# Complete Nucleotide Sequence of the Structural Gene for Alkaline Proteinase from *Pseudomonas aeruginosa* IFO 3455

KENJI OKUDA,<sup>1</sup>\* KAZUYUKI MORIHARA,<sup>1</sup> YUSAKU ATSUMI,<sup>1</sup> HIROAKI TAKEUCHI,<sup>1</sup> SUSUMU KAWAMOTO,<sup>1</sup> HIROSHI KAWASAKI,<sup>2</sup> KOICHI SUZUKI,<sup>2</sup> and JUN FUKUSHIMA<sup>1</sup>

Department of Bacteriology, Yokohama City University School of Medicine, 3-9, Fukuura, Kanazawa-ku, Yokohama 236,<sup>1</sup> and Department of Molecular Biology, The Tokyo Metropolitan Institute of Medical Science, 3-18-2, Honkomagome, Bunkyo-ku, Tokyo 113,<sup>2</sup> Japan

Received 21 February 1990/Accepted 17 September 1990

The DNA-encoding alkaline proteinase (AP) of *Pseudomonas aeruginosa* IFO 3455 was cloned, and its complete nucleotide sequence was determined. When the cloned gene was ligated to pUC18, the *Escherichia coli* expression vector, the gene-incorporated bacteria expressed high levels of both AP activity and AP antigens. The amino acid sequence deduced from the nucleotide sequence revealed that the mature AP consists of 467 amino acids with a relative molecular weight of 49, 507. The amino acid composition predicted from the DNA sequence was similar to the chemically determined composition of purified AP reported previously. The amino acid sequence analysis revealed that both the N-terminal side sequences. The percent homology of amino acid sequences between AP and *Serratia* protease was about 55%. The zinc ligands and an active site of the AP were predicted by comparing the structure of the enzyme with of *Serratia* protease, thermolysin, *Bacillus subtilis* neutral protease, and *Pseudomonas* elastase.

Several extracellular products have been implicated in the pathogenicity of Pseudomonas aeruginosa. These include proteases, phospholipase, hemolysin, exotoxin A, and exoenzyme S (24, 32). P. aeruginosa can produce two or three proteases (21). One of them is alkaline proteinase (AP) (19, 20). Enzyme production was observed in a semisynthetic medium containing Ca as an essential component (22). AP can be regarded as a metalloproteinase, since it is inactivated by the addition of o-phenanthroline (25). However, its properties are completely different from those of Pseudomonas elastase, a known Zn-metalloproteinase (24). The chemically determined molecular weight was reported to be 48,400 (10), and the substrate specificity differed completely from those of Zn-metalloproteinases, including Pseudomonas elastase (23, 24, 26, 27). To determine more detailed characteristics, the entire amino acid sequence of AP was necessary.

We were recently able to clone the AP gene (1). In addition, Guzzo et al. (7) cloned the AP gene (apr), and they found the transposable promoter that secreted AP. In the present study, we were able to determine the entire sequence of *Pseudomonas* AP. The characteristics of the deduced AP amino acid sequence of the AP protein, as well as the amino acid homology of the active site of other Zn-metalloproteinases, were also studied.

### **MATERIALS AND METHODS**

**Bacterial strains and vectors.** *P. aeruginosa* IFO 3455 was obtained from the Institute for Fermentation, Osaka, Japan. Both *Escherichia coli* HB101 and JM101 were used for DNA cloning. M13mp18 phages were used for DNA sequencing. *P. aeruginosa* IFO 3455 produced AP antigen and showed azocasein cleavage activity. *E. coli* strains used in the

present study produced neither AP activity nor antigens that reacted with anti-AP antibody.

**Preparation of inactive** *Pseudomonas* **AP.** Crystal *Pseudomonas* **AP** was obtained from Nagase Biochemical Co., Fukuchiyama-shi, Kyoto, Japan. Inactive *Pseudomonas* **AP** was prepared by using the method described previously (9).

Amino acid sequence determination. For amino acid sequence analysis of peptide fragments, purified AP was digested with a lysyl endopeptidase (Wako Pure Chemical Industries) or trypsin (Biozyme Laboratories Ltd., New South Wales, Australia). The lysyl or tryptic peptides were separated and purified by reversed-phase high-pressure liquid chromatography. The N-terminal sequence of the purified AP, as well as the amino acid sequences of the internal lysyl peptides, was determined by automated Edman degradation with a model 477A gas-phase sequencer (Applied Biosystems, Foster City, Calif.). The resulting phenylthiohydantoin amino acids were analyzed with a model 120A phenylthiohydantoin analyzer (Applied Biosystems).

**Construction and screening of the gene library.** Chromosomal DNA was collected from *P. aeruginosa* IFO 3455 by using conventional methods (14). The method for the construction and screening of the library was described previously (1). In these clones, the AP activity was tested by a modified version of the method of Long et al. (12).

The plasmids of AP-positive bacteria were collected and the appropriate DNA fragments were prepared with various restriction enzymes (Takara Shuzo Co., Ltd., Kyoto, Japan). For nucleotide sequence analysis, several smaller DNA fragments about 500 bp in length were made from the pAPDS2 clone as described previously (1). In addition, staggered deletions were prepared by unidirectional digestion by using the method of Henikoff (8). To determine the span of the AP structural gene, the AP activity of each *E. coli* clone, transformed with deleted DNA, was measured.

DNA sequencing. Various restriction fragments of the AP

<sup>\*</sup> Corresponding author.

														-32	28 GC	GGTC	GACA	GAGO	TCGA	ACA	-308
307	307 ATACGCATGGAATTGTTTTTCAATAAGGCGCTATTTATATGAAACGGATTTTTATAATATCCAGTGTCTTCGGTAGTT -229																				
228	TATC	TTTC	CTGC	CATGA	ATTTO	CATAT	ATG	GGAA	TTCO	TTCC	TGCA	TATO	CTTI	AAGI	GTAG	GCTTA	ATT	TTC	TGA	ACC	-150
149	49 TGTGGACATTGGAAAATGAGCTGTGTCAGTTTGGACTTCGGCCTGATCTGGCCGATAACTGCAATTATCCATCC																				
-70	TTGI	ATCO	GAGI	TAT	гтсст	TTAC	CAACI	TAA	GGGA	TATO	GACA	ATG	TCC	AGC	AAT	TCT	CTT	GCA	TTG	AAA	-1
-9	-9 SD <u>M S S N S L A L K</u> -1																				
1	GGT	CGT	AGC	GAT	GCG	TAT	ACC	CAG	GTA	GAC	AAC	TTC	CTG	CAT	GCC	TAT	GCG	CGG	GGC	GGG	60
1	G	R	S	D	A	Y	Т	Q	. V	D	N	F	L	Н	A	Y	A	R	G	G	20
61	Ma	caa	e pro	oteir GTC	ר געע ג	CCC	CAT	CCC	TCC	ן דאד	I-ter	mina CTC	1 fr	ragme	ent	ccc	CAC	CAG	ለጥር	CTC	120
21	D	E	L	V	N	G	H	P	S	Y	T	V	D	Q	A	A	E	Q	I	L	40
											_			•				•		_	
121	CGC	GAA	CAG	GCG	TCT	TGG	CAG	AAA	GCG	CCG	GGC	GAC	TCG	GTG	CTG	ACC	CTG	TCC	TAT	TCG	180
41	n	E	Ŷ	A	3	w	Ŷ	ĸ				Fra	omer	nt 1							60
181	TTC	CTG	ACC	AAA	CCG	AAC	GAC	TTC	TTC	AAT	ACG	CCG	TGG	AAG	TAT	GTC	AGC	GAT	ATC	TAC	240
61	. F	L	Т.,	<u>.</u> K	. P.	N	D	F	. F	N	T.	Р	W	к	<b>.</b> Y	<u>V</u>	S	D	I.	Y	80
241	тсс	стс	GGC	AAG	TTC	AGC	GCC	ጥጥጥ	TCC	CCC.	CAG	CAG	CAG	GCC	CAG	ccc	Fra	agmer	It 2	CTG	300
81	S	L	G	K	F	S	A	F	S	A	Q	Q	Q	A	Q	A	K	L	S	L	100
											•	•	•		•						
301	CAA	TCC	TGG	TCG	GAC	GTC	ACC	AAT	ATC	CAC	TTC	GTC	GAC	GCC	GGC	CAG	GGG	CAT	CAG	GGC	360
101	Ŷ	3	w	3	D	v	1	IN	T	п	г	• *•••		A	ragn	ient	3				120
361	GAC	CTG	ACC	TTC	GGC	AAC	TTC	AGC	AGT	AGT	GTC	GGC	GGT	GCG	GCG	TTC	GCC	TTC	CTG	CCG	420
121	D	L	т	F	G	N	F	S	S	S	v	G	G	Α	Α	F	Α	F	L	Ρ	140
421	GAT	GTA	CCG	GAT	GCG	стс	AAG	GGG	CAA	тсс	TGG	TAC	CTG	ATC	AAC	AGC	AGC	TAC	AGC	GCC	480
141	D	V	Р	D	Α	L	K	G	Q	S	W	Ŷ	L	I	N	S	S	Y	S	A	160
											~~~							. – –			
481	AAC	GTC	AAT	P	GCC	AAC	GCG A	AAC	TAC	GGA	CGC	CAG	ACC	CTG	ACC	CAC	GAG	ATC	GGC	CAT	540 180
101		•		•		.,		.,	•	ŭ		¥	1	5	•		Ľ	1	ŭ		100
541	ACC	CTG	GGC	CTG	AGC	ACC	CCG	GAC	TAC	AAC	GCC	GGC	GAG	GGC	GAT	ccc	ACC	TAC	GCC	GAC	600
181	т	L	G	L	s	Т	Ρ	D	Y	N	A	G	Е	G	D	Р	Т	Y	Α	D	200
601	GCT	ACC	TAC	GCC	GAG	GAC	ACC	CGC	GCC	ТАТ	TCG	GTG	ATG	AGC	TAC	TGG	GAA	GAG	CAG	AAC	660
201	Α	Т	Y	Α	Е	D	Т	R	Α	Y	S	v	М	S	Y	W	Е	Е	Q	N	220
cc1	• • • •	~~~		~~~	<b>mm</b> /2		000	000	<b></b>	maa	<b>m</b> .c.c.	~~	000	000	000	~ ~ ~	~~~		000	000	700
221	ACC T	G	CAG Q	D	F	K	G	A	Y	S	S	GCA	P	UTG L	UTG L	GAC	GAC	ATC	GCG	GCG A	240
	Fragment 4																				
721	ATC	CAG	AAG	CTC	TAC	GGG	GCC	AAC	CTG	ACC	ACC	CGC	ACC	GGC	GAC	ACG	GTG	TAC	GGC	TTC	780
241	I	Q	К	L	Y	G	Α	N	L	Т	Т	R	Т	G	D	Т	v	Y	G	F	260
781	AAC	тсс	AAC	ACC	GAG	CGC	GAC	TTC	TAC	AGC	GCC	ACC	TCG	тсс	AGT	TCC	AAG	CTG	GTG	TTC	840
101																					

FIG. 1. DNA sequence and deduced amino acid sequence of AP. Both strands were sequenced by using several independent, overlapping clones. A putative ribosome-binding site (Shine-Dalgarno-like sequence) is underlined. The start site of the mature product is amino acid 1. The N terminus of the mature protein was identified on the basis of its homology to the N-terminal sequence of chemically determined amino acids of the AP fragment. Dashed lines indicate chemically identified peptides. Numbers written on both sides of the lines indicate the positions of both nucleotides and amino acids. A palindromelike structure or a putative transcription terminator is indicated by arrows.

gene were subcloned into M13mp18 (44). Sequencing reactions were carried out by dideoxy chain termination (34) with <sup>32</sup>P-labeled nucleotide. The region of the structural gene was sequenced by using universal or synthetic oligonucleotides primers. The oligonucleotides were synthesized by use of an automatic DNA synthesizer (model 381A; Applied Biosystems).

**Hydropathy analysis.** Hydropathy analysis was performed by using a method described previously (11).

Analysis of amino acid homology. The program of homology analysis was as described in the Genetyx manual (Software Development Co., Tokyo).

## RESULTS

**DNA sequencing.** As described previously (1), the AP101 clone was further subcloned into a smaller DNA fragment into *E. coli* (pAPDS2). Several smaller DNA fragments about 500 bp in length were made from the pAPDS2 clone for DNA sequencing. In addition, staggered deletions of about 300 bp each were also prepared by unidirectional digestion with exonuclease III. The nucleotide sequence of the AP gene and the deduced primary structure of the protein encoded by this gene are shown in Fig. 1. The entire *Pseudomonas* AP reading frame was sequenced on both

#### STRUCTURAL GENE FOR *PSEUDOMONAS* ALKALINE PROTEINASE 4085

841 TCG GTG TGG GAC GCC GGC GGC AAC GAC ACC CTG GAC TTC TCC GGC TTC AGC CAG AAC CAG 900 281 S G N D Т L D F S G F S Q N Q 300 v W D Α G 901 AAG ATC AAC CTC AAC GAG AAG GCG CTG TCC GAT GTC GGC GGG TTG AAG GGC AAT GTG TCG 960 D v G Κ G Ν v 320 Е Κ L S G L N Α 301 K I Ν L 961 ATC GCT GCC GGG GTC ACC GTG GAA AAC GCC ATC GGC GGC TCG GGT AGC GAC CTG TTG TAC 1020 G G S S D L L 340 Е N Т G Т v Α 321 I Α Α G v 1021 GGC AAC GAC GTG GCC AAC GTG CTC AAG GGC GGC GCC GGC AAC GAC ATC CTC TAC GGC GGC 1080 D Y 360 G G Ν Т I. G G v К G Α 341 G N D v Α N L 1081 CTC GGC GCG GAC CAG TTG TGG GGC GGC GCG GGG GCC GAC ACC TTC GTC TAC GCG ATA TCG 1140 G Α F v S 380 361 L D Q w G G Α D Т Y Α I G Δ Ĩ. 1141 CCG AGT CCT CCG CGC GCG CCG GAT ACC CTG CGC GAC TTC GTC AGC GGC CAG GAC AAG ATC 1200 381 P S Р Р R Α Ρ D Т L R D F V S G Q DKI 400 1201 GAC CTG TCC GGG CTG GAT GCC TTC GTC AAC GGC GGG CTG GTG CTG CAA TAC GTC GAC GCC 1260 401 <u>D L S G L D A F V N G G L V L Q Y V D A</u> 420 Fragment 5 1261 TTC GCC GGC AAC GCG CAG GGC ATC CTG TCC TAC GAC GCG GCG AGC AAG GCC GGC AGC CTG 1320 Q G I L S Y D Α Α S K Α G S L 440 421 <u>F</u> A G N Α 1321 GCA GTC GAC TTC AGC GGG GAC CGC CAT GCC GAT TTC GCG ATC AAT CTG ATC GGC CAG GCG 1380 460 D F I N L I G 441 A v D F D R Н Α Α Α S G 1381 ACC CAG GCC GAC ATC GTG CTC TGA CGCGCTGAAGCGCTGACGTTCGGTGCCGGTACGGACATGCTCCGTCG 1451 461 Т Q Α D I v L 1531 CTGGGCAGCAAGCAGGGCCCCGGCGCCTGGCGCGCTCGCGGTTCGAAGCGAGGATGGCGGGGGCTTCGCGTGTTATCC 1609 1689 TTTCAGCACAGGTATCAGTATGGCCAGCAGTCTGATTCTTCTCAGCGCTTCCGATCTCGCGGGCAATGGACCCTGCAGC 1767 1768 AGGACGAGGTCAGCCCGCGATCTGCCACCTGGAGCTGGCCGACAGCGAAGTGGCGGAACGCAGTGGCTACGACCTGGCG 1846 1926 GCGGCCTGACCCTGATGCTCCTCGGTCCAGGGCGAGGGCGACTAACCGGGTGCAGCGAGAGCGGCGGGCAGTTGGTGCT 2004 2005 GCGCGGCGACGCCCTAGCTCGCTGCGGGTCAGACCACGAAGCGACCGATCATGCCCTTGAGGTCGTTGGCCAGGCGCGA 2083 2084 CAGCTCGTGGGNGTGCGGGGTCTCGTCGGCGCCCCAGGGTGGTCTGCACGGCAACGTCGCGGATGTTCACCAGGTTGCGAT 2102 2103 CGACTCGCGGGGGGGGGGGGGGCCCCTCTAGAG 2130

FIG. 1—Continued.

strands. Both the N-terminal side amino acid sequence of the purified AP (Gly-1 to Gly-20) and the four internal lysyl peptides, fragment 1 (Ala-49 to Thr-71), fragment 2 (Tyr-75 to Gly-83), fragment 3 (Val-112 to Leu-122), and fragment 4 (Asp-232 to Gly-240), derived from AP corresponded completely to the deduced amino acid sequences. The G+C content of this mature AP gene was 64.0%.

ATG, the initiation methionine codon, is located 25 to 27 bp upstream of the mature protein coding sequence. A Shine-Dalgarno-like sequence was also observed upstream from this ATG codon. However, another possibility is that Leu is the initial amino acid of the mature AP protein (see below), since the termination codon TGA was observed at nucleotide positions 1402 to 1404 bp. In addition, the  $M_r$  determined by Inoue et al. (10) supports these results. We were able to observe a palindromelike structure or a putative transcription terminator (1425 bp to 1433 bp), which also supported the existence of termination codon at this site.

**Comparison of amino acid composition.** In a previous study, Morihara et al. (29) reported the amino acid composition of *Pseudomonas* AP by using a purified protein sample. A comparison of the amino acid composition of Morihara et al. with ours is shown in Table 1. The amino acid composition deduced from our DNA sequence was similar to that determined from purified AP.

**Hydropathy profile.** Hydrophobic amino acids were observed in the initial sequence (amino acid 1 to 325). The amino acid sequence of the C-terminal region (amino acid 326 to 467) was strongly hydrophilic.

Amino acid sequence comparison. The amino acid sequence of *Pseudomonas* AP was compared with that of *Serratia* protease. The *Serratia* protease is a Zn-metalloproteinase (19), and its substrate specificity is very similar to that of *Pseudomonas* AP (18, 23). In addition, the amino acid sequence of *Serratia* protease has also been elucidated (30). The frequency of identity appears to be most evident in the central position of the molecule (Fig. 2).

TABLE 1. Amino acid composition of AP

	No. of res	sidues
Amino acid	Determined from purified enzyme <sup>a</sup>	Predicted from DNA sequence
Asp	67.2 (Asp + Asn)	41
Asn		28
Thr	24.5	26
Ser	42.8	46
Glu	36.1 (Glu + Gln)	11
Gln		25
Pro	10.8	14
Gly	66.3	52
Ala	58.9	54
Val	25.0	27
Met	0	1
Ile	17.1	17
Leu	38.1	40
Tyr	20.9	23
Phe	19.9	23
Lys	15.8	15
His	6.0	7
Arg	7.2	10
Cys	0	0
Trp	6.2	7

" Calculated assuming a molecular weight of 49,507 and by using the data of Morihara et al. (21, 24).

INFECT. IMMUN.

## DISCUSSION

Recent studies into the molecular genetics of *P. aeruginosa* have revealed the amino acid sequences of exotoxin A (13, 14, 33, 39), hemolysin (24), carboxypeptidase G2 (17), pilin (31, 36), and elastase (2, 4, 35, 42), as well as molecular size of proteins (3, 10) or *P. aeruginosa* gene expression (3, 5, 6). AP is also thought to be one of the important pathogenic factors in *Pseudomonas* infection, and a vaccine involving this protein mixture has been shown to be effective (9). Molecular genetic approaches are therefore expected to yield important information in this area of research.

We have cloned the *Pseudomonas* AP gene and determined the entire DNA and amino acid sequence. The details of this gene cloning have been reported previously (1). The translated gene product of the pAPDS2 clone displayed both the AP activity and almost the same  $M_r$  as that of purified AP. In addition, the characteristics of the deduced amino acid sequence of AP were similar to those of purified *Pseudomonas* AP. These results strongly indicate that the gene we cloned is a true AP gene.

We tentatively decided that the initiating amino acid of the mature protein is Gly on the basis of only amino acid sequence data. If this is the case, some other proteolytic enzyme might break the signal peptide between Lys and Gly. Another possibility is that Leu-2 is the initiating amino acid, since Ala is a rather common signal sequence terminal amino acid (16) and Ser-Leu-Ala (-6 to -4) is the typical amino acids of the "(-3, -1)-rule" for a signal sequence cleavage

AP	1	GRSDAYTQVDNYLHAYARGGDELVNGHPSYTVDQAAEQILRGQASWQKAPGDSVLTLS	57
SP	1	AATTGYDAVDDLLHYHERGNGIQINGKDSFSNEQAGLFITRENQIWNGYKVFGQPV-KLT	59
AP	58	YSFLIKPNDFFNIPWKYVSDIYSLGKFSAFSAQQQAQAKLSLQSWSDVINIHFVDAGQGH 1	.17
SP	60	FSFPDYKFSSTNVAGDTGLSKFSAEQQQQAKLSLQSWADVANITFTEVAAGQ	111
AP	118	QEDLIFFENFSSSVGGAAFAFLPDVPDALKGQSWYLINSSYSANVNPANANY **** * * * * * **** * * ** * * *	168
SP	112	KANITFONYSQDRPGHTDYGTQATAFLPNTIWQGQDLGGQTWYNVNQSNVKHPATEDY	169
AO	169	GROTLITHEIGHTIGLSTP-DYNAGEGDPTYADATYAEDTRAYSVMSYWEEONTGODFKGA	227
SP	170	GRQIFTHEIGHALGLSHPGDYNAGEGNPTYRDVTYAEDTRQFSLMSYWSEINIGGDNGGH	229
AP	228	YSSAPLLDDIAAIQKLYGANLTTRTGDTVYGFNSDTERDFYSATSSSSKLVFSVWDAGGN * ********** ***** ****** **********	287
SP	230	YAAAPLLDDIAAIQHLYGANLSTRTGDTVYGFNSNTERDFLSTTSNSQKVIFAAWDAGGN	289
AP	288	DTLDFSGFSQNQKININEKALSDVGGLKGNVSIAAGVIVENAIGGSGSDILYGNDVANVL ** **** ** ****** *******************	347
SP	290	DTFDFSGYTANORININEKSFSDVGGLKGWSIAAGVTIENAIGFROR-LIVGNAANNVL	348
AP	348	KGGAGNDILYGGLGADQLWGGAGADTFVY-AISPSPPRAPDTLRDFVSGQDKIDLSGLDA	406
SP	349	KGGAGNDVLFGGGGADQLWGGAGLDIFVFSAASDSAPGASDWIRDFQKGIDKIDLSFFNL	409
AP	407	FVNGGLVLQYVDAFAQNAQGILSYDAS-KAGSLAVDFSGDRHADFAINLIGQATQADIVL	466
SP	409	EAOSSDFIHFVDHFSGAAGALLSYNASNNVIDLSVNIGGHQAPDFLVKIVGQVDVATDFIV	470

FIG. 2. Comparison of the amino acid sequence of AP with that of *Serratia* protease. The amino acid sequence of AP is from Fig. 1, and that of *Serratia* protease is from Nakahara et al. (30). Gaps (indicated by dashes) were introduced to increase the similarity, and matching amino acids are indicated by asterisks. The probable zinc ligands are boxed.

			nebraaco
A)	AP	TLTH <sup>*</sup> EIGH <sup>*</sup> TL	173-182
	SP	TFTH <sup>*</sup> EIGH <sup>*</sup> AL	173–182
	BN	<b>V Т А Н<sup>*</sup> Е М Т Н<sup>*</sup> G V</b>	361-370
	Т	V V A H <sup>*</sup> E L T H <sup>*</sup> A V	139–148
	Е	VVAH <sup>*</sup> EVSH <sup>*</sup> GF	137–146
B)	AP	LNE <sup>*</sup> KALSDVGG	304-314
	SP	LNE <sup>*</sup> KSFSDVGG	305-315
	BN	LNE <sup>*</sup> – SFSDV <u>G</u> G	165–174
	Т	INE <sup>*</sup> – AISDIFG	164–173
	Е	MNE <sup>*</sup> – AFSDMAG	162-171
C)	AP	QDFKG-A Y	223-229
	SP	G D N G G – H <sup>*</sup> Y	224-230
	BN	G D Y G G V H <sup>*</sup> T	222-229
	Т	Q D N G G V H <sup>*</sup> I	225-224
	Е	– <u>D</u> – – – v н <sup>*</sup> н	221-224

Regidues

FIG. 3. Comparison of the regions containing Zn ligands and active sites of AP, *Serratia* protease, thermolysin, *B. subtilis* neutral protease, and *Pseudomonas* elastase. Abbreviations: SP, *Serratia* protease; T, thermolysin; BN, *B. subtilis* neutral protease; E, *Pseudomonas* elastase. The proposed Zn ligands and the active sites are indicated by single asterisks.

site (40). However, more thorough analysis is necessary to confirm this.

Although it has not yet been determined what metal is essential for the activity of Pseudomonas AP, the substrate specifity is very similar to that of the Serratia protease but unlike that of typical neutral Zn-metalloproteinases such as thermolysin and Pseudomonas elastase, etc. Both thermolysin and Pseudomonas elastase exhibit their specificity against bulky or hydrophobic amino acid residues at the imino-side of the splitting point (23). In contrast, Pseudomonas AP and the Serratia protease do not show any similarity to them in their specificity against oxidized insulin B chain and synthetic peptides. However, both show very similar specificity against oxidized insulin B chain (23). Therefore, we compared the amino acid sequence of AP with that of Serratia protease (30) by empirically aligning the sequences of these two enzymes to maximize homology. The homology of total protein is about 55% (Fig. 2). The frequency of homology is greatest in the central position of the molecule (Fig. 2; AP numbers 36 through 403). In contrast, for the sequences of the N-terminal region (AP numbers 1 through 35) and the C-terminal region (AP numbers 404 through 467), homologies are relatively low.

The pattern of homology is particularly evident in the regions that include structurally and functionally important residues which have been assumed to be regions of Zn ligands in the *Serratia* protease (Fig. 3), where a comparison is also made with thermolysin (37), *B. subtilis* neutral protease (38, 43), and *Pseudomonas* elastase (4, 42). In thermolysin, the Zn ligands are His-142, His-146, and Glu-166, and the active site is His-231 (14, 43). Glu-143 and Asp-226 are

also thought to participate in the catalysis (15). We have previously compared the amino acid sequence of Pseudomonas elastase with the active site sequence of thermolysin (4) and showed that a region containing two histidine residues (at positions 140 and 144) in the elastase was similar to the Zn binding region in thermolysin. The amino acid sequence of AP has also been shown to be similar to that of thermolysin (Fig. 3): three Zn ligands present in thermolysin are present in AP (His-176, His-180, and Glu-314). The Zn ligands and the active site of AP were searched for by comparing the structure of the enzyme with the structure of Serratia protease (30), B. subtilis neutral protease (38, 43), Serratia thermolysin (37), and Pseudomonas elastase (4, 42) (Fig. 3). The region from residue 224 to 230 in Serratia protease is similar to the active-site region in thermolysin (from residue 225 to 232) and to the region from residue 222 to 229 in B. subtilis protease. However, we were not able to observe this active-site histidine in AP.

#### ACKNOWLEDGMENTS

We greatly thank T. Takigami, T. Nishino, and Y. Shibano for their kind suggestions and Y. Sato for her technical assistance. The secretarial assistance of K. Chiba is also appreciated.

This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science, and Culture of Japan (01570242).

#### LITERATURE CITED

- Atsumi, Y., K. Morihara, J. Fukushima, H. Takeuchi, N. Mizuki, S. Kawamoto, and K. Okuda. 1989. Cloning and expression of the alkaline proteinase gene from *Pseudomonas aeruginosa* IFO 3455. J. Bacteriol. 171:5173-5175.
- Bever, R. A., and B. H. Iglewski. 1988. Molecular characterization and nucleotide sequence of the *P. aeruginosa* elastase structural gene. J. Bacteriol. 170:4309–4314.
- Finlay, B. B., B. L. Pasloske, and W. Paranchych. 1986. Expression of the *Pseudomonas aeruginosa* PAK gene in *Escherichia coli*. J. Bacteriol. 165:625–630.
- Fukushima, J., S. Yamamoto, K. Morihara, Y. Atsumi, H. Takeuchi, S. Kawamoto, and K. Okuda. 1989. Structural gene and complete amino acid sequence *Pseudomonas aeruginosa* IFO 3455 elastase. J. Bacteriol. 171:1698–1704.
- Gingeras, T. R., and J. E. Brooks. 1983. Cloned restriction modification system from *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. USA 80:402–406.
- Golderg, J. B., and D. E. Ohman. 1987. Activation of an elastase precursor by the *lasA* gene product of *Pseudomonas aerugi*nosa. J. Bacteriol. 169:4532–4539.
- Guzzo, J., M. Murgier, A. Filloux, and A. Lazdunski. 1990. Cloning of the *Pseudomonas aeruginosa* alkaline protease gene and secretion of the protease into medium by *Escherichia coli*. 172:942–948.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints from DNA sequencing. Gene 28: 351-359.
- 9. Homma, J. Y., C. Abe, K. Tanamoto, Y. Hirano, K. Morihara, H. Tsuzuki, R. Yanagawa, E. Honda, Y. Aoi, Y. Fujimoto, M. Goryo, N. Imazeki, H. Noda, A. Goda, S. Takeuchi, and T. Ishihara. 1978. Effectiveness of immunization with single and multi-component vaccines prepared from a common antigen (OEP) protease and elastase toxoids of *Pseudomonas aeruginosa* on protection against hemorrhagic pneumonia in mink due to *P. aeruginosa*. Jpn. J. Exp. Med. 48:111-133.
- Inoue, H., T. Nakagawa, and K. Morihara. 1963. Pseudomonas aeruginosa proteinase. II. Molecular weight and molecular dimension. Biochim. Biophys. Acta 73:125–131.
- 11. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.
- 12. Long, S., M. A. Mothibeli, F. T. Robb., and D. R. Woods. 1981.

Regulation of extracellular alkaline proteinase activity by histidine in a collagenolytic *Vibrio alginolyticus* strain. J. Gen. Microbiol. **127:1**93–199.

- Lory, S., M. S. Strom, and K. Johnson. 1988. Expression and secretion of the cloned *Pseudomonas aeruginosa* exotoxin A by *Escherichia coli*. J. Bacteriol. 170:714–719.
- 14. Lukac, M., and R. J. Collier. 1988. Restoration of enzymic activity and cytotoxicity of mutant, E553C, *Pseudomonas aeruginosa* exotoxin A by reaction with iodoacetic acid. J. Biol. Chem. 263:6146-6149.
- 15. Matthews, B. W., J. N. Jansonius, P. M. Colman, B. P. Schoenborn, and D. Dupourque. 1972. Three-dimensional structure of thermolysin. Nature (London) New Biol. 238:37-41.
- Michaelis, S., and J. Bechwith. 1982. Mechanism of incorporation of cell envelope proteines in *Escherichia coli*. Annu. Rev. Microbiol. 36:435–465.
- 17. Minton, N. P., T. Atkinson, C. J. Bruton, and R. F. Sherwood. 1984. The complete nucleotide sequence of the *Pseudomonas* gene coding for carboxypeptidase G2. Gene **31**:31–38.
- Miyata, K., K. Tomoda, and M. Isono. 1970. Serratia protease purification and general properities of enzyme. Agric. Biol. Chem. 34:1457-1462.
- 19. Miyata, K., K. Tomoda, and M. Isono. 1971. Serratia protease characteristics of enzyme as metalloenzyme. Agric. Biol. Chem. 35:460-467.
- Morihara, K. 1957. Studies on the protease of *Pseudomonas*. II. Crystallization of the protease and its physicochemical and general properties. Bull. Agric. Chem. Soc. Jpn. 21:11–17.
- Morihara, K. 1963. Pseudomonas aeruginosa proteinase. I. Purification and general properties. Biochim. Biophys. Acta. 73:113-124.
- 22. Morihara, K. 1964. Production of elastase and proteinase by *Pseudomonas aeruginosa*. J. Bacteriol. 88:745–757.
- Morihara, K. 1974. Comparative specificity of microbiol proteinase. Adv. Enzymol. 41:179–243.
- Morihara, K., and J. Y. Homma. 1985. Pseudomonas protease, p. 41-79. In I. A. Holder (ed.), Bacterial enzymes and virulence. CRC Press, Inc., Boca Raton, Fla.
- Morihara, K., and H. Tsuzuki. 1964. Pseudomonas aeruginosa peptide peptidohydorolase. III. Some characteristics of a Cametalloenzyme. Biochim. Biophys. Acta. 92:351-360.
- Morihara, K., and H. Tsuzuki. 1966. Substrate specificity of elastolytic and non-elastolytic proteinase from *Pseudomonas* aeruginosa. Arch. Biochem. Biophys. 114:158–165.
- Morihara, K., and H. Tsuzuki. 1974. Effect of cobalt ion on the enzymatic activity of *Pseudomonas aeruginosa* alkaline proteinase. Agric. Biol. Chem. 38:621-626.
- Morihara, K., H. Tsuzuki, and T. Oka. 1973. On the specificity of *Pseudomonas aeruginosa* alkaline proteinase with synthetic peptides. Biochim. Biophys. Acta 309:414–429.
- Morihara, K., N. Yoshida, and K. Kuriyama. 1964. Pseudomonas aeruginosa peptide peptidohydorolase. IV. Optical rotary dispersion and amino acid composition. Biochim. Biophys. Acta 92:361-366.

- Nakahara, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Hase, and H. Matsubara. 1986. Cloning and sequencing of *Serratia* protease. Nucleic Acids Res. 14:5843-5855.
- 31. Pasloke, B. P., B. B. Finlay, and W. Paranchych. 1985. Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. FEBS Lett. 183:408-412.
- Pavlovaskis, O. R., and B. Wretlind. 1982. Pseudomonas aeruginosa toxins, p. 97-128. In C. S. Easmon, and J. Jeljaszewicz, (ed.), Medical microbiology, vol. 1. Academic Press, Inc. (London), Ltd., London.
- 33. Pritchard, A. E., and M. L. Vasil. 1986. Nucleotide sequence and expression of a phosphate-regulated gene encoding a secreted hemolysin of *Pseudomonas aeruginosa*. J. Bacteriol. 167:291-298.
- 34. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- Schad, P. A., R. A. Bever, T. I. Nicos, F. Leduc, L. F. Hanne, and B. H. Iglewski. 1987. Cloning and characterization of elastase genes from *Pseudomonas aeruginosa*. J. Bacteriol. 169:2691-2696.
- Strom, M. S., and S. Lory. 1986. Cloning and expression of the pilin gene of *Pseudomonas aeruginosa* PAK in *Escherichia coli* J. Bacteriol. 165:367–372.
- Titani, K., M. A. Hermodoson, L. H. Ericsson, K. A. Walsh, and H. Neurath. 1972. Amino-acid sequence of termolysin. Nature (London) New Biol. 238:35–37.
- 38. Vasantha, N., L. D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula. 1984. Genes for alkaline protease and neutral protease from *Bacillus amyloliquefaciens* contain a large open reading frame between the regions coding for signal sequence and mature protein. J. Bacteriol. 159:811–819.
- Vasil, M. L., C. Chamberlain, and C. C. R. Grant. 1986. Molecular studies of *Pseudomonas* exotoxin A gene. Infect. Immun. 52:538-548.
- 40. Von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683–4690.
- 41. Weaver, L. H., W. R. Kester, and B. W. Matthews. 1977. Acrystallographic study of the complex of phosphoramidon with thermolysin. A model for the presumed catalytic transition state and for the binding of extended substrates. J. Mol. Biol. 114:119-132.
- 42. Yamamoto, S., J. Fukushima, Y. Atsumi, H. Takeuchi, S. Kawamoto, K. Okuda, and K. Morihara. 1988. Cloning and characterization of elastase structural gene from *Pseudomonas* aeruginosa IFO 3455. Biophys. Biophys. Res. Commun. 152: 1117–1122.
- 43. Yang, M. Y., E. Ferrain, and D. J. Henner. 1984. Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an in vitro-derived deletion mutation. J. Bacteriol. 160:15-21.
- 44. Yanisch, P. C., V. Jefferey, and M. Joachim. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33:103–119.