Cloning of cDNA for Proteinase 3: A Serine Protease, Antibiotic, and Autoantigen from Human Neutrophils

By David Campanelli, Maxine Melchior, Yiping Fu, Munehiro Nakata, Howard Shuman,* Carl Nathan, and Joelle E. Gabay

From the Beatrice and Samuel A. Seaver Laboratory, Division of Hematology-Oncology, Department of Medicine, Cornell University Medical College, New York, New York 10021; and the *Department of Microbiology, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Summary

Closely similar but nonidentical NH2-terminal amino acid sequences have been reported for a protein or proteins in human neutrophils whose bioactivities is/are diverse (as a serine protease, antibiotic, and Wegener's granulomatosis autoantigen) but that share(s) several features: localization in the azurophil granules, a molecular mass of \sim 29 kD, reactivity with diisopropylfluorophosphate, and the ability to degrade elastin. We previously purified one such entity, termed p29b. Using ^a monospecific antibody, we have cloned from human bone marrow ^a cDNA encoding the complete p29b protein in its mature form, along with pre- and pro-sequences . The predicted amino acid sequence agrees closely with the NH₂-terminal sequence obtained previously from purified p29b, as well as with sequences newly obtained from CNBr fragments . The primary structure is highly homologous to elastase, cathepsin G, T cell granzymes, and other serine proteases, and shares both the catalytic triad and substrate binding pocket of elastase. Hybridization of the full-length cDNA with restriction enzyme digests of human genomic DNA revealed only one fragment. This suggests that the closely related species described previously are the same, and can be subsumed by the term used for the first-described activity, proteinase 3. Proteinase 3 is more abundant in neutrophils than elastase and has a similar proteolytic profile and specific activity. Thus, proteinase 3 may share the role previously attributed to neutrophil elastase in tissue damage, and has the potential to function as an antimicrobial agent.

The azurophil granules of human neutrophils (PMN) contain two ~29-kD proteins, cathepsin G and elastase, that are active both as neutral serine proteases and as antibiotics he azurophil granules of human neutrophils (PMN) contain two \sim 29-kD proteins, cathepsin G and elastase, that (1-4) . These molecules are thought to help PMN form abscesses and damage other inflamed tissues (5), and may participate in killing micoorganisms . Recently, we purified another azurophil granule protein of \sim 29 kD, termed p29b, whose abundance is intermediate between that of cathepsin G and elastase, and whose NH2-terminal amino acid sequence is highly homologous to theirs (3, 4). Like the latter proteins, p29b is bifunctional, displaying broad-spectrum antimicrobial as well as elastase-like enzymatic activity (3, 4) . The proteolytic profile of p29b resembled that of a diisopropylfluorophosphate (DFP)¹-inhibitable protease of \sim 29 kD purified by Kao et al. (6) from azurophil granules and shown to cause emphysema after intratracheal instillation in ham-

sters. The latter enzyme was termed proteinase 3 (PR-3), because it shared the reported properties of an organic ester hydrolase that Baggiolini et al. (7) had so designated after separating it from cathepsin G and elastase on nondenaturing SDS-PAGE. The NH2-terminal sequence of PR-3 (8) matched that of p29b in only 9 of 14 residues, leaving their relationship unclear.

Interest in p29b and/or PR-3 was heightened by the discovery that autoantibodies characteristic of Wegener's granulomatosis activate PMN (9), stain PMN granules, and can be used to affinity purify a 29-kD, DFP-binding protease, termed p29, ACPA, or C-ANCA (10-13), whose $NH₂$ -terminal sequence, as first reported (11, 12), was related but not identical to those of $p29b(3)$ or PR-3 (8) . Anti-PR-3 mAbs could block the staining of PMN by sera from patients with Wegener's granulomatosis (13) . Meanwhile, ^a cDNA cloned from differentiating HL60 promyelocytic leukemia cells was predicted to encode a serine protease, termed myeloblastin (14), whose six NH2-terminal residues matched residues 15-20 in p29b. Addition of ^a myeloblastin antisense oligode-

 $¹$ Abbreviations used in this paper: DFP, diisopropylfluorophosphate; IPTG,</sup> isopropyl- β -D-thiogalactoside; PR-3, proteinase 3.

¹⁷⁰⁹ J. Exp. Med. ® The Rockefeller University Press - 0022-1007/90/12/1709/07 \$2 .00 Volume 172 December 1990 1709-1715

oxynucleotide construct to H1r60 cells induced their myeloid differentiation (14). Very recently, a preliminary report by Jenne et al. (15) has revised the NH₂-terminal sequence for the Wegener's autoantigen (12), such that it now matches all 20 NH_2 -terminal residues reported for p29b (3), except for Gln¹⁹ in place of Glu¹⁹. Based on PCR analysis, Jenne et al. (15) also proposed a revision of the 5' region of the myeloblastin cDNA, according to which the deduced NH2 terminal 20 residues of myeloblastin would likewise match those of the Wegener's autoantigen and p29b.

We have used ^a monospecific anti-p29b IgG to clone ^a cDNA that encodes the entire mature protein. Below, we provide the sequence of this clone, compare its deduced amino acid sequence with the observed sequence of proteolytic fragments of p29b, note the extensive homology of this molecule to cathepsin G, elastase, and the granzymes of T cells, and demonstrate that only one human genomic DNA fragment appears to hybridize with the p29b cDNA. Thus, it is likely that the cDNA presented here encodes ^a single PMN azurophil granule protein that has the potential for elastinolysis, microbial killing, and the regulation of myeloid differentiation, and serves as an autoantigen.

Materials and Methods

Amino Acid Sequence of CNBr Fragments. 50 μ g of purified p29b (3, 4) was incubated with ^a 500-fold excess (with respect to methionine content) of CNBr in 0.1 N HCl for ⁴⁸ h at room temperature in the dark. The samples were subjected to SDS-PAGE. Resolved peptides were transferred to polyvinylidine difluoride membranes (Millipore Continental Water Systems, Bedford, MA) (16) and sequenced directly by automated Edman degradation on ^a gasphase sequenator (470; Applied Biosystems, Inc., Foster City, CA).

 $cDNA$ Cloning. A λ gt11 bone human marrow $cDNA$ library (Clontech, Palo Alto, CA) was plated on Escherichia coli Y1090 (17) . Nitrocellulose filter lifts saturated with isopropyl- β -D-thiogalactoside (IPTG; Sigma Chemical Co., St . Louis, MO) were dried, blocked with 5% nonfat dry milk and 0.05% Tween ²⁰ in PBS, rinsed, probed with a 1:500 dilution of monospecific rabbit anti-p29b IgG (4) that had been preabsorbed with E. coli Y1090, rinsed, and treated with goat anti-rabbit IgG conjugated with alkaline phosphatase (Boehringer Mannheim Biochemicals, Indianapolis, IN) . Plaques positive by a colorimetric phosphatase reaction were purified to homogeneity, as monitored by rescreening with the antibodies. Plate cultures were lysed in chloroform, clarified by centrifugation, and rid of bacterial DNA with DNAse (Pharmacia Fine Chemicals, Piscataway, NJ). Phage were collected by ultracentrifugation (89,500 g, 2 h, 4°C), purified on spin columns of DEAE cellulose (5-Prime 3-Prime, Westchester, PA) (18), and the DNA was isolated according to the manufacturer's instructions. The cDNA insert was recovered by restriction with EcoR1 (Boehringer Mannheim Biochemicals) and electrophoresis in low melt/low gelling agarose (Bio-Rad Laboratories, Richmond, CA) as visualized with long-wavelength UV light. The appropriate gel slice was melted and extracted with phenol and ether, and the DNA was precipitated with ethanol. pBluescript II SK^+ (Stratagene, La Jolla, CA) was digested with EcoRl and dephosphorylated (calf intestinal alkaline phosphatase; Boehringer Mannheim Biochemicals) for insertion of the p29b cDNA with T4 DNA ligase (Stratagene). The resultant plasmid was used to transform competent E. coli XL1 Blue (Stratagene) (19) . Recombinants were screened by blue/white color

 $(\beta$ -galactosidase) selection on luria broth agar with ampicillin, 20 mM IPTG, and 80 μ g/ml 5-bromo-4-chloro-3-indolyl- β -Dgalactoside . After transformation of DH5 cells (20), the recombinant plasmids were amplified and purified (21). cDNA was sequenced by the dideoxynucleotide chain termination method using double-stranded plasmid DNA as ^a template (22) with ^a kit from U.S. Biochemical Corp. (Cleveland, OH) and deoxyadenosine 5'-[³⁵S]triphosphate (New England Nuclear, Boston, MA). After initial use of universal primers, sequencing was continued independently on both strands with sequentially constructed 20-25-mer oligonucleotides so as to obtain sequences in regions overlapping by 20-40bp. Results were analyzed by DNASIS and PROSIS software (Hitachi America, Ltd., San Bruno, CA). Alignment to other proteases was scored according to Lipman and Pearson (23) with the program FASTR Structural predictions were obtained as described (24).

Southern Hybridization. Human placental genomic DNA (10 μ g; Clontech) was digested to completion with EcoRI, HindIII, BamHl, Pstl, or BglII (Boehringer Mannheim Biochemicals) . The digests were transferred onto ^a nylon membrane (Schleicher & Schuell, Inc., Keene, NH) and hybridized with p29b cDNA [32P] labeled by random priming (25). Prehybridization (2 h in $5 \times$ Denhardt's solution) and hybridization (18 h in 10% dextran sulfate) were performed at 42'C (high stringency) or 32°C (low stringency) in 50% (vol/vol) formamide in $6 \times$ SSC/0.5% SDS with 100 μ g/ml sonicated salmon sperm DNA. Filters were washed twice for ¹⁵ min at room temperature with $6 \times$ SSC/0.5% SDS, twice at 37°C and once at 65° C with $1 \times$ SSC/0.5% SDS (high stringency), or twice at room temperature with $2 \times$ SSC/0.1% SDS and twice at 37° C with $0.25 \times$ SSC/0.1% SDS (low stringency).

Results

Cloning. A human bone marrow cDNA expression library in λ gt11 was screened with monospecific anti-p29b IgG (4). One of the positive clones, containing an insert of 1,014 bp, was purified, subcloned in pBluescript SK', and sequenced (Fig. 1). The open reading frame of 762 bp encodes a 254-amino acid polypeptide, including a 26-residue NH2 terminal peptide and a 228-residue mature protein. The latter corresponds to p29b, based on the following evidence. First, the deduced amino acid sequence for residues 1-20 (numbering for the mature protein) matches exactly the sequence obtained from the $NH₂$ terminus of the purified protein $(Glu¹⁹$ was originally observed [3], but $Gln¹⁹$ was detected in the CNBr peptide sequenced in the present study and conforms to the cDNA). Second, two peptides from CNBr cleavage of purified p29b provided the sequence of 55 residues, and ^a [3H]DFP-binding tryptic fragment of p29b reported by Wilde et al. (26) provided another 22 residues; of these 77 residues, 70 were identical to the deduced sequence. Third, the deduced amino acid composition corresponded closely to that determined for purified p29b (not shown).

Features of the Predicted Structure. The 5' end of the cDNA encodes a hydrophobic stretch (residues -22 to -9) (Fig. 2) typical of ^a signal sequence followed by a consensus signal peptidase cleavage site (Ala-X-Ala) (27) at -5 to -3 (Fig. 1). A 24-residue pre-signal peptide (Arg⁻²⁶ to Ala⁻³) 50% homologous to the pre-sequence in the elastase cDNA (28) is followed by a two-residue pro-peptide $(Ala^{-2}-Glu^{-1})$, as

1 -26	CGGCACCGGCCCCCCAGCCCTGCCCTGGCGTCCGTGCTGCTGGCCTTGCTGCTGAGCGGT R H R P P S P A L A S V L L A L L L S G	60 -7
61 -6	SCTSCCCGASCTSCSGAGATCGTGGGCGGGCACGAGGCGCAGCCACACTCCCGGCCCTAC A A R A A E I V G G H E A O P H S R P Y	120 14
		180
121 15	ATGGCCTCCCTGCAGATGCGGGGGAACCCGGGCAGCCACTTCTGCGGAGGCACCTTGATC NASLONRGNPGSHFCOGGTLI	34
181	CACCCCASCTTCGTGCTGACGGCCGCACTGCCTGCGGGACATACCCCAGCGCCTGGTG	240
35	HPSFYLTAAHOLRDIPQRLV	54
241		300
55	N Y Y L G A H N Y R T O E P T O O H F S	74
301	GTGGCTCAGGTGTTTCTGAACAACTACGACGCGGASAACAAACTGAACGACGTTCTCCTC	360
75	V A Q V F L N N Y D A E N K L N IDI V L L	94
361	ATCCAGCTGAGCAGCCCAGCCAACCTCAGTGCCTCCGTCGCCACAGTCCAGCTGCCACAG	420
95	O L S S P A N L S A S V A T V O L P O r	114
421	CAGGACCAGCCAGTGCCCCACGGCACCCAGTGCCTGGCCATGGGCTGGGCCGCGTGGGT	480
115	O D O P V P H G T O COL A M <u>G W G R V G</u>	134
481	<i>GCCCACGACCCCCCAGCCCAGGTCCTGCAGGAGCTCAATGTCACCGTGGTCACCTTCTTC</i>	540
135	A H D P P A O V L O E L N V T V V T F F	154
541	TGCCGGCCACATAACATTTGCACTTTCGTCCCTCGCCGCAAGGCCGGCATCTGCTTCGGA	600
155	CRPHNICTFVPRRKA_G_ICOFG	174
601 175	GACTCAGGTGGCCCCCTGATCTGTGATGGCATCATCCAAGGAATAGACTCCTTCGTGATC DISIGGPLICODGIIOGIDSFVI	660 194
661 195	T6666AT6T6CCACCC6CCTTTTCCCT6ACTTCTTCAC6C666TA6CCCTCTAC6T66AC W G CO A T R L F P D F F T R V A L Y Y D	720 214
721 215	TGGATCCGTTCCACGCTGCGCCGTGTGGAGGCCAAGGGCCCCTGAACCGCCCCTCCCACA WIRSTLRRVEAKGP +	780 228
781	GCGCTGGCCGGGACCCCGAGCCTGGCTCCAAACCCTCGAGGCGGATCTTTGGACAGAAGC	
		840
841		900
901	GGCTCCGGAAGACAGGCCGGCCCTGCACCTCACCCCACCGTGACCTCAATAAACGTTGAA	960
961	ACTCCAAAAAAAAAA	975

Figure 1. Nucleotide and deduced amino acid sequences of p29b (PR-3). The nucleotide sequence is numbered from the first base of the cDNA insert after the EcoRl linker site . The deduced amino acid sequence (singleletter code) is numbered from the NHz-terminal residue of the mature protein $(+1)$ (3). Regions for which amino acid sequence have been determined directly are underlined. These include residues 1-20, as determined earlier from the intact protein (3), 16-44 and 129-154, as determined in this study from CNBr fragments, and 169-192, as determined by Wilde et al. (26) from a tryptic peptide; the latter contained gaps corresponding to positions 182 and 189 of p29b. The following discrepancies are discussed in the text (giving the position, the deduced amino acid, and the assignment from amino acid sequencing): 29, C for I; 31, G for P; 35, H for L/V; 37, S for A/V; 44, H for I; 130, W for P; 147, N for R. Components of the catalytic site triad are boxed. Cysteines are circled. Potential N-linked glycosylation sites are marked by dots. The stop codon is marked with a large asterisk. The polyadenylation signal is underlined twice. Amino acid assignments that differ from those deduced for myeloblastin (14, 15) are marked by small asterisks; the Arg that forms the penultimate residue of the reported myeloblastin sequence (14) was not detected in p29b. Amino acids -8 to $+14$ are missing from the myeloblastin sequence of Bories et al. (14) but are present in the revised sequence proposed by Jenne et al. (15); amino acids -26 to -9 are not included in either of the reported sequences. These sequence data have been submitted to the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the accession number X55668 for human mRNA HUMPR-3.

in other leukocyte lysosomal proteases (28-30) . However, the cDNA insert lacks an ATG translation initiation codon. The ³' end of the cDNA contains ^a TGA termination codon, as well as a polyadenylation signal (AATAAA) (31) and a poly(A) tail, suggesting that the clone encompasses the full-length sequence for the mature protein.

Mature p29b is predicted to consist of 228 amino acids with a M_r of 25,000 for the polypeptide backbone. Two

Figure 2. Hydropathy plot of p29b (PR-3) according to the algorithm of Kyte and Doolittle (42) with a window of $n = 6$. The lower panel indicates the distribution of basic $(+)$ and acidic $(-)$ residues.

potential N-linked glycosylation sites (Asn-X-Ser/Thr) are present, and may account for the apparent isoforms (3, 4). Eight Cys residues (five clustered in the COOH-terminal third of the molecule) suggest that there may be four intramolecular disulfide bonds, as in elastase (32). His⁴⁴, Asp⁹¹, and Ser¹⁷⁶ correspond to the active site triad typical of serine proteases (Fig. 1) .

After a hydrophobic region $(-22 \text{ to } -9)$ corresponding to the presumed signal sequence, the NH2-terminal half of p29b contains six regions averaging 17 residues each, in which there is a marked, periodic alternation of hydrophilicity and hydrophobicity (Fig. 2). The mature protein is predicted by the Chou-Fasman algorithm (33) to contain three short stretches of α -helix totalling 8.3%, and 10 β -sheets totalling 40%. Remarkably, the β -sheets are all predominantly hydrophobic. With 27 basic and 15 acidic amino acids in the mature protein (Fig. 2, bottom), p29b has an overall calculated pI of 7.9 . Regions of the molecule predicted to be surface exposed (34) correspond to residues 7-12, 63-72, 83-88, 113-118, 165-168; these 32 residues include four acidic and eight basic amino acids.

Fig. 3 compares the p29b sequence with that of other serine proteases, aligned to maximize homology. p29b exhibits the highest homology with elastase (54%). There is also considerable homology with the T cell enzymes human lymphocyte protease (36%) and granzyme B (33%), and with cathepsin G (35%), rat mast cell protease II (34%), and chymotrypsin (30%). Residues 9-16 of p29b comprise the PHSRPYMA sequence strictly conserved among granzymes (35) . All eight Cys in p29b align with the Cys residues of elastase; five align with corresponding residues in the other proteases listed. One of the potential N-linked glycosylation sites in $p29b$ (Asn¹⁷³) aligns with a corresponding site in elastase; both enzymes have two such sites. The members of the catalytic triad (His, Asp, Ser) are located at homologous positions in all the aligned sequences, flanked by wellconserved peptides. The residues at -6 , 15-17, and 28 relative to the active site serine are thought to determine the

PR-3 Elastase : Cat G Gra B RMCP II HLP Chymo.		IIV G G <u>H E AIQIP H SR PY</u> NA SL QINR GIN PIGIS IH F C G <u>G T LIIHIP</u> ISIF V <u>L T</u> A A H CIL R D I P QIRILIV IN V V L G HECGATLIAPNFVHSAAHCVANVNVRAVRVLG IVGGRRARPHAMPEMMSLOLRG G ITTGGESTRPHSRPYMATIOTOSPAGO SRCGGELIVREDFVLTAAHCHGS IIIIGGWEISIPHSRPYNAMLDIVTEKGLRVICGGFLISROFVLTAAHCKGR IIIKGHEAKPHSRPYNAWLMIWDQKSLKR CGGFLIQDDEVLTAAHCWGS DKTGF HIFCGGSLINENTWATAAHCGVT I VINGELE AV PIGSMPNOVISLO	'N'INVITIL G 1 IN VITIL GI EITNILG SINVILG TSDVVVAGI
PR-3 Elastase : Cat G Gra B RMCP II HLP Chymo.		OMFLIANT ON EPACE NOW LEITOLS SPANE SASTA TIZOLPIOOD IFFISIVIA RTOEPTOO JAHNVI SRRIEP TIRIO VIFLAMOR "IFFEINGY OP VINLE NOI VILLO LETE SAT TRANVO VALO L PALOG AHNE AHNI QRRENTOQ INTYARRATRHPONRORTIONOI NOLO LERRYRRRRNN MODEL PRAO NSKYNSLTINNOITULKLSTAASFSOTUS AUCLPSAS : EFDQGSSSEKIQKLKIAKVFK	
PR-3 Cat G Gra B RMCP II HLP Chymo.		LNVTVVTFF CRPH ONLOE : OPMPHGTOCLANGWGRVGAHDPPA Swige LI DTLREVOLI NTLQEVELI ILNVTVVTISL IC RIR S Elastase : RRLGNGMOCLAMGWGLLGRNRGIA RVQRDRQCLRIFGSYDPRR : EGLRPGTLLCTVAGWGRVTS MRRGT : VNWKPGDVCCYVAGWGRMAPM <u>GK</u> YS TVOKDRECESYFKNRYNKT NOLCA YTLREVELL RIMDEKACVDYRYYEYKF : DFTHPIGAMICHANGWGKTGVRDPTS <mark>TV</mark> IQEDRKICESDLRHYYDSTI E∐CÍV A OMK PIGIO TICIS VAIGN GIO TAP LEKHS HTLO EIVKH PLLSNTNCKKYWG TKIKDAM <mark>IC</mark> A DDFAAGTTCVTTGWGLTRYTNANTPDRLQQASLI $\ddot{}$	itcT $\frac{1}{2}$ ⊽ст .o∏dv 'o'Viciv
PR-3 Elastase : Cat G Gra B RNCP II HLP Chymo.	$\ddot{}$	۰ ITO GILDISFVIN GCATRILIFIP DIFFTRVALLYVIDMI оQ GDSGGPLIC FVPRRKAGICT NG LITHG TAISF V ROGCALS GLYP DIAIFA PIV AN FLYINH I LIVIR GRIGA GMC FI IGD SGSPLIVIC VAHG IMSYEKS SGVP GDRIRERKAA FKGDSGGPLLC ΝÑ VAAGINSYGYKDGSP iririgo s g g p Livici KΚ GDIPIK TIKIRAS ۸ū VAHGIWSYGHPDAKP IFINGD SGGPLILICI GSIPITTERAA V ALG I VISIY GRNNGHP GDIPLE INCK TS FIKGDSGGPLIVIC NΚ SSCI MGDSGGPLWCKKNGAWTLVKGLWSMG SSTOSLIJST GASGV	IPIE VIFTR VISSFL PM 1 IPIR AIF TIRMS S F L SIN I IPIA IIF T R VIS TIY VIP TII PRACTIONS S FMMTI IPG VY AIR VIT A LIVINMV
PR-3 Cat G Gra B RMCP II HLP Chymo.	$\ddot{}$ ŧ. $\ddot{}$: RSTLRRWEAKGP Elastase : DSIIORSEDNPCPHPRDPDPASRTH : RITTMRISFKLLDOMETPL : KKTIMKSS NAVIN KKTIMKRY QQTTAAN	228 238 235 227 224 227 230

specificity of the substrate-binding pocket in serine proteases (36). In all these positions (Gly¹⁷⁰, Ser¹⁹¹, Phe¹⁹², Val¹⁹³, Asp²⁰⁴; marked by filled diamonds in Fig. 3), p29b matches elastase, but in four of the five cases there is no match to the other proteases listed. This suggests that the substrate specificity of p29b is very similar to that of elastase.

Southern Hybridization. Human genomic DNA was restricted with five enzymes and hybridized with [32P]-labeled p29b cDNA. At high stringency, ^a single strongly hybridizing band was noted in digests with HindIII (\sim 20 kb), BamH1 (\sim 12 kb), EcoR1 (\sim 9.6 kb), and BglII (\sim 18 kb). Pst1 digestion yielded no hybridizing bands within the resolving range of the gel. One or two faint additional bands were detected after digestion with EcoRl and Bg1II. At low stringency, no changes in band intensity and no additional bands were observed (not shown). Thus, there appears to be a single human gene for p29b. This gene may contain introns with sites for BglII and EcoRl.

Discussion

Close relatedness among PR-3, p29b, and the cytoplasmic autoantigen of Wegener's granulomatosis has emerged from their subcellular localization, relative abundance, molecular mass, enzymatic activity, ability to bind DFP, and NH2 terminal sequences. High homology with the predicted protein termed myeloblastin was also evident from its cDNA sequence. However, apparent discrepancies in their partial amino acid sequences have left unclear the precise interrelatFigure 3. Alignment of the predicted p29b (PR-3) sequence with other serine proteases. Gaps were introduced to maximize alignment. Conserved residues are boxed. The His, Asp, Ser residues comprising the catalytic sites are indicated by asterisks . Residues lining the substrate binding pocket of serine proteases (36) are marked by diamonds. The human elastase sequence is from reference 28, human cathepsin G (Cat G) fromreference 29, mouse granzyme B (Gra B) from reference 43, rat mast cell protease ¹¹ (RMCP 11) from reference 30, human lymphocyte protease (HLP) from reference 44, and bovine chymotrypsin A (Chymo.) from reference 45.

edness of these entities, which were isolated by independent techniques according to distinct bioactivities. By obtaining what to our knowledge is the first cDNA sequence encoding the complete mature form of any one of these species, we have been able to perform Southern blotting experiments that strongly suggest the existence of a single hybridizing gene, even at low stringency. It is highly likely that PR-3, p29b, the predominant cytoplasmic autoantigen of Wegener's granulomatosis, and myeloblastin are all encoded by this gene. We propose that the gene product be called proteinase ³ (PR-3), a term with historical precedence (7) .

The predicted amino acid sequence of PR-3 differs in 7 of92 residues from the sequences determined from the purified protein or its fragments . Six of the seven discrepant positions are highly conserved among serine proteases; in each case, the conserved amino acid is that predicted by the cDNA. The discrepancies may arise from ambiguities in amino acid sequencing. Alternatively, PR-3 may be polymorphic. The nonconservative nature of the substitutions militates against this; nonetheless, we plan to test for polymorphism by PCR techniques. Finally, we cannot exclude the remote possibility that an additional gene(s) is linked in tandem to PR3 without intervening sites for the restriction enzymes used in the Southern blotting experiments. Genomic cloning will resolve this issue .

The PR-3 cDNA encodes ^a precursor with signal (pre) and propeptides typical of serine proteases. The hydrophobic prepeptide of²⁴ or more residues (the cDNA may be incomplete at the ⁵' end) may target PR-3 to the endoplasmic retic-

Figure 4. Southern blot of human genomic DNA with p29b (PR-3) cDNA as a probe. Human placental genomic DNA was digested to completion with five restriction enzymes and subjected to electrophoresis in 0.7% agarose. (A) The gel was stained with ethidium bromide and photographed under short-wavelength UV light to show the extent of digestion of the genomic DNA and the migration of size markers (left lane) derived from phage λ DNA digested with HindIII (23, 9.6, 6.6, 4.4, 2.3, 2 .0 kb) . (B) The digests were transferred to ^a nylon membrane, hybridized with [32PJ-labeled p29b (PR-3) cDNA under high stringency, and washed as described in Materials and Methods, then autoradiographed with an intensifying screen for 4 d at -70° C.

ulum and be removed in the process. The two-residue propeptide, which resembles that of elastase, cathepsin G, and rat mast cell protease II, may confer zymogen status. The mature protein shares not only the His, Asp, Ser catalytic triad with other serine proteases, but also the sequence at residues 9-16, which constitutes a hallmark of granzymes. Finally, PR-3 shares the substrate binding pocket of elastase (32). Given the ability of PR-3 to hydrolyze the same substrates as elastase (4, 6) with ^a similar specific activity (4), and considering that twice as much PR-3 as elastase can be purified from PMN (4), PR-3 may share ^a major role in inflammatory states where PMN elastase has been implicated (5, 28).

The difference between the calculated pI of PR-3 (7.9) and the experimental value (9.1, reference 6; $>$ 9.5, reference 4) might reflect the predicted preferential distribution of positively charged amino acids on the surface of the protein. Such is the case with elastase, in which Arg residues are exclusively surface disposed (32). It will be of interest to determine the three-dimensional structure of PR-3, both to confirm the predicted distribution of surface charges and to learn whether the alternation of hydrophilic and hydrophobic regions in the primary sequence (Fig. 2) is associated with amphiphilicity, a functionally important property proposed for a variety of antimicrobial polypeptides (reviewed in reference 37). By X-ray crystallography, elastase shares the two short α -helices and predominant, hydrophobic β -sheets predicted for PR-3 (32) . However, the hydrophilic-hydrophobic periodicity seen in PR-3 is not observed in elastase or cathepsin G (not shown) .

Although PR-3 is 35% homologous to cathepsin G, PR-3 lacks the peptides IIGGR and HPQYNQR, which were proposed to contribute to cathepsin G's antibacterial activity (38). Nor does PR-3 display homology to any of the nonproteolytic antibacterial proteins from human or bovine neutrophils whose primary sequences have been reported (37, 39-41) . Of 10 antimicrobial proteins of known sequence in the azurophil granules of human PMN, two are thought to be unique in primary structure (lysozyme [39] and bactericidal/permeability increasing factor [41]), while the remaining eight fall into two families of four members each: the defensins (3, 40) on the one hand, and cathepsin G, elastase, PR-3, and azurocidin (3, 4) on the other hand. Azurocidin has a Gly for Ser substitution in the catalytic triad (26) and therefore lacks the ability to bind DFP (4, 26) or hydrolyze proteins (4), but is closely related to the other members of the group by its NH2 terminal sequence, molecular mass, basic pI, and broadspectrum antibiotic activity. Collectively, these latter four proteins could be termed "serprocidins," to denote that they represent, or are closely related to, serine proteases with microbicidal activity.

We thank A. Sievertsen and A. Marra for help and suggestions, and D. Romeo for communicating results before publication.

This work was supported by National Institutes of Health grants AI-23807, CA-45218, and CA-43610.

Address correspondence to J. Gabay, Box 57, Cornell University Medical College, 1300 York Avenue, New York, NY 10021.

Received for publication 18 September 1990.

References

- 1. Odeberg, H., and I. Olsson. 1975. Antibacterial activity of cationic proteins from human granulocytes . J. Clin. Invest. 56:1118 .
- 2. Elsbach, P., and J. Weiss. 1988. Phagocytic cells: oxygenindependent antimicrobial systems. In Inflammation: Basic Principles and Clinical Correlates. J.I . Gallin, ^I .M. Goldstein, and
- 1713 Campanelli et al.

R. Synderman, editors. Raven Press, Ltd., New York. 445-470.

- 3. Gabay, J.E ., R.W Scott, D. Campanelli, J. Griffith, C. Wilde, M.N. Marra, M. Seeger, and C.F. Nathan. ¹⁹⁸⁹ . Antibiotic proteins of human polymorphonuclear leukocytes. Proc. Natl. Acad. Sci. USA. ⁸⁶ :5610.
- 4. Campanelli, D., P.A . Detmers, C.F. Nathan, and J.E. Gabay. 1990. Azurocidin and a homologous serine protease from neutrophils . Differential antimicrobial and proteolytic properties. J. Clin. Invest. 85:904.
- 5. Weiss, S.J. Tissue destruction by neutrophils. 1989. N. Engl. J. Med. 320:365 .
- 6. Kao, R.C ., N.G. Wehner, K.M. Skubitz, B.H. Gray, and J.R. Hoidal. 1988. Proteinase 3: A distinct human polymorphonuclear leukocyte proteinase that produces emphysema in hamsters. J. Clin. Invest. 82:1963.
- 7. Baggiolini, M., U. Bretz, B. Dewald, and M.E . Feigenson. 1978. The polymorphonuclear leukocyte. Agents Actions. 8:3.
- 8. Wehner, N.G. ¹⁹⁸⁷ . Biochemistry and function of proteases isolated from the cytoplasmic granules of polymorphonuclear leukocytes. Ph. D. Thesis, University of Minnesota, Minneapolis, MN. ¹⁵⁵ pp.
- 9. Falk, R.J., R.S. Terrell, L.A. Charles, and J.C. Jennette. 1990. Anti-neutrophil cytoplasmic autoantibodies induce neutrophils to degranulate and produce oxygen radicals in vitro. Proc. Natl. Acad. Sci. USA. 87:4115.
- ¹⁰ . Goldschmeding, R., C.E . van der Shoot, D. ten Bokkel Huinink, C.E . Hack, M.E. van den Ende, C.G.M. Kallenberg, and A.E.G. kr. von dem Borne. 1989. Wegener's granulomatosis autoantibodies identify a novel diisopropylfluorosphate-binding protein in the lysosomes of normal human neutrophils. J. Clin. Invest. 84 :1577.
- ¹¹ . Niles, J.L ., R.T. McCluskey, M.F. Ahmad, and M.A . Arnaout. 1989. Wegener's granulomatosis autoantigen is a novel serine proteinase. Blood. 74:1888.
- 12. Ludemann, J., B. Utecht, and W.L. Gross. 1990. Anti-neutrophil cytoplasm antibodies in Wegener's granulomatosis recognize an elastinolytic enzyme. J. Exp. Med. 171:357.
- 13. Jennette, J.C., J.R. Hoidal, and R.J. Falk. 1990. Correspondence. Specificity of anti-neutrophil cytoplasmic autoantibodies for proteinase 3. Blood. 75:2263.
- ¹⁴ . Bories, D., M.C. Raynal, D.H. Solomon, Z. Darzynkiewicz, and Y.E. Cayre. 1989. Down-regulation of a serine protease, myeloblastin, causes growth arrest and differentiation of promyelocytic leukemia cells. Cell. 59:959.
- 15. Jenne, D.E., J. Tschopp, J. Ludemann, B. Utecht, and W.L. Gross. 1990. Scientific Correspondence. Wegener's autoantigen decoded. Nature (Lond.). 346:520.
- 16 . Matsudaira, F. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. 262:10035 .
- ¹⁷ . Young, R.A ., and R.W. Davis. ¹⁹⁸³ . Yeast RNA polymerase II genes. Isolation with antibody probes. Science (Wash. DC). 222:778.
- 18. Helms, C., M.Y. Graham, J.E. Dutchik, and M.V. Olson. 1985. A new method for purifying lambda DNA from phage lysates. DNA (NY). 4:39.
- 19. Bullock, W.O., J.M. Fernandez, and J.M. Short. 1987. XL1-Blue. Ahigh efficiency plasmid transforming recA Escherichia $coll$ strain with β -galactosidase selection. Biotechniques. 5:376.
- 20. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Second edition, vol. 3. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. pg. A10.
- 21. Kraft, R.J., K.S. Tardiff, L. Krauter, and L. Leinwand. 1988.

Using mini-prep plasmid DNA for sequencing double-stranded
templates with Sequenase^{rM}. Biotechniques. 6:544.

- Using mini-prep plasmid DNA for sequencing doub
templates with Sequenase^{rM}. *Biotechniques*. 6:544.
Chen, E.Y., and P.H. Seeburg. 1985. Supercoil seq
fast and simple method for sequencing plasmid Dl
(NY). 4:165. 22. Chen, E.Y., and P.H. Seeburg. 1985. Supercoil sequencing: a feet and simple method for some is a physical SNA. DNA fast and simple method for sequencing plasmid DNA. DNA (NY) . 4:165.
- 23. Lipman, D.J., and W.R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science (Wash. DC). 22:1435.
- 24. Devereux, J., P. Haeberi, and O. Smithies. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids. Res. 12:387.
- 25. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: ALaboratory Manual . Second edition, vol. 3. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. A1-A13.
- 26. Wilde, C.G., J.L. Snable, J.E. Griffith, and R.W. Scott. 1990. Characterization of two azurophil granule proteases with activesite homology to neutrophil elastase. J. Biol. Chem. 265:2038.
- ²⁷ . von Heijne, G. ¹⁹⁸⁶ . A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683.
- ²⁸ . Takahashi, H., T Nukiwa, K. Yoshimura, C.D. Quick, D.J . States, M.D. Holmes, J. Whang-Peng, T. Knutsen, and R.G. Crystal. 1988 . Structure of the human neutrophil elastase gene. J. Biol. Chem. 263:14739 .
- ²⁹ . Salvesen, G., D. Farley, J. Shuman, A. Przybyla, C. Reilly, and J. Travis. 1987. Molecular cloning of human cathepsin G: Structural similarity to mast cell and cytotoxic T lymphocyte proteinases. Biochemistry. 26:2289.
- 30. Benfey, P.N., F.H. Yin, and P. Leder. 1987 . Cloning of the mast cell proteinase, RMCP II: evidence for cell-specific expression and ^a multi-gene family. J. Biol. Chem. 262:5377.
- 31. Dayhofl; M.O., WC. Barker, and L.T. Hunt. ¹⁹⁸³ . Establishing homologies in protein sequences. Methods Enzymol. 91:524.
- 32. Navia, M.A., B.M. McKeever, J.P. Springer, T.-Y. Lin, H.R. Williams, E.M. Fluder, C.P. Dorn, and K. Hoogsteen. ¹⁹⁸⁹ . Structure of human neutrophil elastase in complex with ^a peptide chloromethyl ketone inhibitor at 1.84-A resolution . Proc. Natl. Acad. Sci. USA. 86:7.
- 33. Chou, P.Y., and G.D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47 :45.
- 34. Emini, E.A ., J.V. Hughes, D.S. Perlow, and J. Boger. ¹⁹⁸⁵ . Induction of hepatitis Avirus-neutralizing antibody by ^a virusspecific synthetic peptide. J. Virol. 55:836.

Ianna D. C. Bou J. A. Headings B. V. Ci.
- ³⁵ . Jenne, D., C. Rey, J.-A. Haefliger, B.-Y. Qiao, P. Groscurth, andJ. Tschopp. ¹⁹⁸⁸ . Identification and sequencing of cDNA clones encoding the granule-associated serine proteases gran zymes D, E, and F of cytolytic T lymphocytes. Proc. Natl. Acad. Sci. USA. ⁸⁵ :4814.
- 36 . Kraut, J. 1977 . Serine proteases: structure and mechanism of catalysis. Annu. Rev. Biochem. 46:331 .
- 37. Frank, R.W., R. Gennaro, K. Schneider, M. Przybylski, and D. Romeo. 1990. Amino acid sequences of two proline-rich bactenecins. Antimicrobial peptides of bovine neutrophils. J. Biol. Chem. In press.
- 38. Bangalore, N., J. Travis, V.C. Onunka, J. Pohl, and W.M. Shafer. 1990. Identification of the primary antimicrobial domains in human neutrophil cathepsin G. J . Biol. Chem. 265:13584.
- ³⁹ . Chung, L.P., S. Keshav, and S. Gordon . ¹⁹⁸⁸ . Cloning the human lysozyme cDNA: Inverted Alu repeat in the mRNA and in situ hybridization for macrophages and Paneth cells. Proc. Natl. Acad. Sci. USA. 85:6227.
- 40. Ganz, T., M.E. Selsted, D. Szklarek, S.S.L. Harwig, K. Daher, D.F. Bainton, and R.I. Lehrer. 1985. Defensins. Natural pep-

tide antibiotics of human neutrophils. J. Clin. Invest. 76:1427.

- 41. Gray, P.W., G. Flaggs, S.R. Leong, R.J. Gumina, J. Weiss, C.E. Ooi, and P. Elsbach. 1989. Cloning of the cDNA of a human neutrophil bactericidal protein. J. Biol. Chem. 264:9505.
- 42. Kyte, J., and R.F. Doolittle. ¹⁹⁸² . A simple method for displaying the hydropathic character of ^a protein. J. Mol. Biol. 157:105 .
- 43. Lobe, C.G., B.B. Finlay, W. Paranchych, V.H. Paetkau, and R.C. Bleackley. ¹⁹⁸⁶ . Novel serine proteases encoded by two

cytotoxic T lymphocyte-specific genes. Science (Wash. DC.). 232:858 .

- 44. Schmid, J., and C. Weissmann. ¹⁹⁸⁷ . Induction of mRNA for a serine protease and a β -thromboglobulin-like protein in mitogen-stimulated human leukocytes. J. Immunol. 139:250.
- 45. Dayhoff, M.O. 1978. Atlas of Protein Sequences and Structure. Vol. 5, suppl. 3. National Biomedical Research Foundation. Washington, D.C. 73-93.