2 h in assay buffer (see "Materials and Methods") to provide conditions for autolytic processing. After this time the full-length wild-type proenzyme was almost completely converted to the mature form (Fig. 3, *A* and *B*). The fully mature recombinant enzyme appeared to be identical in mass to that of the native enzyme indicating cleavage at the same site. The possibility that a contaminating yeast proteinase was responsible for processing is unlikely because the general serine protease inhibitors phenylmethanesulfonic acid (2.5 mM) and 3,4-dichloroisocoumarin (0.1 mM), the metalloprotease inhibitor 1,10-phenanthroline (1.0 mM), and the aspartic protease inhibitor pepstatin (0.015 mM) had no effect on the self-processing (not shown). In contrast, the gingipain inhibitor leupeptin at 100 μM completely prevented proenzyme autoprocessing. Moreover, the catalytic mutant $Cys^{244} \rightarrow$ Ala incubated in the same conditions remained unprocessed (Fig. 4*A*). The double mutant ($Arg^{-1} \rightarrow$ Ala/$Arg^{-103} \rightarrow$ Ala) subjected to self-processing did not generate any bands recognized by antisera (Fig. 5, *A* and *B*). However, affinity labeling revealed additional bands that resulted from an aberrant cleavage (Fig. 5, *C* and *D*). The aberrant product did not accumulate during self-processing conditions, and we did not observe an increase in activity as determined by cleavage of Boc-QGR-AMC (Fig. 6) or affinity labeling (Fig. 5, *C* and *D*).

We determined the order of processing by using differential Western blot using antisera against the N-terminal and C-terminal (FLAG-tagged) regions. An intermediate with the N-terminal propeptide intact but lacking the C-terminal FLAG was not detected. This reveals that removal of the C-terminal region (residues 434–506) requires prior removal of the N-terminal pro segment. Surprisingly, conversion to the mature form was drastically retarded in the RgpB-ΔC-WT mutant zymogen exposed to autoprocessing conditions, indicating the importance of the C-terminal part of the protein in maturation (Fig. 4*B*). The RgpB-ΔC-WT must be relatively inactive because it does not autopro-

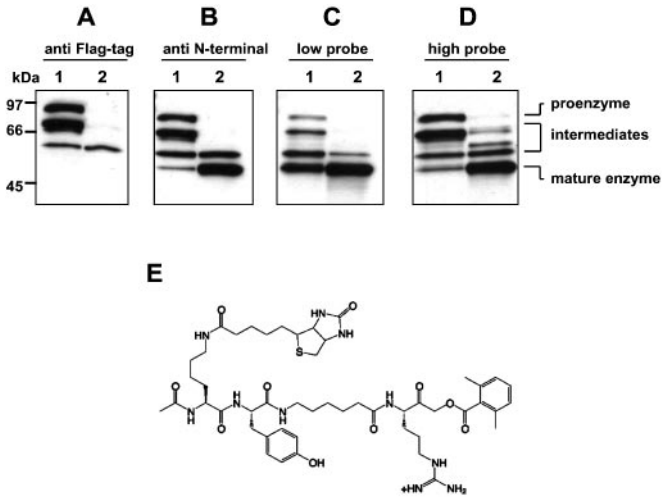

FIG. 3. **Self-processing/activation of pro-RgpB.** Purified recombinant full-length wild-type zymogen was incubated for 2 h at 37 °C, and the reaction products were visualized by Western blot or affinity labeling with BiRK. *A*, proteins probed with anti-FLAG; *B*, proteins probed with specific anti-N-terminal RgpB antiserum; *C*, proteins labeled with a trace concentration of BiRK; *D*, proteins labeled with a saturating amount of BiRK. In each panel *lane 1* represents untreated protein, and *lane 2* represents incubated protein. The different forms of the recombinant enzymes are indicated to the right. *E*, structure of the affinity label BiRK.

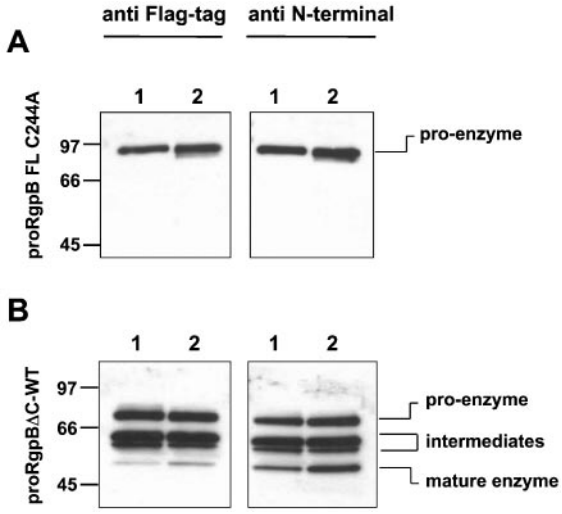

FIG. 4. **Self-processing of pro-RgpB mutants.** Pro-RgpB mutants were subjected to self-processing conditions and analyzed by Western blot using anti-FLAG and anti-N-terminal RgpB antiserum. *A*, full-length catalytic mutant; *B*, ΔC-WT mutant. In each panel *lane 1* represents untreated protein, and *lane 2* represents incubated gingipain.

cess *in vitro*, although conditions in yeast must have been more favorable to produce the initial cleavage events.

Activity Determination of Recombinant Progingipain Forms—Our data demonstrate that mature gingipain can be formed *in vitro* by sequential autoprocessing with the ultimate product lacking both the N-terminal propeptide and the C-terminal extension exactly as seen in the active enzyme isolated from *P. gingivalis*. However, it is not clear whether the final product represents the most active species or what the relative activities of the zymogen and intermediate forms are. To address this issue and to test the relevance of the cleavages we used active site labeling of recombinant zymogen and its derivatives. This technique relies on the inherent ability of a single active site-directed affinity probe to react with kinetics that parallel the enzyme catalytic competence (25). All intermediates in the processing of pro-RgpB have a potential active

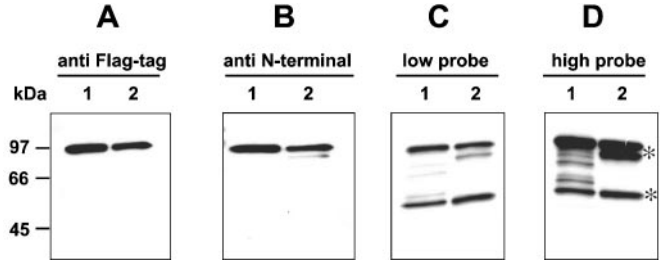

FIG. 5. **Self-processing of pro-RgpB double mutant.** Purified recombinant $Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$ mutant was incubated for 2 h at 37 °C, and the reaction products were visualized by Western blot or affinity labeling with BiRK. *A*, proteins probed with anti-FLAG; *B*, proteins probed with specific anti-N-terminal RgpB antiserum; *C*, proteins labeled with a trace concentration of BiRK; *D*, proteins labeled with a saturating amount of BiRK. In each panel *lane 1* represents untreated protein, and *lane 2* represents incubated protein. Aberrant products are highlighted by an *asterisk*.

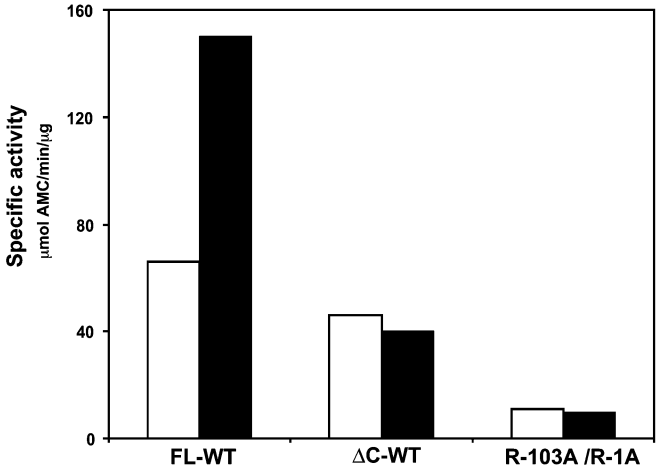

FIG. 6. **Activation of pro-RgpB by self-processing.** The indicated recombinant pro-RgpB forms were subjected to self-processing by incubating for 2 h at 37 °C. Activity was measured by Boc-QGR-AMC substrate hydrolysis. The untreated samples (*white bars*) were stored on ice for the same period of time as the treated ones (*black bars*) and incubated for 10 min at 37 °C in the assay buffer prior to assay. *R-103A/R-1A*, $Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$.

site and an intact catalytic apparatus as demonstrated by BiRK labeling (Fig. 3, *C* and *D*), but this does not mean that they have equivalent activities. If two species of an enzyme with different activities are present in an equimolar amount, they will be equally labeled only if inhibitor is in excess. Conversely, if the inhibitor is in deficit (enzyme is saturating) the relative labeling should reflect the activity of the forms: the most active form will be more heavily labeled. In other words, low BiRK concentrations reflect enzymatic activity, whereas high BiRK concentrations simply reflect total concentration. This useful property enables us to quantitate enzyme activity as a function of BiRK labeling.

The probe we designed, BiRK, was intended for high reactivity with Arg-specific cysteine proteases but was broadly tolerant of extended subsite occupancy. Labeling was performed using a low probe concentration (10 nM) and short time of incubation (5 min) under conditions predetermined (see "Materials and Methods") not to saturate the potential active sites present (Fig. 3*C*). To verify this procedure we also incubated the intermediates with probe under conditions calculated to completely saturate all available active sites (5 μM, 45 min) (Fig. 3*D*). The intensity of the bands labeled at high probe concentration was comparable to the intensity of bands probed with antibodies, reflecting the total protein of each form (Fig. 3, *B* and *D*, compare *lanes 1*). In stark contrast, low probe con-

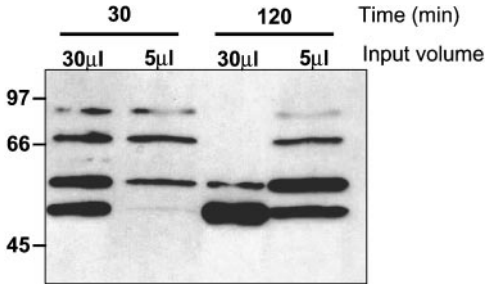

FIG. 7. **Effect of proenzyme concentration on *in vitro* autoprocessing.** Varying amounts of purified recombinant wild-type pro-RgpB were incubated for 30 or 120 min in assay buffer at 37 °C. The samples were treated with 100 μM leupeptin to terminate processing, and an equal amount of protein was analyzed by Western blot using specific anti-N-terminal RgpB antiserum. *Input volume* refers to the volume of pro-RgpB (12 μg/ml) added to the 40-μl total reaction volume.

centrations revealed a completely different pattern with the fully processed enzyme, even though it is present in the lowest amount, displaying highest labeling (Fig. 3, *B* and *C*, compare *lanes 1*). Full-length zymogen, representing the highest concentration among the proteins, showed the lowest labeling under low probe conditions. These findings indicate that the enzyme gains activity with each of the three processing events.

The same samples were also analyzed for enzymatic activity. When incubating with the fluorometric substrate Boc-QGR-AMC a 3-fold increase in substrate cleavage was detected in wild-type digest, and no increase in activity was observed for C-terminally truncated enzyme and double mutant (Fig. 6). Taken together, the data argue that pro-RgpB activation correlates with the enzyme processing and is accomplished by sequential cleavage of N-terminal propeptide followed by trimming of the C terminus.

Processing of Pro-RgpB Is a Bimolecular Process—Zymogen processing could be either intramolecular (by the catalytic site within the precursor) or intermolecular (where a different catalytic site attacks the bonds in an adjacent molecule). These possibilities can be distinguished by determining whether the process is unimolecular or bimolecular (26). The rate of autoprocessing was dependent on precursor concentration (Fig. 7), implying a second order process. Since this is consistent with a bimolecular mechanism, we conclude that processing is predominantly intermolecular, although we cannot rule out some intramolecular component.

Catalytic Properties of Recombinant Gingipain Forms—Probe labeling is consistent with an increase in activity during processing, and we attempted to confirm this by determining the kinetic parameters of full-length wild-type protein converted to the mature form by self-processing and full-length double mutant ($Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$), which represents the zymogen form. Labeling with BiRK revealed the full-length zymogen and an aberrant product of about 45 kDa (Fig. 5, *C* and *D*). The latter probably corresponds to a cleavage C-terminal to Arg^{-1} since it is not recognized by the specific anti-RgpB antiserum raised to the first 35 residues of mature RgpB. The aberrant cleavage appears to be a result of mutating Arg^{-103} and Arg^{-1} and must be taken into account for quantitating activity due to the full-length zymogen. Quantitative image analysis revealed that the full-length zymogen corresponds to 66% of the total RgpB as determined under saturating BiRK conditions (Fig. 5*D*). Thus we estimated the protein concentration to be 66% of that determined by a dye binding assay (see "Materials and Methods"). The full-length form corresponds to 55% of activity as determined under trace BiRK conditions (Fig. 5*C*). Therefore we estimated that 55% of the velocity against Boc-QGR-AMC is due to the full-length zymogen. This is based on the assumption that probe labeling efficiency parallels catalytic efficiency. Therefore these values must be taken into account for an accurate estimate of enzyme concentration and estimate of k_{cat}/K_m.

TABLE I

Kinetic parameters of recombinant mature RgpB and zymogen mutant form ($Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$)

Recombinant pro-RgpB FL-WT was converted to the mature form by self-processing. The frozen zymogen ($Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$) cannot be converted, thus this form represents the full-length zymogen. The parameters describe the respective properties with Boc-QGR-AMC as the substrate.

	K_m	k_{cat}	k_{cat}/K_m
	M	s^{-1}	$M^{-1}\ s^{-1}$
FL-WT	25×10^{-6}	28.0	1.1×10^{6}
FL $Arg^{-1} \rightarrow Ala/Arg^{-103} \rightarrow Ala$	83×10^{-6}	1.2	1.4×10^{4}
Native RgpB	22×10^{-6}	47.0	2.1×10^{6}

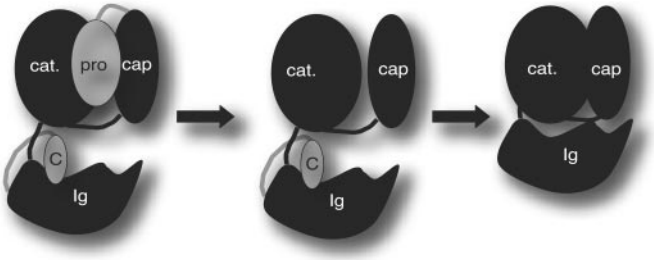

FIG. 8. **Proposed model for RgpB activation.** In the schematic the *gray ovals* represent the N- and C-terminal extensions not found in the fully processed enzyme. The first phase in activation is intermolecular removal of the N-terminal propeptide (*pro*), which occurs in two steps. This is followed by a second phase during which the C-terminal extension (*C*) is removed. According to the crystal structure (6) the catalytic domain is composed of a core with catalytic sites (*cat.*) and a structure that stabilizes the conformation of the active form (*cap*). The active domain sits on top of an Ig domain that has intimate contacts with both the catalytic site and the active form. We propose that autolytic processing removes constraints to promote active site formation through conformational coupling of the three main structural units.

Based on the estimated k_{cat}/K_m values (Table I) we observed an approximately 80-fold increase in catalytic efficiency during conversion of zymogen to the mature enzyme. Significantly the kinetic parameters of fully processed wild-type pro-RgpB are close to that of the native enzyme isolated from *P. gingivalis*.

DISCUSSION

The initial translation product deduced from the *rgpB* gene is a precursor protein with a 24-residue signal peptide, 229-residue N-terminal pro region, and a 72-residue C-terminal region that are not found in the mature native enzyme. Although previously postulated (27, 28), the presence of an RgpB precursor has not previously been demonstrated in *P. gingivalis*, and production of recombinant active gingipains in *E. coli* with activity comparable to the natural enzymes has been unsuccessful (29, 30). The latent activity of almost all proteases is usually restrained by embedding them in a precursor that must be processed to generate the active form, thus protecting biosynthetic machinery and allowing for activation control (12). This study aimed to produce the precursor form of RgpB with a view to understanding the reason for the precursor: was it to secure the zymogen or was it required for folding of the catalytic form? Because RgpB is naturally a secreted product we reasoned that a heterologous secretion/expression system would most faithfully simulate the natural folding and processing events, hence our choice of a yeast expression system.

We provide direct evidence that recombinant pro-RgpB can be autocatalytically processed to generate a mature form of the enzyme equivalent to that isolated from *P. gingivalis* culture

supernatants. Maturation of pro-RgpB occurs through the sequential appearance of intermediates leading to the ultimate product (Fig. 8). The first cleavage at Arg^{-103} is essential for further processing since mutation at this site abolishes generation of authentic downstream intermediates. The second processing step removes the remainder of the N-terminal propeptide, and the final step yields the mature enzyme by removal of a C-terminal peptide. Unfortunately we were not able to identify the cleavage site at the C terminus because of the difficulty in obtaining the peptide. Nevertheless, conversion of pro-RgpB into the mature enzyme can be attributed to the action of the enzyme itself. The conversion can take place during expression and/or purification but also *in vitro* with the purified components. Naturally other bacterial factors may influence processing *in vivo*, but our data demonstrate that an inherent processing pathway exists. This implies that the precursor must have at least a small degree of proteolytic activity, hence our attempts to measure the relative activity of the various intermediates.

It proved impossible for us to obtain full-length wild-type pro-RgpB because of processing in yeast, and therefore we resorted to obtaining material mutated at the −1 and −103 processing sites, which we define as the "frozen zymogen." This full-length material was somewhat active with an apparent k_{cat}/K_m value 80-fold lower than that of the mature protein. It is not unusual for protease zymogens to have proteolytic activity, but normally this activity is many orders of magnitude lower than that of the active enzyme (12). There are a few examples of protease precursors with substantial activity compared with its processed product, including tissue-type plasminogen activator (31), coagulation factor VII (32), and caspase-9 (13, 33). However, in each of these cases intrinsic activity of the processed form is low and requires enhancement by cofactors. This appears not to be the case with pro-RgpB since the frozen zymogen form has considerable catalytic activity with an estimated k_{cat}/K_m of 1.4×10^4 $M^{-1}s^{-1}$, higher than some fully processed proteases on equivalent synthetic substrates. Interestingly the activity of the enzyme seemed to increase with each processing step as demonstrated by an enhanced interaction with the affinity probe (Fig. 3*C*). Although it is difficult to quantitate the probe data, the increase in probe labeling is consistent with an 80-fold increase in catalysis.

Conceivably the 80-fold lower activity of the zymogen may reflect the need to restrain activity during protein synthesis and delivery in *P. gingivalis*, but there is another plausible explanation for the existence of the precursor. We were not able to obtain protein from a construct encoding the mature protein. This indicates that the propeptide may serve an intramolecular chaperone function analogous to the propeptide of subtilisin-like proteases. Extensive work on subtilisin demonstrates that the 77-residue subtilisin propeptide is not required for enzymatic activity and is removed intramolecularly by autoprocessing upon the completion of the protein folding (for a review, see Ref. 34). This intramolecular chaperone function for protease N-terminal propeptides is not restricted to subtilisins but also encompasses other families (35, 36). Significantly the intramolecular chaperone propeptides frequently also act as potent inhibitors of the enzymes, decreasing enzyme activity several orders of magnitude (37). Since the RgpB propeptide seems to decrease activity 80-fold it may function as an inhibitor, but it is more plausible that its role is to serve in the folding of the catalytic domain of RgpB. In support of this is the inability of our group and several other groups to express the catalytic domain (residues 1–506) in an active or soluble form (29, 30).

Perhaps the greatest challenge is to relate the structure/function of the RgpB propeptide to the activation mechanism of its cousins, separase, the legumains, and caspases. On the basis of conservation of mechanism one would expect the fundamental activation mechanisms to be essentially identical. Both yeast and human separase undergo autolytic processing, yet this does not seem to be responsible for activation of the enzymes (38, 39). Separase may require a separate chaperone/inhibitor known as securin to allow correct folding and generation of a latent active site. In contrast, human legumain has been proposed to undergo an activating C-terminal cleavage (40) similar to the final activating cleavage demonstrated above for RgpB. On the other hand, recent x-ray structures of zymogen forms of caspase-7 (14, 15) and caspase-9 (13) demonstrate a fundamentally different mechanism. In both cases the zymogens show at least 3 orders of magnitude less activity than that of the fully active forms (33, 41), and this is caused by a dislocation of two loops that contain the major activity and specificity determinants of each protease. Although the driving forces are distinct (cofactor binding for caspase-9 and proteolysis for caspase-7) activation results from reordering of these two loops. Like RgpB, the precursors of caspases retain a reduced ability to incorporate affinity probes (25, 42, 43), which is not typical of most protease zymogens, and we take this as preliminary evidence that pro-RgpB is activated in a similar manner (Fig. 8). This would mean that the propeptide may not block access of substrate to a fully formed active site as seen in many protease families. More likely the propeptide docks with the region between the catalytic domain and cap domain thereby restricting ordering of the activation loop (Fig. 8). Proteolytic removal would then allow activation by a caspase-like mechanism, but further speculation should await structural evidence.

Acknowledgment—We thank Scott Snipas for outstanding technical support.

REFERENCES

1. Nakayama, K., Kadowaki, T., Okamoto, K., and Yamamoto, K. (1995) *J. Biol. Chem.* **270,** 23619–23626
2. Curtis, M. A., Aduse-Opoku, J., and Rangarajan, M. (2001) *Crit. Rev. Oral Biol. Med.* **12,** 192–216
3. Genco, C. A., Potempa, J., Mikolajczyk-Pawlinska, J., and Travis, J. (1999) *Clin. Infect. Dis.* **28,** 456–465
4. Veith, P. D., Talbo, G. H., Slakeski, N., Dashper, S. G., Moore, C., Paolini, R. A., and Reynolds, E. C. (2002) *Biochem. J.* **363,** 105–115
5. Kadowaki, T., Nakayama, K., Yoshimura, F., Okamoto, K., Abe, N., and Yamamoto, K. (1998) *J. Biol. Chem.* **273,** 29072–29076
6. Eichinger, A., Beisel, H. G., Jacob, U., Huber, R., Medrano, F. J., Banbula, A., Potempa, J., Travis, J., and Bode, W. (1999) *EMBO J.* **18,** 5453–5462
7. Thornberry, N. A., and Lazebnik, Y. (1998) *Science* **281,** 1312–1316
8. Chen, J. M., Rawlings, N. D., Stevens, R. A., and Barrett, A. J. (1998) *FEBS Lett.* **441,** 361–365
9. Uhlmann, F., Wernic, D., Poupart, M. A., Koonin, E. V., and Nasmyth, M. (2000) *Cell* **103,** 375–386
10. Barrett, A. J., and Rawlings, N. D. (2001) *Biol. Chem.* **382,** 727–733
11. Aravind, L., and Koonin, E. V. (2002) *Proteins* **46,** 355–367
12. Neurath, H. (1989) *Trends Biochem. Sci.* **14,** 268–271
13. Renatus, M., Stennicke, H. R., Scott, F. L., Liddington, R. C., and Salvesen, G. S. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 14250–14255
14. Chai, J., Wu, Q., Shiozaki, E., Srinivasula, S. M., Alnemri, E. S., and Shi, Y. (2001) *Cell* **107,** 399–407
15. Riedl, S. J., Fuentes-Prior, P., Renatus, M., Kairies, N., Krapp, R., Huber, R., Salvesen, G. S., and Bode, W. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 14790–14795
16. Johnston, J. R. (1994) *Molecular Genetics of Yeast: A Practical Approach*, Vol. 141, pp. 123–134, IRL Press at Oxford University Press, Oxford
17. Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122,** 19–27
18. Nielsen, T. L., Holmberg, S., and Petersen, J. G. (1990) *Appl. Microbiol. Biotechnol.* **33,** 307–312
19. Gietz, D., St. Jean, A., Woods, R. A., and Schiestl, R. H. (1992) *Nucleic Acids Res.* **20,** 1425
20. Salvesen, G., and Nagase, H. (1989) in *Proteolytic Enzymes: A Practical Approach* (Beynon, R. J., and Bond, J. S., eds) pp. 83–104, IRL Press, Oxford
21. Potempa, J., Pike, R., and Travis, J. (1995) *Infect. Immun.* **63,** 1176–1182
22. Lee, A., Huang, L., and Ellman, J. A. (1999) *J. Am. Chem. Soc.* **121,** 9907–9914
23. Robinson, A. J., Pauls, H. W., Coles, P. J., Smith, R. A., and Krantz, A. (1992) *Bioorg. Chem.* **20,** 42–54
24. Potempa, J., Mikolajczyk-Pawlinska, J., Brassell, D., Nelson, D., Thogersen, I. B., Enghild, J. J., and Travis, J. (1998) *J. Biol. Chem.* **273,** 21648–21657
25. Yamin, T.-T., Ayala, J. M., and Miller, D. K. (1996) *J. Biol. Chem.* **271,** 13273–13282

26. Nagase, H., Enghild, J., Suzuki, K., and Salvesen, G. (1990) *Biochemistry* **29,** 5783–5789
27. Mikolajczyk-Pawlinska, J., Kordula, T., Pavloff, N., Pemberton, P. A., Chen, W. C., Travis, J., and Potempa, J. (1998) *Biol. Chem.* **379,** 205–211
28. Nakayama, K. (1997) *Microbiol. Immunol.* **41,** 185–196
29. Aduse-Opoku, J., Muir, J., Slaney, J. M., Rangarajan, M., and Curtis, M. A. (1995) *Infect. Immun.* **63,** 4744–4754
30. Margetts, M. B., Barr, I. G., and Webb, E. A. (2000) *Protein Expr. Purif.* **18,** 262–268
31. Madison, E. L., Kobe, A., Gething, M., Sambrook, J. E., and Goldsmith, E. J. (1993) *Science* 419–421
32. Ruf, W., and Dickinson, C. D. (1998) *Trends Cardiovasc. Med.* **8,** 350–356
33. Stennicke, H. R., Deveraux, Q. L., Humke, E. W., Reed, J. C., Dixit, V. M., and Salvesen, G. S. (1999) *J. Biol. Chem.* **274,** 8359–8362
34. Inouye, M. (1991) *Enzyme* **45,** 314–321
35. Baker, D., Shiau, A. K., and Agard, D. A. (1993) *Curr. Opin. Cell Biol.* **5,** 966–970
36. Winther, J. R., and Sorensen, P. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9330–9334
37. Peters, R. J., Shiau, A. K., Sohl, J. L., Anderson, D. E., Tang, G., Silen, J. L., and Agard, D. A. (1998) *Biochemistry* **37,** 12058–12067
38. Hornig, N. C., Knowles, P. P., McDonald, N. Q., and Uhlmann, F. (2002) *Curr. Biol.* **12,** 973–982
39. Waizenegger, I., Gimenez-Abian, J., Wernic, D., and Peters, J. (2002) *Curr. Biol.* **12,** 1368
40. Chen, J. M., Fortunato, M., and Barrett, A. J. (2000) *Biochem. J.* **352,** 327–334
41. Zhou, Q., and Salvesen, G. S. (1997) *Biochem. J.* **324,** 361–364
42. Muzio, M., Stockwell, B. R., Stennicke, H. R., Salvesen, G. S., and Dixit, V. M. (1998) *J. Biol. Chem.* **273,** 2926–2930
43. Roy, S., Bayly, C. I., Gareau, Y., Houtzager, V. M., Kargman, S., Keen, S. L., Rowland, K., Seiden, I. M., Thornberry, N. A., and Nicholson, D. W. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 6132–6137
44. Pavloff, N., Potempa, J., Pike, R. N., Prochazka, V., Kiefer, M. C., Travis, J., and Barr, P. J. (1995) *J. Biol. Chem.* **270,** 1007–1010