

Nucleotide Sequence of the Gene for the Hole-Forming Toxin Aerolysin of *Aeromonas hydrophila*

S. PETER HOWARD,¹ WILLIAM J. GARLAND,² MARGARET J. GREEN,² AND J. THOMAS BUCKLEY^{2*}

Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, V8W 2Y2, Canada,² and Centre de Biochimie et de Biologie Moléculaire, Centre National de la Recherche Scientifique, 13402 Marseille, France¹

Received 5 December 1986/Accepted 13 March 1987

The gene for the hole-forming toxin aerolysin from *Aeromonas hydrophila* was sequenced. Although most of the sequence seems unrelated to that of *Staphylococcus aureus* alpha-toxin, both proteins are very hydrophilic, and they each contain a nearly identical string of 10 amino acids.

Aerolysin is a cytolytic toxin exported by the gram-negative bacterium *Aeromonas hydrophila* (6-9, 14). There are two precursor forms of the toxin. It crosses the inner bacterial membrane as a preprotoxin containing a typical signal sequence, which is removed cotranslationally (8). The resulting protoxin is exported from the bacteria and activated by proteolytic removal of about 25 amino acids from the carboxy terminus (9). Like *Staphylococcus aureus* alpha-toxin (3), the mature protein binds to eucaryotic cells and aggregates to form holes approximately 3 nm in diameter, leading to destruction of the membrane permeability barrier and osmotic lysis (6). We have recently cloned the structural gene for preproaerolysin into *Escherichia coli* (10). Here we present its nucleotide sequence.

The plasmid pPH4 (10) was digested with the restriction enzymes shown in Fig. 1. The resulting fragments were inserted into the polylinker cloning regions of M13 mp18 and mp19 (15) and sequenced (19). The universal 17-base single-stranded primer was used to begin sequencing from the 3' end of the inserts. Sequencing was continued with 18-base oligonucleotide primers prepared by using a Sam One DNA synthesizer (Biosearch Inc., San Rafael, Calif.).

The sequence of the region containing the aerolysin gene is shown in Fig. 2. The longest open reading frame extends from base pairs (bp) 532 to 1989. This would produce a protein with a molecular weight of 53,800, very close to the observed molecular weight of preproaerolysin (8). Several features of the region preceding the gene, including the presence of sequences analogous to Shine-Dalgarno (20) and -10 and -35 regions, have been discussed earlier (10). The translation product of the first 69 bp beginning at 532 is a typical 23-amino-acid signal peptide (17). This is followed by a sequence corresponding to the amino terminus of purified aerolysin (10). In addition, the translated amino acid sequence matches the amino-terminal sequences of two of the cyanogen bromide fragments which were isolated and sequenced after cleavage of purified aerolysin with cyanogen bromide (5). There are no extended open reading frames in the 350-bp 5' untranslated sequence upstream of the gene, which appears to contain a promoter, and there is a palindromic sequence analogous to a Rho-independent termination signal just downstream of the termination codon

(18). These observations, which indicate that aerolysin is translated from a monocistronic message, are inconsistent with the results of Chakraborty et al. (1). Those authors have recently cloned and mapped a fragment of *A. hydrophila* DNA containing a gene for a hemolytic toxin, as well as flanking regions downstream and upstream which they suggest modulate its expression and activity. Our earlier observation that hemolytic activity is not affected in recombinants containing plasmids with Tn5 insertions at or downstream of the *Bam*HI site at bp 2176 (10) also contrasts with their report that hemolysis is changed by transposon insertions in comparable regions of the gene they have cloned. In addition, Chakraborty et al. (1) found that plasmid deletions obtained by digestion with *Pst*I produced nonhemolytic recombinants, yet the sequence reported in Fig. 2 contains no *Pst*I sites (we have confirmed this by restriction analysis of pPH4). Indeed, the restriction map of the aerolysin structural gene in Fig. 2 differs in several other respects from the map reported by Chakraborty et al. (1). This may mean either that the aerolysin genes of the two isolates of *A. hydrophila* have diverged enormously, or that two different hemolysins have been cloned. *A. hydrophila* is known to release a weakly hemolytic phospholipase (14), and it may produce another cytolytic (21).

The experimentally determined amino acid composition of purified aerolysin (Table 1) is very similar to the composition calculated assuming that the site of activation is between Arg-442 and Leu-443. This site was located by digesting the mature toxin with carboxypeptidase Y (12) and determining the amino acids released from the C terminus (not shown here). The nucleotide sequence predicts five cyanogen bromide fragments in both aerolysin and proaerolysin, including an unusually large fragment of 27,900 molecular weight, in good agreement with our experimental observations (9).

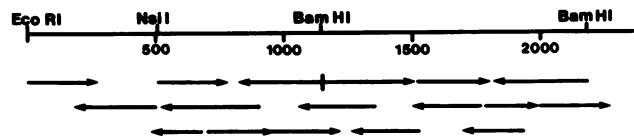


FIG. 1. Strategy for sequencing the aerolysin gene. Arrows represent sequences obtained by using fragments produced by restriction enzyme digestion at the sites indicated or sequences obtained using synthetic oligonucleotides as primers.

* Corresponding author.

15 30 45 60 75 90 105
 GCG CCC GAG TCA GCT GGG GGC GTT CAC TGG CGA CGG GCA CAG GCC CCT TGC TTG CGG TGG CCG GTC ACT CGC TGC AAT TGC AGG GGT TGG GCA CAA TCA CCT TCG
 120 135 150 165 180 195 210
 ATG CCG GCA CCC GCT GGC TCA ACG GCG GTC CCG CCG ATC TGC AAC CCG GTC GCC AAC TGG TGC TGA GCC CCG ATG AAA CCG GTC GGG CAA CCG AGA TCC TGA TCC
 225 240 255 270 285 300 315
 CCA ACC CCG AGG ATG AAC CCG AAT AAG GAT CAT GCA GCC AAA GCG TTA ATA TTT ATT TTG CTA AAT TAG AAA TAT CTT TTT TAT CTA TAT TCC AAA AGA TGA TTA
 330 345 360 375 390 405 420
 AGT GAC GAA TAA AAT AAT AGA GCG AGT GCT CTG ATA TTA TAT CAA TCA ATA TTG AAT GAA GTT CAA TTT ATG ATT TTG TTA ATA TAT TGC GCA TAT TAA AAT GTG
 435 450 465 480 495 510 525
 GGC TGG ATC GCA TAT TGA GAT TAA TGT GAC TGA TAT TGT CGT ACT CAC ATG CCA CCC GCT GAT ATA TAA GGT TGG TGA ATG CAT GTC AAT GTT CAA TAT ATT GGG
 540 555 570 585 600 615 630
 GTT GCT ATG CAA AAA ATA AAA CTA ACT GCC TTG TCA TTA ATC ATA TCC CTG CTG ATG GCA CAG GCG CAA GCG GCA GAG CCC GTC TAT CCA GAC CAG CTT CGC
Met Gln Lys Ile Lys Leu Thr Gly Leu Ser Leu Ile Ile Ser Gly Leu Leu Met Ala Gln Ala Gln Ala Ala Glu Pro Val Tyr Pro Asp Gln Leu Arg
 -23 -20 -10 1 1
 645 660 675 690 705 720 735
 TTG TTT TCA TTG GGC CAA GGG GTC TGT GGC GAC AAG TAT CCG CCC GTC AAT CGA GAA GAA GCC CAA AGC GTT AAA AGC AAT ATT GTC GGC ATG ATG GGG CAA TGG
 Leu Phe Ser Leu Gly Gln Gly Val Cys Gly Asp Lys Tyr Arg Pro Val Asn Arg Glu Glu Ala Gln Ser Val Lys Ser Asn Ile Val Gly Met Met Gly Gln Trp
 750 765 780 795 810 825 840
 CAA ATA AGC GGG CTG GCC AAC GGC TGG GTC ATT ATG GGG CCG GGT TAT AAC GGT GAA ATA AAA CCA GGG ACA CCG TCC AAT ACC TGG TGT TAT CCG ACC AAT CCT
 Gln Ile Ser Gly Leu Ala Asn Gly Trp Val Ile Met Gly Pro Gly Tyr Asn Gly Glu Ile Lys Pro Gly Thr Ala Ser Asn Thr Trp Cys Tyr Pro Thr Asn Pro
 855 870 885 900 915 930 945
 GTT ACC GGT GAA ATA CCG ACA CTG TCT GCC CTG GAT ATT CCA GAT GGT GAC GAA GTC GAT GTG CAG TGG CGA CTG GTA CAT GAC ACT GCG AAT TTC ATC AAA CCA
 Val Thr Gly Glu Ile Pro Thr Leu Ser Ala Leu Asp Ile Pro Asp Gly Asp Glu Val Asp Val Gln Trp Arg Leu Val His Asp Ser Ala Asn Phe Ile Lys Pro
 960 975 990 1005 1020 1035 1050
 AOC AGC TAT CTG GCC CAT TAC CTC GGT TAT GCC TGG GTG GGC GGC AAT CAC AGC CAA TAT GTC GGC GAA GAC ATG GAT GTG ACC CGT GAT GGC GAC GCC TGG CTG
 Thr Ser Tyr Leu Ala His Tyr Leu Gly Tyr Ala Trp Val Gly Gly Asn His Ser Gln Tyr Val Gly Glu Asp Met Asp Val Thr Arg Asp Gly Asp Gly Trp Val
 1065 1080 1095 1110 1125 1140 1155
 ATC CGT GGC AAC AAT GAC GGC GGC TGT GAC GGC TAT CCG TGT GGT GAC AAG ACG GCC ATC AAG GTC AGC AAC TTC GCC TAT AAC CTG GAT CCC GAC AGC TTC AAG
 Ile Arg Gly Asn Asn Asp Gly Gly Cys Asp Gly Tyr Arg Cys Gly Asp Lys Thr Ala Ile Lys Val Ser Asn Phe Ala Tyr Asn Leu Asp Pro Asp Ser Phe Lys
 1170 1185 1200 1215 1230 1245 1260
 CAT GGC GAT GTC ACC CAG TCC GAC CCG CAG CTG GTC AAG ACT GTG GTG GGC TGG GCG GTC AAC GAC AGC GAC ACC CCC CAA TCC GGC TAT GAC GTC ACC CTG CCG
 His Gly Asp Val Thr Gln Ser Asp Arg Gln Leu Val Lys Thr Val Val Gly Trp Ala Val Asn Asp Ser Asp Thr Pro Gln Ser Gly Tyr Asp Val Thr Leu Arg
 1275 1290 1305 1320 1335 1350 1365
 TAC GAC ACA GCC ACC AAC TGG TCC AAG ACC AAC ACC TAT GGC CTG AGC GAG AAG GTG ACC ACC AAG AAC AAG TTC AAG TGG CCA CTG GTG GGG GAA ACC CAA CTC
 Tyr Asp Thr Ala Thr Asn Trp Ser Lys Thr Asn Thr Tyr Gly Leu Ser Glu Lys Val Thr Thr Thr Ser Leu Ser Gln Ser Phe Lys Trp Pro Leu Val Gly Gly Glu Thr Gln Leu
 1380 1395 1410 1425 1440 1455 1470
 TCC ATC GAG ATT CGT GCC AAT CAG TCC TGG GCG TCC CAG AAC GGG GGC TCG ACC ACC ACC TCC CTG TCT CAG TCC GTG CGA CCG ACT GTG GCG GCG CCG TCC AAG
 Ser Ile Glu Ile Arg Ala Asn Gln Ser Trp Ala Ser Gln Asn Gly Gly Ser Thr Thr Thr Ser Leu Ser Gln Ser Val Arg Pro Thr Val Pro Ala Arg Ser Lys
 1485 1500 1515 1530 1545 1560 1575
 ATC CCG GTG AAG ATA GAG CTC TAC AAG GCC GAC ATC TCC TAT CCG TAT GAG TTC AAG GCC GAT GTC AGC TAT GAC CTG ACC CTG AGC GGC TTC CTG CCG TGG GGC
 Ile Pro Val Lys Ile Glu Leu Tyr Lys Ala Asp Ile Ser Tyr Pro Tyr Glu Phe Lys Ala Asp Val Ser Tyr Asp Leu Thr Leu Ser Gly Phe Leu Arg Trp Gly
 1590 1605 1620 1635 1650 1665 1680
 GGC AAC GCC TGG TAT ACC CAC CCG GAC AAC CGT CCG AAC TGG AAC CAC ACC TTC GTC ATA GGT CCG TAC AAG GAC AAG GCG AGC AGC ATT CGC TAC CAG TGG GAC
 Gly Asn Ala Trp Tyr Thr His Pro Asp Asn Arg Pro Asn Trp Asn His Thr Phe Val Ile Gly Pro Tyr Lys Asp Lys Ala Ser Ser Ile Arg Tyr Gln Trp Asp
 1695 1710 1725 1740 1755 1770 1785
 AAG CGT TAC ATC CCG GGT GAA GTG AAG TGG TGG GAC TGG AAC TGG ACC ATA CAG CAG AAC GGT CTG TCT ACC ATG CAG AAC AAC CTG GCC AGA GTG CTG CCG CCG
 Lys Arg Tyr Ile Pro Gly Glu Val Lys Trp Trp Asp Trp Asn Trp Thr Ile Gln Gln Asn Gly Leu Ser Thr Met Gln Asn Asn Leu Ala Arg Val Leu Arg Pro
 1800 1815 1830 1845 1860 1875 1890
 GTG CCG GCG GGG ATC ACC GGT GAT TTC AGT GCC GAG AGC CAG TTT GCC GGC AAC ATA GAG ATC GGT GCT CCC GTG CCG CTC GCG GCT GAC AGC AAG GTG CGT CGT
 Val Arg Ala Gly Thr Gly Asp Phe Ser Ala Glu Ser Gln Phe Ala Gly Asn Ile Glu Ile Gly Ala Pro Val Pro Leu Ala Ala Asp Ser Lys Val Arg Arg
 1905 1920 1935 1950 1965 1980 1995
 GCT CCG AGT GTG GAC GGC GCT GGT CAA GCC CTG AGG CTG GAG ATC CCG CTC GAT CCG GAA GAG CTC TCC GGG CTT GGC TTC AAC AAG TCA GCC TCA CCG TGA CCC
 Ala Arg Ser Val Asp Gly Ala Gly Gln Gly Leu Arg Leu Glu Ile Pro Leu Asp Arg Glu Glu Leu Ser Gly Leu Gly Phe Asn Lys Ser Ala Ser Ala
 2010 2025 2040 2055 2070 2085 2100
 CTG CTG CCA ATC AAT AAC GGC AGC CCG TTG TAG TGA TGG AAC CGG GCC TCT GTG GCC CGG TTT TTG TTT GCA CTG GTC GGG CTT GTT AAA GGC TTG TGC TTT CCA
 2115 2130 2145 2160 2175 2190
 TTT CCC CAC TTA TAC TGG CCG CAT CTT GTC GGA GTG CCA ACC GTC GAA CGA CCG GAG GCT GAG ACC GTT AAT TCG GGA TCC GTG GAA CCT GAT CAG GCT AGC ACC
 2220 2235 2250 2265 2280 2295 2310
 TGC GAA GCG AAA CAA GGG TAA CTT CCG GGT TGC CCG GCC GGG GGA GGG ACA AGC CTC TCC CCG TCA TCA AGA GGA GCC ATT CCT CGA TGA GTC AGG CCG CAC AAG
 2325 2340
 AGG GAG TCT GTC CCG TCC GGT CTG CCC AGG AGG GGC

FIG. 2. DNA sequence of the aerolysin gene. Sequences which could be part of a promoter region and a Rho-independent terminator are underlined, as is the amino acid sequence of the signal peptide. The experimentally determined N-terminal amino acid sequences of two cyanogen bromide fragments are also underlined. Sites of proteolytic processing of preproaerolysin and proaerolysin are marked with arrows.

TABLE 1. Comparison of the amino acid composition of aerolysin derived from the DNA sequence with that obtained by analysis of the purified protein

Amino acid(s)	Composition (mol/mol of protein) determined by:	
	Derivation ^a	Protein analysis ^b
Asp + Asn	59	63
Thr	29	29
Ser	34	33
Glu + Gln	35	39
Pro	23	23
Gly	44	41
Ala	28	26
Cys	4	ND
Val	34	33
Met	5	5
Ile	21	21
Leu	25	24
Tyr	21	21
Phe	10	10
His	6	6
Lys	22	23
Arg	24	20
Trp	18	ND

^a These calculations are based on the assumption that proaerolysin is processed at Arg-442.

^b Average of two determinations. ND, Not determined.

Aerolysin	223	232
	Thr.Ala.Thr.Asn.Trp.Ser.Lys.Thr.Asn.Thr.	
Alpha toxin	Thr.Ser.Thr.Asn.Trp.Lys.Gly.Thr.Asn.Thr.	311
	302	

FIG. 3. Comparison of regions of potential sequence homology between aerolysin and alpha-toxin.

Analysis of the predicted amino acid sequence (2, 11) indicates that the protein is remarkably hydrophilic [average (H)19 hydropathy of -0.57] with no obvious hydrophobic stretches long enough to span a membrane. Because aerolysin forms holes in biological membranes, this is at first surprising; however, the pore-forming outer membrane porins also lack hydrophobic sequences (16). Like these proteins, aerolysin may aggregate to accomplish hole formation. It should also be noted that *S. aureus* alpha-toxin is even more hydrophilic than aerolysin (4).

A search of the National Biomedical Research Foundation protein sequence library (13) did not reveal any sequences homologous to aerolysin, but a direct comparison of aerolysin and alpha-toxin (4) uncovered similar strings of 10 amino acids in each protein. Seven of the 10 are identical, including a Trp in the middle of each sequence (Fig. 3). Another position is occupied by Ser in one protein and homologous Ala in the other. Because the two toxins are functionally so similar, it is tempting to speculate that this small area of similarity is significant.

This research was supported by grants from the National Sciences and Engineering Research Council of Canada and the British Columbia Health Care Research Foundation.

LITERATURE CITED

1. Chakraborty, T., B. Huhle, H. Bergbauer, and W. Goebel. 1986. Cloning, expression, and mapping of the *Aeromonas hydrophila* aerolysin gene determinant in *Escherichia coli* K-12. *J. Bacteriol.* **167**:368-374.
2. Eisenberg, D. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**:125-142.
3. Freer, J. H., and J. P. Arbutnot. 1983. Toxins of *Staphylococcus aureus*. *Pharmacol. Ther.* **19**:55-106.
4. Grant, G. S., and M. Kehoe. 1984. Primary sequence of the alpha toxin gene from *Staphylococcus aureus* Wood 46. *Infect. Immun.* **46**:615-618.
5. Gross, E. 1967. The cyanogen bromide reaction. *Methods Enzymol.* **11**:245-255.
6. Howard, S. P., and J. T. Buckley. 1982. Membrane glycoprotein receptor and hole-forming properties of a cytolytic protein toxin. *Biochemistry* **21**:1662-1667.
7. Howard, S. P., and J. T. Buckley. 1983. Intracellular accumulation of extracellular proteins by pleiotropic export mutants of *Aeromonas hydrophila*. *J. Bacteriol.* **154**:413-418.
8. Howard, S. P., and J. T. Buckley. 1985. Protein export by a gram-negative bacterium: production of aerolysin by *Aeromonas hydrophila*. *J. Bacteriol.* **161**:1118-1124.
9. Howard, S. P., and J. T. Buckley. 1985. Activation of the hole-forming toxin aerolysin by extracellular processing. *J. Bacteriol.* **163**:336-340.
10. Howard, S. P., and J. T. Buckley. 1986. Molecular cloning and expression in *Escherichia coli* of the structural gene for the hemolytic toxin aerolysin from *Aeromonas hydrophila*. *Mol. Gen. Genet.* **204**:289-295.
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. Lee, H.-M., and J. F. Riordan. 1978. Does carboxypeptidase Y have intrinsic endopeptidase activity? *Biochem. Biophys. Res. Commun.* **85**:1135-1142.
13. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science* **227**:1435-1441.
14. MacIntyre, S., L. N. Halasa, K. D. Lund, and J. T. Buckley. 1981. Purification and some properties of the hemolytic toxin aerolysin. *Can. J. Biochem.* **59**:430-435.
15. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:20-77.
16. Mizuno, T., M.-Y. Chou, and M. Inouye. 1983. A comparative study on the genes for the three porins of the *Escherichia coli* outer membrane. *J. Biol. Chem.* **258**:6932-6940.
17. Perlman, D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eucaryotic and procaryotic signal peptides. *J. Mol. Biol.* **167**:391-409.
18. Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**:319-353.
19. Sanger, F., S. Nicklen, and A. R. Coulson. 1979. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
20. Shine, J., and L. Dalgarno. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature (London)* **254**:34-38.
21. Thelestam, M., and A. Ljungh. 1982. Two membrane-damaging toxins from *Aeromonas hydrophila*. *Toxicon* **20**:159-163.