Molecular Analysis of a Sphingomyelinase C Gene from Leptospira interrogans Serovar hardjo

RUUD P. A. M. SEGERS,* ABRAHAM van der DRIFT, ARNE de NIJS, PAOLA CORCIONE, BERNARD A. M. van der ZEIJST, and WIM GAASTRA

Department of Bacteriology of the Institute of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, University of Utrecht, P. O. Box 80.165, 3508 TD Utrecht, The Netherlands

Received 29 January 1990/Accepted 10 April 1990

A thermolabile hemolysin from *Leptospira interrogans* serovar hardjo, strain Sponselee, was shown to specifically degrade sphingomyelin. Nucleotide sequence determination revealed that sphingomyelinase activity was encoded by an open reading frame of 1,668 nucleotides. Although a putative signal sequence could be identified, no evidence for protein export in either *L. interrogans* or *Escherichia coli* was obtained. The apparent molecular mass of the expression product in *E. coli* minicells was 41.2 kilodaltons, whereas open reading frame 1 encoded a protein of 63,268 daltons. The observed difference may be explained by processing at the carboxy-terminal part of the hemolysin in *E. coli*. A high degree of similarity on the DNA and protein levels with *Staphylococcus aureus* β -hemolysin and sphingomyelinase C from three *Bacillus cereus* strains was observed. The presence of various sphingomyelinase genes within the *L. interrogans* species is demonstrated.

Leptospira interrogans is the etiologic agent of leptospirosis, which is a worldwide zoonosis. The bacteria of this species are divided into 19 serogroups and subdivided into more than 180 serovars on the basis of a microscopic agglutination test (23). Members of the serovar hardjo (belonging to the serogroup Sejroe) cause leptospirosis in dairy cattle, resulting in serious economic losses due to agalactia and abortion (11, 36). The infection can be transmitted by the urine to humans, resulting in dairy fever, characterized by headache, severe fever, meningitis, and icterus. In spite of the medical and economical importance of leptospirosis, very little is known about the virulence factors involved in pathogenesis. Hemolysins are involved in the pathogenesis of infections by Escherichia coli (16), Staphylococcus aureus (5), Listeria monocytogenes (7, 26), and Streptococcus pneumoniae (4) and have been claimed to be important in leptospiral infections (2, 35, 38, 39). Hemolysis by leptospirae is caused by phospholipases (38); both phospholipase A and sphingomyelinase C activities have been demonstrated (3, 6). Whereas pathogenic L. interrogans and nonpathogenic Leptospira biflexa strains have phospholipase A activity, sphingomyelinase C activity has only been demonstrated in strains of L. interrogans. A DNA fragment containing a sphingomyelinase C gene has been cloned from a Dutch field strain, causing bovine leptospirosis (10). In this study the molecular properties of the gene and its expression product are analyzed.

MATERIALS AND METHODS

Bacterial strains, plasmids, media, and transformation. L. interrogans strains were obtained from the World Health Organization/Food and Agricultural Organization Collaborating Center for Reference and Research on Leptospirosis at the Royal Tropical Institute in Amsterdam, The Netherlands, and were grown as described previously (10). Bluescript SK M13⁺ (pBS) was used as plasmid cloning vehicle, and all experiments with pBS, unless stated otherwise, were performed with XL1-Blue (Stratagene, La Jolla, Calif.) as the *E. coli* host strain. *E. coli* cells were grown in LuriaBertani (LB) medium or on LB agar plates (28) containing 100 μ g of ampicillin per ml. Competent *E. coli* cells were prepared by the CaCl₂ method (28). Plasmid pHL2-B3 has been described previously (10). Plasmid pHL2-B4 contains the same DNA fragment as pHL2-B3 in the opposite orientation.

Production of mutants and sequence analysis. DNA from clones pHL2-B3 and pHL2-B4 was digested with restriction enzymes *XbaI* and *SstI*, and unidirectional deletion mutants were produced by exonuclease III digestion with the Erase-a-Base kit (Promega Biotec, Madison, Wis.). Single-stranded DNA from these clones was prepared as described by the manufacturers of pBS (Stratagene). Nucleotide sequences were determined by the dideoxy-chain termination method of Sanger et al. (33). Nucleotide sequences were analyzed with the Beckman Microgenie (release 6.0; Beckman Instruments, Palo Alto, Calif.) and the PC/Gene (release 6.0; Genofit S.A., Geneva, Switzerland) computing programs.

The FASTA program, release 1.0, April 1988 (30), was used to compare nucleotide and amino acid sequences with the following data bases: EMBL (release 19.0), NBRF/PIR (release 21.0), NBRF/NEW (release 39.0), Swiss-prot (release 11.0), and Brookhaven (July 1989). Similar sequences were aligned by using the Clustal computer program (19, 20).

Hemolysin plate assay. Colonies of *E. coli* harboring recombinant plasmids were streaked on plates containing LB broth, 1% agar, 20% (vol/vol) fresh sheep erythrocytes (washed twice in 0.9% NaCl), 25 mM MgCl₂, and 100 μ g of ampicillin per ml. Cells were grown for 18 h at 37°C, and hemolytic zones appeared after an additional incubation of 24 h at room temperature. Omission of MgCl₂ from the agarplates resulted in much smaller hemolytic zones.

Phospholipase assay and analysis by thin-layer chromatography. Sphingomyelinase activity was tested in a biphasic system, essentially as described previously (24), consisting of an ether-methanol (9:1, vol/vol) organic phase containing 2 mg of sphingomyelin isolated from bovine brain (Sigma Chemical Co., St. Louis, Mo.) per ml and a water phase containing 10 mM Tris hydrochloride (pH 7.4), 25 mM MgCl₂, and (sonicated) bacteria and/or culture medium. The

^{*} Corresponding author.

INFECT. IMMUN.

| 10 | 20 | 30 | 40 | 50 | 60 | | 910 | 93 | 20 | 930 | 94 |) | 950 | 960 | 1 | .7 <u>50</u> | 1760 | | 1770 | 1780 | 1790 | 1800 |
|-----------------------|----------------------|----------------------|--------------------|---------------------|---------------|------------|---------|---------|--------------|---------|----------|------------|---------------|--------------|----------|--------------|---------|--------------|----------|------------|--------------|-------------|
| (5'OH) GGATOCTITITIOG | TIATTTTACT | AGAATTIGTIC | XCAAGGAOGI | GGAAACTIT | LAOGT | ACAM | ACGAAAG | AGCACA | ACGTAI | OGTGAGT | CCAATTA | ATOCA | MACCAN | GAOGTCAT | CGATOGA | GTTTUC | TUGTAT | CIGIA S V | GOCACIGG | AAAAAAGAT | ICAAGCIAA | LTTCTGAAAA |
| 70 | 80 | 90 | 100 | 110 | 120 | ų | NER | , n d | K 1 | v 3 | 5 N I | Τų | мų | D • 1 | 183 | • • | | | A 1 0 | | 4 n r | |
| TTTAGATGTTCGTG | GATAGAACTTT | ITTICCTCCTAAT | TTIGTICAL | CAAGTAGCT | GATC | | 970 | 9 | B0 | 990 | 100 |) | 1010 | 1020 | | | | | | | | |
| | | | | | | TGTGT | TOGATGA | AGCTTT | IGATAC | AGATOCA | GAAAAAT | CICIT | GACOGA | GTTOGTTC | 1 | 810 | 1820 |) | 1830 | 1840 | 1850 | 1860 |
| 130 | 140 | 150 | 160 | 170 | 180 | v | FDE | C A F | DT | DA | RKI | LL | DG | VRS | ATOGAAC | GCATGG | TIGAAGG | TGAAT | GCTACGAC | GGAAAOOGA | TTTAACAA | ATTCAATTT |
| TTIGTAAATCIGAT | TIGAGAACITT | TACATTACATA | CAAGTOCAT | AAAAOGOCAI | TATAT | | | | | | | | | | SN | A W | LK | V N | ATI | BTD | LTI | CFNL |
| 100 | 200 | 210 | | | ~~~ | | 1030 | 10 | 40 | 1050 | 106 |) | 1070 | 1080 | , | 970 | 1990 | | 1990 | 1900 | 1010 | 1920 |
| CONCASTICAACTICAATT | 200 | 210 | ZZU TTCAAATCA | 2.30 TCAAATATA1 | 240 | ALAAN P | | | D U | TC | | WALG: | NGGGAN V D | | GGTGCA | ACGAAT | GATOOOG | ATTO | AATCOCTC | CTICATIGAA | AAGOOGOGC | TGTOOGAT |
| | 10112111011 | | 1 I GENALICE | | | 5 | | | | 1 0 | | ~ 0 | | | ΥQ | TN | DP | DS | NPS | CMK | SGI | IVRI |
| 250 | 260 | 270 | 280 | 290 | 300 | | 1090 | 11 | 00 | 1110 | 112 |) | 1130 | 1140 | | | | , | | | | |
| TATTTTAATATAA | TATAGCTTATA | MAGOOGTTTA | TCTGCCCCG | TGTGGGATT | CAAA | TGGC | TGTACAG | AACOGA | IGOGTT | TAOGAAD | GACCAGT | GTTAT | IGTAAGT | AAATGGCC | 1 | .930 | 1940 | | 1950 | 1960 | 1970 | 1980 |
| | | | | | | G | LYB | R T D | AF | TN | GGV | V I | V S | KWP | OGAATCI | TCICAT | TCTTTA | ATTAC | TTTTGGA | TIGGIGGCT | COGREGATE | TAAAGGTAA |
| ATCATAAAACATTA | 320 CAAATOCTTTT | 330 Pataticacatac | 340 Yeattaticaa | 350 | 360 | | 1150 | 11 | 6 0 | 1170 | 110 | ` | 1100 | 1200 | 6 3 | эп | 5 1, | A 1 | | | | , K G U |
| ALCALASSIANI IA | | | | ATTICITUA | | CATTO | | | OU ACATCT | | CAAAAA | , 76766 | COCCAT | GTITTTC | 1 | 1990 | 2000 |) | 2010 | 2020 | 2030 | 2040 |
| 370 | 380 | 390 | 400 | 410 | 420 | I | EEB | CIO | HV | FK | EKG | CG | A D | VFS | CIACCO | TACTAT | COCAAAT | TCAAO | GACOGGTO | GAATOGAAT | CCAAATCA | TAAACTTAGA |
| TCTTGAAACATGAT | CTAAAAAGTAA | ACAAAGACTTTG | TTOGCTTT | AATTIGTTL | TTTT | • | | 1 | | | | | | | Y A | YY | PK | FN | DGS | ; N R I | QI | LNLD |
| | | | | | | | 1210 | 12 | 20 | 1230 | 124 | D | 1250 | 1260 | | | | | | | | |
| 430 | 440 | 450 | 460 | 470 | 480 | CAAC | AAGGATT | TIGCTTA | TGTGAG | GATOGAT | AAAAACGG | AGAAA | GITTCAT | ATCATTOG | 2 | 2050 | 2060 |) | 2070 | 2080 | 2090 | 2100 |
| TATACATTTAAAAT | TTIGAAGGTIG | CTTTAGGAATTG | GCATTITIA | ATTICAATT | CTTG | N | KGE | FAY | VR | ID | KNG | RK | FH | IIG | COGAGOG | FIGITIG | CAAGAC | UGAGI | UCACIUX | ATTCAAAGA | TEATGATA | TATTICANG |
| 400 | F00 | E 10 | E20 | E20 | E/ 0 | | | | ~~ | | | | 1010 | 1000 | 66 | СĽ | ųр | 63 | K V P | . F K D | , I D 1 | LISK |
| AGOTTATATGAAA | TICATIONITY | DIU TATAAATATA | 320 CTTC:ITC:IT | DBU CTTICAACTICT | J4U TATICT | 4407 | 1270 | | ACATTC | 1290 | | | 1310 | ACACTICAA | 2 | 2110 | 2120 | , , | 2130 | 2140 | 2150 | 2160 |
| | | | | -35 | | T | HVO | | DS | GC | ANL | GV | VS | RVN | MGACA | TATTT | CICACIO | TTIGG | GAAGGTGG | AAACTOOGA | TAAATATC | CTATCTTIG |
| | | | | | | - | | | | | | | | | RQ | YF | LT | V W | EGG | ; N W D | • K Y J | LYLW |
| 550 | 560 | 570 | 580 | 590 | 600 | | 1330 | 13 | 40 | 1350 | 136 | 0 | 1370 | 1380 | | | | | | | | |
| TATAGTATOGGCAT | TGAT <u>TTAAAA</u> A | CAATOGAGA TAG | TGTATGAGA | ATAAAAAAA | TATAC | TCAA | TCAAOGA | AGATCAG | GGATTI | TATOGAT | TOCAAAAA | ATCCO | GAAAAAO | GAAATGGT | 2 | 2170 | 2180 |) | 2190 | 2200 | 2210 | 2220 |
| | -10 | \$.D. | MR | IKK | Y T | Q | FNE | EIR | DF | ID | SKK | IP | K N | EMV | GAGAAG | CATATO | OCTCTA | COCAA | ATTTTT | TCTCAAGCT | GGATTOCT | TOOOGAAAT |
| | | | (star | : OR F 1) | | | | | | IR | 1 | | | | ĸs | нт | GL | KE | 1 6 1 | | | SPEM |
| 610 | 620 | 630 | 640 | 650 | 660 | | 1390 | 14 | 00 | 1410 | 142 | 0 | 1430 | 1660 | : | 2230 | 2240 |) | 2250 | 2260 | 2270 | 2280 |
| AAAAGTGAGACTTC | TIGTAAATIGT | GICICITATIA | TTTTTTCTA | ATAGATIGIC | GACC | TCTG | ATOGOGG | GAGATTI | GAAOGT | AATCAAA | OGTACTAG | AGAATA | TCATCAG | AIGCITIG | GAATTO | GAGTAAG | AAATTA | TTEAT | OCTICACO | CTTAAAAAAT | CAAATGAA | AGACOGCAGG |
| KVRL | LVNC | CLLL | FFL | IDC | GΑ | L | IAC | GDL | . N V | IK | GSR | EY | НQ | MLC | NW | S K | KL | IY | R | | | |
| | | | | | | | | | | | | _ | | | | | | end O | RF1) | | | |
| 6/0 | 680 | 690 | 700 | 710 | 720 | | 1450 | 14 | 60 | 1470 | 148 | 0 | 1490 | 1500 | | 2200 | 2200 | | 2210 | 2220 | 2220 | 2260 |
| DROS | I. Y. K. D. | | LTY | TSD | NR | | T N T | UNANLAG | | | C V F | 1111AL | U V D | TET | GATICTIC | 10000 | CITTIO | , XXXTAC | ATAGAAT | TTTATATO | ATOCAGGA | CTTTIGOGATA |
| | | 2 2 4 0 | | | N K | • | 5 | | | ••• | • • • | • • | | | IR4 | | | | | | Ľ | 5 |
| 730 | 740 | 750 | 760 | 770 | 780 | | 1510 | 15 | 20 | 1530 | 154 | 0 | 1550 | 1560 | | | | | | | | |
| AAACATTGGTTCTA | CAAATTCTGAT | TAACGGGTTCI | OGCTCICIA | AGTICIAGIO | XXXXX | CAAT | GAAATIG | COCCTTT | TIATI/ | TAAAAAA | GTOGAGOC | OGCATA | CTTGGAT | TACATATT | | 2350 | 236 |) | 2370 | 2380 | 2390 | 2400 |
| NIGS | TNSD | LTGS | GSV | SSS | ΡA | N | EIA | A A F | YY | . K K | VEF | AY | LD | YIF | GATATA | ATGACGC | TCTAAT | CTAA | TTTATTA | TOGGAOGTA | JGAAAAAA | TTOCATACAG |
| | | | | | | | | | | | | | | | | | | | | | | |
| /90 | 800 | STO STO | 820 | 830 | 840 | ~ | 1570 | 15 | 80 | 1590 | 160 | 0 | 1610 | TPSC | | | | | | | | |
| | RNST | L A N S | TPF | NMC | | USTT V | CAABI | | | | U O X | | V D | DTS | | | | | | | | |
| 2 | | | | | | • | | | | • • | | | | | | | | | | | | |
| 850 | 860 | 870 | 880 | 890 | 900 | | 1630 | 16 | 540 | 1650 | 166 | 0 | 1670 | 1680 | | | | | | | | |
| AATTCTAACTCACA | ACGITITICCIG | TIGOOGAAAAOG | CTTOCAGGI | TGGGGAAATT | GOOG | TGCA | AAAACTT | GGACCCC | AAAGG | ATATAC | ACTGATG | ATTTTC | OGATCA | TATCOOFT | | | | | | | | |
| ILTH | NVFL | LPKT | LPG | WGN | WG | A | KT | WTA | ĸ | 5 Y T | SDI | FS | DH | YPV | | | | | | | | |
| | | | | | | | 1690 | 17 | 700 | 1710 | 172 | 00 | 1730 | 1740 | | | | | | | | |
| | | | | | | TTAC | GGTTTTA | TCTATCO | OGATIC | ATOCACI | COGACGA | GTOOOG | ACGTAA | ACGAAATA | | | | | | | | |
| | | | | | | Y | GF | IYA | DS | SST | PTI | SG | RK | RKY | | | | | | | | |
| | | | | | | | | | | IR2 | | | | | | | | | _ | _ | | |

FIG. 1. Nucleotide and derived amino acid sequences from the 3,987-bp BamHI insert in pHL2-B3. The putative signal sequence is boxed, and -10 and -35 transcription signals and putative ribosome-binding sites (RBS) are underlined. Inverted repeats (IRs) are numbered and overlined by arrows. Gibbs free energy differences were calculated by using the Microgenie program and are indicated in kilocalories (1 kcal is equal to ca. 4.184 kJ) per mole: IR1, -9.6; IR2, -8.8; IR3, -9.6; IR4, -25.4; IR5, -10.8; IR6, -8.0. Sequences similar to the potential regulatory sequence in L. biflexa (42) are overlined with a boldface bar. These data have been submitted to the EMBL Data Library under accession no. X52176.

samples were vigorously shaken for 4 h at 37°C, and 2 µl of the organic phase was applied on a silica gel-60-coated glass plate (E. Merck AG, Darmstadt, Federal Republic of Germany). When MgCl₂ was omitted from the reaction mixture, sphingomyelinase activity was much lower. The chromatogram was developed with a chloroform-methanol-water-25% ammonia (58:35:3.5:3.5, vol/vol) mixture as the mobile phase. (Phospho)lipids were visualized by spraying the plates with 30% sulfuric acid, followed by heating at 110°C for 5 min. Purified sphingomyelinase C (0.08 U) from S. aureus (Sigma) was used as a positive control. Sonicated E. coli cells containing pBS were used as a negative control. Degradation of other phospholipids was performed as described above, chromatograms were developed with mobilephase mixtures as previously described (24, 25), and reaction products were visualized as described above. Degradation of L-a-lysophosphatidylcholin (egg yolk; Sigma), L-a-phosphatidylethanolamine (bovine brain; U.S. Biochemical Corp., Cleveland, Ohio), phosphatidyl-L-serine (bovine brain; U.S. Biochemical), L-a-phosphatidic acid (egg yolk; U.S. Biochemical), and L- α -lecithin (U.S. Biochemical) as substrates was tested.

RESULTS

Nucleotide sequence analysis. The nucleotide sequence of a cloned 3,987-base-pair (bp) BamHI DNA fragment, which was shown to code for sphingomyelinase activity (10), was determined (Fig. 1; data submitted to EMBL Data Library under accession no. X52176). For this purpose, a number of deletion clones (Fig. 2C) were produced. After translation of the nucleotide sequence, three open reading frames (ORFs) were identified (Fig. 2A). The percentage of G+C in the whole fragment (35.9%) as well as in the individual ORFs (Fig. 2A) was within the range of 34.1 to 39.1%, which corresponds to the known G+C percentage of genomic DNA from L. interrogans (23) and from L. interrogans serovar hardjo strains (27).

Upstream of ORF1 and ORF2, putative -10 and -35 transcription and Shine-Dalgarno translation-initiation sequences similar to those in *E. coli* were identified. In ORF2 a second putative translation initiation site was identified at bp 3264 (Fig. 1). Sequences upstream of ORF1 and ORF2 and the internal transcription initiation site in ORF2 were similar to the previously reported potential regulatory se-

3670

3730

3790

3850

3910

3070

3690

3740

3800

3860

3920

TCTATCAACTCTTCCTTTTTGGGGATCC (3'OH)

AGATAGTTGAGAAGGAAAACCCCTAGG (5°CH)

ILEEKQPD (08273)

3610 3620 3630 3640 3650 3660 TOCTGTTTTTCTCTATCTOCGAGAGTTTTATAGGTTCAGAATTTCCAAGTTCATTTGAA

AGGACAAAAAGAGATAGAGGCTCTCAAAATATCCAAGTCTTAAAGGTTCAACTAAACTT E O K K E I E S L K I P E S N G L O N S

ATATTTTCTAGATCTTTGAGATACAACTCATTCAACGGAGCTTGTTTTTTTGGGAGAGG

TICTTIAIGATTIGTIGTIGAATCACTTITTAGAATATTICAAATTTICTIGTIGGIGGCCTTTT

AAGAAATACTAAACAACTTAGAGAAAATCTTATAAGTTTAAAGAACACACGGGAAAAA N K I I Q Q Q I E K L I N L N E Q A S K

GCATCAAGTOGATCTTCTTTCAAAGTCTTAAATAAACTTTTTTTCCOGACTTCCCCATAGA

OGTAGTTCAGCTAGAAGAAAGTTTCAGAATTTATTTGAAAAAAGGGCTGAAGGGCTATCT

TTTAGAGGTTTTATTTCAAAATAAGAATTAGAGATATCTTCCGAAGAAGTGTCCACCCAT

AAATCTCCAAAATAAAGTTTTATTCTTAATCTCTATAGAAGGCTTCTTCACAGGTGGGTA

N L P K I E F Y S N S I D E S S T D V W

CTTTTTCCAATTOGGTCTTTTGCCCGAAGTTTGAACCOGTTTGTTTATTCAAAGAAT

V K E L E T K Q G F N S G T Q K N L S H

3930

A D L R D E K L T K F L S K E R S G S L

3870

3700

3760

3820

3880

3940

3710

3770

3830

3950

3720

3780

3960

3900

3960

3600

INELDKLYLENLPAOKKOS

3750

3810

| ~ | 24. | 10 | | 24 | 20 | | 2 | 430 |) | | 244 | 0 | | 24 | 50 | | - 2 | 460 |
|--|---|---|--|--|---|--|---|---|--|---|---|--|--|---|---|--|---|--|
| Ganu | AGT | GL | | GAA | GOC | TIC | GGI | TTG | ATC | TT/ | CAG | CTI | CAA | ∞ | TOC | GOC | GAI | CAT |
| | | | | | | | | (| 3'0 | E) | GTC | GAA | GTI | GGG | ACC | CCC | CT/ | GTA |
| | | | | | | | | | | | L | ĸ | L | G | G | G | I | М |
| | | | | | | | | | | | (* | nd (| DR.F2 |) | | | | |
| | | | | | | | | | | | | | | | | | | |
| | 24 | /0 | | 24 | 80 | | 2 | 490 | | | 250 | 0 | | 25 | 10 | | 2 | 520 |
| ~~~ | | STIC | GII | | AIG | A14 | | TIG | 161 | ~ | | AA.) | | ATU | 71G1 | TTI | TU. | GIA |
| Trici | AGO | <u>.</u> | | ITT . | TAL | TAI | | ANC A | | GU | uu. | TL | ATT | TAG | ACA | ~~ | MGC. | CAT |
| F | P | E | N | F | н | Ŷ | S | Q | т | ۸ | G | I | L | D | т | ĸ | R | Y |
| | 25 | 30 | | 25 | 40 | | 2 | 2550 |) | | 256 | ю | | 25 | 570 | | 2 | 2580 |
| TTGT/ | GAC | TAT | ATT | | | GLA | AGA | TTC | TAG | TR | XCA | TIC | AAT | ccc | CGI | TO | CAC | |
| AACAI | ста | ATA | TAA | TTI | TTI | CAT | TCI | AAG | ATC | *** | IGGT | AAC | TTA | ۵GG | GCA | AGG | GTG | TTT |
| QI | | I | N | F | F | Y | S | E | L | Е | W | Q | I | G | Т | G | V | F |
| | | | | | | | | | | | | , | | | | | | |
| | 25 | 90 | | 26 | 00 | | 2 | 610 | • | | 262 | 0 | | 26 | 30 | | 2 | 640 |
| AAGG/ | COG: | TTIG | TAC | TTI | GGA | ATG | TAC | TOC | ACI | TIC | AAT | ACT | TTI | TOO | GAC | TOC | CAA | GOC |
| TTCCT | GCC | ww. | ATG | | CCT | TAC | ATG | ACC | TGA | *** | TTA | TC/ | | AGG | CIG | ACC | GT | COGG |
| L V | Т | Q | V | ĸ | S | H | V | G | S | E | I | S | K | G | V | G | L | G |
| | 26 | 50 | | 26 | 60 | | | 670 | | | 269 | 0 | | 26 | - | | | 700 |
| TACT | CAC! | | <u></u> | | | - | <u> </u> | 1070 | | - | 200 | | | 20 | 190 MAT | - | ~ | .700 |
| | CTC. | | ~ | | | - | | | | | | | | | | 100 | ~ | |
| v | | 2 1 1 | נשט. ער | | - 10 T | uUI e | ندين. ج | - C | 10 | | | undi V | דיטי. ע | | 1010 M | نانية: ح | ы. Б | J |
| • 4 | | £ | | r | Ŧ | 5 | ъ | э | ъ | r | r | I | I | x | m | G | P | r. |
| | 27 | | | | | | | | | | | | | | | | | |
| | - 21 | 10 | | 27 | 20 | | 2 | 730 | | | 274 | 0 | | 27 | 50 | | 2 | 760 |
| AAGTG | | LO SAGT | TIC | 27 TAT | 20 20 | GAA | 2 111 | 730 ATC | GTI | 000 | 274 | 0 TTI | ATC | 27 | 50 ATT | *** | 2 AT1 | .760 TCC |
| AAGTO | 273 2273 | LO LAGT CTCA | TTC | 27 TAT | 20 ACC TGG | GAA | 2 .TTT | 730 ATC | GTI | 000 | 274 711G | 0 | ATC | 27 AAA | '50 ATT TAA | | 2 AT1 TA4 | .760 TCC |
| AGTO FTCAC | | LO SAGT CTCA T | TTC AAG E | 27 TAT ATA | 20 ACC IGG G | GAA CTI F | 2 ITT AAA K | 730 ATC TAG D | GTI CAA N | 000 | 274 711G | O TTI AAA | ATC TAG | 27 AAA TTT F | 50 ATT TAA N | AAA TTI F | 2 ATI TAA N | 2760 TCC ACC |
| AAGTO TTCAC L S | | LO EAGT CTCA T | TTC AAG E | 27 TAT ATA | 20 20 20 20 20 20 20 20 20 20 20 20 20 2 | GAA CTI F | 2 IIII AAA K | 730 ATC TAG D | GTI CAA N | 000 600 6 | 274 711G AAC Q | 0 TTI AAA K | ATC TAG D | 27 AAA TTT F | '50 ATT TAA N | TTT F | 2 ATI TAA N | 760 TCC AGG G |
| AAGTO TTCAC L S | 273 | LO EAGT CTCA T 70 | TTC AAG E | 27 TAT ATA I 27 | 20 20 20 20 20 20 30 80 | GAA CTT F | 2 ITT XAAA K 2 | 730 ATC TAG D | GTI CAA N | 000 600 6 | 274 711G AAC Q 280 | 0 TTT AAA K | ATC TAG D | 27 AAA TTT F 28 | 50 ATT TAA N | TTI F | 2 ATT TAA N | 2760 TCC AGG G 820 |
| AAGTO TTCAC L S | 273 71110 5 L 277 | LO AGT TCA T 70 | TTC AAG E | 27 TAT ATA I 27 ATT | 20 ACC IGG G 80 TAT | GAA CTT F | 2 IIII AAA K 2 AAI | 730 ATC TAG D 790 | GTI CAA N | G | 274 TTG AAC Q 280 | 0 TTI AAA K 0 TIG | ATC TAG D | 27 AAA TTT F 28 | 50 ATT TAA N 10 | TTI F | 2 ATT TAA N 2 AGA | 2760 TCC AGG G 820 GTT |
| AAGTO TTCAC L S | 273 71110 5 L 277 ACAC | LO AGT T T 70 | TTC AAG E TAA | 27 TAT ATA 1 27 ATT TAA | 20 ACC TGG G 80 TAT ATA | GAA CTI F GGA | 2 1111 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 | 730 ATO TAG D 790 GCT OGA | GTI CAA N GGA | G | 274 TTG AAC Q 280 OGT | O TTT AAA K O TTG | ATC TAG D GTA | 27 AAA TTT F 28 TTC AAG | 50 ATT TAA N 10 GOG | GGI | 2 ATI TAA N 2 AGA | 2760 TCC AGG G 820 GTT CAA |
| AAGTG ITCAC L S ITCAA AAGTI E F | 273 TTTC 277 ACAC | LO AGT T T 70 AAA F | TAA E TAA ATT L | 27 TAT ATA I 27 ATT TAA N | 20 ACC TGG G 80 TAT ATA I | GAA CTI F GGA | 2 AAA K 2 AAT TTA I | 730 ATC TAG D 790 GCT CGA S | GTI CAA N GGA CCT | G G TAA ATT | 274 711G AAC Q 280 0G1 GCA T | O TTT K O TTG AAC Q | ATC TAG D GTA CAT | 27 AAA TTT F 28 TTC AAG E | 50 ATT TAA N 10 GOO P | GGT GGT CCA | 2 ATI TAA N 2 AGA ICI S | 760 TOC AGG G 820 GTT CAA N |
| AAGTG L S TTCAA TTCAA AGTT E F | 277 TTTC 277 ACAC | LO SAGT T T 70 SAAA F | TAA E TAA ATT L | 27 TAT ATA I 27 ATT TAA N | 20 ACC TGG 80 TAT ATA I | GAA F GGA CCT S | 2 ITI XAAA K 2 AAT TIA I | 730 ATO TAG D 790 GCT CGA S | GTI CAA N GGA CCI S | G G TAA L | 274 711G AAC Q 280 0G1 GCA T | 0 TTT AAA K O TTG AAC Q | ATC TAG D GTA CAT | 27 AAA TTT F 28 TTC AAG E | 50 ATT TAA N 10 GCC P | F GGI CCA T | 2 ATT TAA N 2 AGA ICT S | 760 TOC AGG G 820 GTT CAA N |
| AAGTG TTCAC L S TTCAA AAGTI E F | 273 71110 277 ACAC 71GT0 71L 283 | LO AGT TCA T 70 AAAA F 80 | TAA E TAA ATT L | 27 TAT ATA I 27 ATT TAA N 28 | 20 ACC G 80 TAT ATA I 40 | GAA CTI F GGA CCT S | 2 IIII AAA K 2 AAI II 2 AAI | 730 ATC TAG D 790 GCT CGA S 850 | GTT CAA N GGA CCT S | G G TAM ATT L | 274 TTG Q 280 OGT GCA T 286 | 0 TTT AAA K 0 TTG AAC Q 0 | ATC TAG D GTA CAT Y | 27 AAA TTT F 28 TTC AAG E 28 28 | 50 ATT TAA N 10 0000 P 70 | GGT GGT T | 2 ATT TAA N 2 AGA TCT S 2 AGA | 2760 TOC AGG G 2820 GTT CAA N 880 |
| AAGTO TTCAC L S TTCAA AAGTI E F | 277 AAAC TTTTC 277 ACAC TGTC L 283 TAGA | IO MGTI TCA T 70 MAAA TTTI F 80 MGTTI | TAA E TAA ATT L | 27 TAT ATA I 27 ATT TAA N 28 AAG | 20 ACC IGG G 80 TAT ATA I 40 AAT | GAA CTI F GGA CCT S | 2 1111 AAA K 2 AAT TTA I 2 AAC | 730 ATC TAG D 790 GCT CGA S 850 AAG | GTI CAA N GGA CCT S | G G TAM ATT L | 274 711G AAC Q 280 031 336 AAT | 0 TTI AAA 0 TTIG AAC Q 0 CTTI | ATC TAG D GTA CAT Y | 27 AAA TTT F 28 TTC AAG E 28 AATO | 50 ATT TAA N 10 GGG P 70 GGG | GGI GGI COA T | 2 ATT TAA N 2 AGA TCT S 2 AGA | 2760 TCC AGG G 820 GTT CAA N 880 ATA |
| AAGTO ITCAC L S ITCAA AAGTI E F | 273 AAAC TITIC 277 ACAC TIGIC 283 TAGA ATCI | IO AGT T 70 AAAA TTT F 80 AGTT CAA | TIC AAG E TAA AIT L CAT | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC | 20 ACC IGG 6 80 TATA 1 40 AAT | GAA CTI F GGA CCI S TCC | 2 ITTI AAA K 2 AAT TTA I 2 AAC TTIG | 730 ATC TAG D 790 GCT CGA S 850 AAG | GGA N GGA COTT S | G G TAM ATT L | 274 TTG Q 280 OGT GCA T 286 AAT | 0 TTI AAA K 0 TTIC AAC Q 0 CTTI GAA | ATC TAG D GTA CAT Y GOG COGC | 27 AAA TTT F 28 TTC AAG E 28 ATO TAG | 50 ATT TAA N 10 0000 P 70 0000 P 70 | GGT T | 2 ATT TAA N 2 AGA TCT S 2 AGA TCT | 2760 TCC AGG G 820 GTT CAA N 880 ATA TAT |
| AAGTG L S TTCAA AAGTI E F AAAAT F N | 277 AAAC TITIC 277 ACAC TIGIC L 283 TIAGA TIGIC I S | IO SAGT TCA T 70 SAAA TTT F SO SCAA N | TIC AAG E IAA AIT L CAI GIA | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L | 20 ACC TGG 6 80 TATA 1 40 AAT TTA I | GAA CTI F GGA CCT S TOC AGG G | 2 TTT AAA K 2 AAT TTA I 2 AAC TTIG V | 730 ATC TAG D 790 GCT AGA S 850 AAG TTC L | GGA N GGA COT S AGG TOC P | G G G TAM ATT L AGA | 274 TTG Q 280 OGT T 286 AAT TTA I | 0 TTI AAA K O TTIG AAC Q O CTTI GAA K | ATC TAG D GTA CAT Y GOG CGC R | 27 AAA TTT F 28 TTC AAG E 28 ATC IAG D | 50 ATT TAA N 10 GGC P 70 GGC P 70 GGC P | AAA ITT F GGI T T AAA | 2 ATT TAA N 2 AGA TCT S | 760 TCC AGG G 820 GTT CAA N 880 ATA TAT Y |
| AAGTO L S TTCAA AAGTTI E F AAAATI F N | 277 AAAC 277 ACAC 7GTC 283 TAGA ATCT 5 285 | IO AGT TCA T CAAAA CTTT F CAAA N CAAA N CAAA N | TIC AAG E IAA AIT L GIA M | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L 29 | 20 ACC G 80 TATA ATA I 40 AAT TTA I 00 | GAA CTI F GGA CCT S CCT S CCT G AGG G | 2 ITTI XAA K 2 AAT TTA I 2 AAC TTIG V 2 | 730 ATC TAG D 790 GCT CGA S 850 AAG TTC L 910 | GGA N GGA CCT S AGG TOC P | G G G TAM ATT L AGA | 274 TTG Q 280 OGT T 286 AAT TTA I 292 | 0 TTI AAA K 0 TTI Q 0 CTTI GAA K 0 | ATC TAG D GTA CAT Y GOG CAT R | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 | 50 ATT TAA N 10 GCC P 70 GCC P 30 | GGI GGI CCA T AAA | 2 ATT TAM N 2 AGA TCT S 2 AGA TCT S 2 2 2 | 2760 TOC AGG G 820 GTT CAA N 880 ATA TAT Y 940 |
| AAGTO L S TTCAA AAGTI E F AAAAT F N GAGT | 277 AAAA 2777 ACAC TGTC 283 TAGA ATCI 5 285 TTAGA | IO AGT TCA T CAAA CTTT F CAA N O CTTC | TIC AAG E IAA AIT L GIA M GAT | 27 TAI ATA 27 ATI TAA N 28 AAG L 29 TOO | 20 ACC G 80 TAT ATA I 40 AAT TTA I 00 GAA | GAA CTT F GGA CCT S CCT S TCC AGG G TAA | 2 IIII AAA K 2 AAI I 2 AAI TIG V 2 AAA | 730 ATC TAG D 790 GCT CGA S 850 AAG TTC L 910 GCT | GGA N GGA COT S AGG TOC P GAG | COCC GGG G TAM TCT S AGA | 274 TTG Q 280 OGT T 286 AAT TTA I 292 CAA | 0 TTI AAA K 0 TTG AAC Q 0 CTTI GAA K 0 ATT | ATC TAG D GTA CAT Y GOG CAT R GGG R | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 ACG | 50 ATT TAA N 10 COGG GOC P 70 COGG GOC P 30 ATA | AAA F GGI CCA T T AAA F TTO | 2 ATT TAA N 2 AGA TCT S 2 AGA TCT S 2 GOC | 2760 TOC AGG G 820 GTT CAA N 880 ATA TAT Y 940 ATT |
| AAGTG L S FTCAA AAGTT E F AAAATT F N AGAGT ICTCA | 277 AAAC TTTIC 277 ACAC TGTIC 283 TACA ATCI S 289 TTACA | LO AGI TCA T VO AAAA CTTI F SO AGTI CAA N O O TTO AAG | TIC AAG E IAA AIT L GIA M GAT CIA | 27 TAT ATA I 27 ATT ATA N 28 AAG TTC L 29 TOO AGG | 20 ACC G S0 TATA I 40 AAT I GAA CTT | GAA CTT F GGA CTT S TOC AGG G TAA | 2 IIII AAA K 2 AAI I 2 AAI V 2 AAA III V 2 AAA | 730 ATC TAG D 790 GCT CGA 850 AAG TTC L 910 GCT CGA | GTI CAA N GGA COT S AGG TOC P GAG CTC | GGG G TAM ATT L AGA TCI S ATT | 274 TTG Q 280 OGT T 286 AAT TTA I 292 CAA GTT | 0 TTI AAA K 0 TTG AAC Q 0 CTI GAA K 0 ATTI | ATC TAG D GTA CAT Y GOG R GTA CAT | 27 AAA TTT F 28 AG E 28 ATO D 29 AG TGC | 50 ATT TAA N 10 0000 P 70 0000 P 70 0000 P 30 ATA TAT | AAA TTT F GGT T AAAA TTT F | 2 ATI TAM N 2 AGA TCT S 2 AGA TCT S 2 QCC QCC QCC | 2760 TOC AGG G 820 GTT CAA N 880 ATA Y 940 ATT TAA |
| AAGTCAL L S TTCAL TTCAL E F AAAAAT F N AGAGTI TTTTA F N AGAGTI S N | 277 AAAA TITIC L 277 ACAA TIGTO L 2883 TIAGA AATCI S 2885 TIAGA | LO AGI T 70 AAAA TTT F 80 AGTT CAAA N 80 FTTO CAAG E | TIC AAG E IAA AIT L CAE GIA GIA I | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L 29 TOO G | 20 ACC G S0 TATA A A A A A A A TTA I O O GAA F | GAA CTI F GGA CCT S CCT S CCT S CCT S CCT S CCT S CCT I S CCTI I S C CCTI I S C CCTI S C C C C C C C S C C C C C C C C C C | 2 IIII AAA K 2 AAT I 2 AAC TIG V 2 AAA TIG V 2 AAA | 730 ATC TAG D 790 GCT CGA S 850 AAG TTC L 910 GCT S | GGA N GGA CCT P GAG CTC L | COCC GC GC TAM ATI L AGA TCI S ATI | 274 TTG Q 280 007 T 286 AAT TTA I 292 CAA CAT L | 0 TTI AAA K O TTI Q O CTTI GAA K O ATTI AATI | ATC TAG D GTA CAT Y GOG R GTA CAT Y | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 ACG R C R | 50 ATT TAA N 10 0000 P 70 0000 P 70 0000 P 30 ATA TAT Y | AAA ITTI F GGT CCA T AAAA TTTI F TTCO AAG E | 2 ATI TAA N 2 AGA TCI S 2 AGA TCI S 2 GOC GOC G | 760 TCC AGG G 820 GTT CAA N 880 ATA TAT Y 940 ATT TAA N |
| AAGTCI L S TTCAA AAAGTI E F AAAAATI TTTTA F N AGAGTI ICTCA S N | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | LO AGI T 70 AAAA TT F 80 AGTT KAAA N 80 AGTT KAAG E | TIC AAG E TAA AIT L GTA GTA M GAT CTA I | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L 29 TCO AGG G | 20 ACC G B0 TATA I 40 AAT I I GAA CTE F | GAA CTI F GGA CCT S CCT S TCC AGG G TAA | 2 IIII AAA K 2 AAT I AAT I I AAT I I AAT I I AAT I I AAT I I AAT I I AAT I I AAT I I AAT I I A AAT I I I A A A A A A I I A A A A A I I A A A A A A A A A A A A A | 730 ATC TAG D 790 CT CGA S 850 AAG S TTC L 910 GCT CGA S | GTT CAA N GGA CCT S AGG TOC P GAG CTC L | COC GGG TAM ATI L AGA TCI S ATI | 274 TIG Q 280 OGT T 286 AAT TIA I 292 CAA GTT L | 0 TTI AAA K O TTI Q AAC Q O CTTI GAA K O ATTI TAA N | ATC TAG D GTA CAT Y GOG CAT Y CAT Y | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 ACG R R C R | 50 ATT TAA N 10 CGG GCC P 70 CGG GCC P 70 CGG C P 30 ATA TAT | AAA ITTI F GGI COA T T TTO AAA E | 2 ATT TAM N 2 AGA TCT S 2 AGA TCT S 2 COC G C G C C | 760 TCC AGG G 820 GTT CAA N 880 ATA TAT Y 940 ATT TAA N |
| AAGTC TTCAA AAGTTI E F AAAAAT TTTTAA F N AGAGT TCTCA S N | 277 AAAA TTTI E 277 ACAC TGTC 283 TAGA ATCI S 285 TTAGA AATCI S 285 TTAGA | LO SAGI T T SAAA TTT SAAA STTT SAAA N SO STTC SAAG E SO | TIC AAG E TAA AIT L CAT GTA M GAT CTA I | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L 29 TCO G G 29 | 20 ACC G B0 TATA I 40 AAT I GAA CTTE F 60 | GAA CTI F GGA CCT S TOC AGG G TAA ATT L | 2 IIII AAA K 2 AAT I 2 AAC II V 2 AAA II V 2 AAA II V 2 AAA II V 2 AAA II V 2 AAA II V 2 AAA II V AAA I V AAAA I V AAAA I V AAAA I V AAAA I V AAAA I V AAAA I V AAAA I V AAAA I V AAAA I I V AAAA I I V AAAA I I V AAAA I I I I I I I I I I I I I | 730 ATC TAG D 790 CT C C AAG S 850 AAG C TTC L 910 GCT C GA S 970 | GTI CAA N GGA CCT S AGG TOC P GAG CTC L | GGG G TAM ATI L AGA TCT S ATT | 274 TTG Q 280 0GT T 286 AAT TTA I 292 CAA GTT L 298 | 0 TTI AAA K O TIG AAC Q O CTTI GAA K O ATTI TAA N O | ATC TAG D GTA CAT Y GOG CAT Y GTA CAT Y | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 ACG R 29 ACG R 29 ACG R 29 | 50 ATT TAA N 10 CGG GCC P 70 CGG F 70 CGG F 70 CGG F 30 ATA TAT Y 90 | AAA ITTI F GGI COA T TTO AAA E | 2 ATT TAA N 2 AGA TCT S 2 AGA TCT S 2 COC GCC GCC GCC GCC GCC GCC GCC GCC GCC | 760 TCC G B20 GTT CAA N B80 ATA TAT Y 940 ATT TAA N 0000 |
| AAGTC TTCAA AAGTT E F AAAATT TTTTA F N AGAGT S N | 277 AAAA TTTTC TTTC TTTC TTTC TTTC TTTC T | LO SAGI T 70 SAAA 70 70 70 70 70 70 70 70 70 70 70 70 70 | TAA E TAA ATT L CAT GTA M GAT CTA I TTT | 27 TATA ATA I 27 ATT TAA N 28 AAG TTC L 29 TOO GGT GGT | 20 ACC IGG G 80 TATA AAT I 40 AAT I GAA CTE F 60 AAT | GAA CTI F GGA CCT S TCC AGG G TAA ATT L TAC | 2 ITTI AAA K 2 AAT TTA I 2 AAT TTG V 2 AAA TTTF F 2 AGT | 730 ATC TAG D 790 CT C C A A S 850 A A G C T T C C A S 910 G C T T C C A S 910 A TAC S 850 A A C T T C S 850 A T C C A C C C A C C C C A C C C A C C C C A C C C C C A C C C C A C C C C C A C C C C C A C C C C C C A C C C C A C C C C C C C C C C C C A C C C C C C C C C A C | GTI CAA N GGA COT S AGG TOC P GAG CTC L | COC GGG G TAM ATI L AGA TCI S ATI TAM | 274 TIG Q 280 OGT T 286 AAT TTA I 292 CAA GTT L 298 CTC | 0 TTI AAA K 0 TTG AAC Q 0 CTTI GAA K 0 ATTI TAA N 0 AATT | ATC TAG D GTA CAT Y GOG CR GTA CAT Y | 27 AAA TTT F 28 TTC AAG E 28 ATC D 29 AGG R 29 AGG | 50 ATT TAA N 10 GCC P 70 GCC P 70 GCC P 30 ATA TAT Y 90 GTC | AAA TTT F GGT T AAAA TTT F TTO AAAG E TAT | 2 ATT TAA N 2 AGA TCT S 2 AGA GOC GGG G G 3 ATA | 760 TCC AGG G 320 GTT CAA N 880 ATA TAT Y 940 ATT TAA N 000 GAT |
| AAGTC TTCAA L S TTCAA AAGTI E F AAAATI TTTTAA F N AGAGTI TTTTCA S N AGAGTI | 277 AAAA TTTTC TTTTC TTTTC TTTTC TTTC TTT | LO SAGI T T T T T T T T T T T T T T T T T T T | TAA E TAA ATT L CAT GTA M GAT CTA I TTT | 27 TAT ATA I 27 ATT TAA N 28 AAG TTC L 29 TOO G G G G G G G G G G G G G G G G G G | 20 ACC IGG 6 80 TATA AAA I 40 AAAT I 60 AATT F 60 AATT | GAA CTT F GGA CCT S CCT S TCC AGG G TAA ATT L TAC | 2 ITT AAA K 2 AAT TTA I 2 AAT TTG V 2 AAT TTG V 2 AAT | 730 ATC TAG D 790 GCT CGA S 850 AAG TTC L 910 GCT CGA S 970 ATA TAT | GGA N GGA CCT S GAG CTC L TCT AGA | GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG | 274 TTG Q 280 OGT T 286 AAT TTA I 292 CAA GTT L 296 CTC GAG | 0 TTI AAA K O TTIG AAC Q O CTTI GAA K O AATI TAA N O AATI | ATC TAG D GTA CAT Y GOG CAT Y ATA | 27 AAA TTT F 28 ATC D 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 29 ACC R 20 ACC ACC R 20 ACC R 20 ACC R 20 ACC R 20 ACC ACC ACC ACC ACC ACC ACC ACC ACC AC | 50 ATT TAA N 10 COGC P 70 COGC P 70 COGC P 70 COGC P 30 ATAT Y 90 CAG | AAA TTT F CCA T TTT F TTC AAA E TTT F | 2 ATT TAA N 2 AGA TCT S 2 AGA TCT S 2 GOO G G G G G G G G G G G G G G G G G | 760 TCC AGG G 820 GTT CAA N 880 ATA TAT Y 940 ATT TAA N 0000 GAT CTA |

3070 3080 3090 3100 3110 3120 THOSTIGGICALACTICICALANATICALATITITIATATIATIATIC AMACCACASGICALCISTEMAGISTICITITIAGUENTAMATICALATIATICALANA K T S W S S L N W S F Y F K K I N N H E 3130 3140 3150 3160 3170 3180

3190
3200
3210
3220
3230
3240

ASTANAGATHTHTUCACTHTUCGUAGENACIONACTURATION
GOLDAN

2250 2260 2270 2280 2290 3290 3300 CARTAMOTINGATIOCATITITICATION CRAINING THE AND THICK AND THE AND TH

3310 3320 3330 <u>3340 3350 3360</u> AATCCRATGGATCTAATHGGAGGGATATACAAAAAATTTOCATGGGAACTOOGG TTAGGTCTAGCCTAATTATCCOGTAATAGGTATTTTAAAGGTA <u>CAGC</u> I W I P D L L L C I C F I E M IB6 s.D. (start OED2)

3370
3380
3390
3400
3410
3420

AMACTITITICAMANGUATION
GAMAGUATION
GAMAG

3490
3500
3510
3520
3530
3540

ATTGATTTCOTTAMOSTATIOSTICUTATIONSTICATIONATIONSTICATIONATIONS
TIMOTAMOSANTICATACONSCIENCESCONSCIENCONSCIENCESCONSCIENCONSCIENCESCONSCIENCESCONSCIENCES

FIG. 1-Continued

quence TCAAAAT/CGAAT, found upstream of L. biflexa trpE and trpG genes (42) (Fig. 1). In ORF1 and ORF2 this sequence was closer to the putative translation initiation site than was found with the trp genes. Inverted repeats that could function as rho-independent transcription termination signals were located directly downstream of ORF1 and ORF3. The first inverted repeat downstream of ORF2 was after 80 bp. The initiation methionine codon and the regulatory sequences of ORF3 were not located on clone pHL2-B3. Therefore we only sequenced the coding region for the carboxy-terminal part of the protein.

Localization of region encoding hemolytic and sphingomyelinase activity. A number of deletion mutants (Fig. 2C) were tested for hemolytic activity in blood agar plates and for sphingomyelinase C activity (Fig. 3). In all clones tested in both assays, the presence of sphingomyelinase activity coincided with the presence of hemolytic activity. In both experiments, pHL2-B4-51 was the clone with the smallest DNA insert, still expressing hemolytic and sphingomyelinase C activities at the same level as clones pHL2-B3 and pHL2-B4. E. coli containing recombinant pHL2-B4-53, coding for 45.6 kilodaltons (kDa) of the amino-terminal part of the protein, lost both enzymatic activities. Recombinant pHL2-B4-06 contained the coding region of ORF1 except for the last 10 amino acids; E. coli cells harboring this recombinant still contained both enzymatic activities, but at lower levels (data not shown). Therefore the whole coding region of ORF1 is required for full expression of hemolytic and sphingomyelinase activities in E. coli. Since E. coli containing pHL2-B3-05 was not hemolytic and had no sphingomyelinase activity, expression also required the presence of at least a portion of the 580-bp noncoding DNA region upstream of ORF1. A DNA sequence controlling the expression of sphingomyelinase activity may be located on this DNA region.

Temperature sensitivity and substrate specificity. Sphingomyelinase activity in lysates from *E. coli* cultures harboring pHL2-B3 and *L. interrogans* Sponselee cultures was destroyed by heating the samples for 10 min at 56°C. Lysates from clone pHL2-B3 were also tested for the ability to degrade L- α -lysophosphatidylcholin, L- α -phosphatidylethanolamine, phosphatidyl-L-serine, L- α -phosphatidic acid, and L- α -lecithin, but no activity was detected.

Protein sequence analysis and homology with other sphingomyelinases. At the amino terminus of ORF1, the first 27 amino acids have the consensus of a procaryotic signal sequence (40): a basic amino-terminal region, central hydrophobic region, and a polar carboxy-terminal region. The cleavage site of signal peptides is always preceded by a small amino acid residue at the -1 position and a small uncharged amino acid residue at the -3 position (40). Therefore, cleavage is most likely to occur after the alanine residue at position 27.

Comparison of nucleotide and amino acid sequences of ORF1 with those in data bases revealed homology with a β -hemolysin from *S. aureus* (32) and three sphingomyelinases from different *B. cereus* strains (14, 22, 41) on both the DNA level (data not shown) and the protein level (Fig. 4).

pHL2-B3



FIG. 2. Restriction map of pHL2-B3 and sequence strategy. (A) Properties of the three ORFs deduced from the nucleotide sequence analysis. The hatched bar indicates a putative signal peptide. (B) Restriction map of the 3,987-bp BamHI insert in pHL2-B3 (pHL2-B4 contains the same insert in the opposite orientation) (B, BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PvuI). (C) Deletion clones generated by exonuclease III digestion and used for nucleotide sequence analysis and localization of hemolytic and sphingomyelinase activity. Arrows pointing to the right indicate clones derived from pHL2-B3; the 3,987-bp BamHI insert was digested by exonuclease III exclusively from the side of bp 1, and the remaining insert DNA extends from the base of the arrow up to bp 3987. Arrows pointing to the left are derived from pHL2-B4; the 3,987-bp BamHI insert was digested by exonuclease III exclusively from the side of bp 3987, and the remaining insert DNA extends from the base of the arrow represents the part of which the nucleotide sequence was determined. Deletion clones are named after the clone from which they were derived (i.e., pHL2-B3 or pHL2-B4), followed by the number indicated above. A number of clones were tested for hemolytic activity; boldface and broken-line arrows indicate the presence and absence of hemolytic activity, respectively. The other arrows indicate clones that were not tested.

Both on the DNA and protein levels, the similarity was present in the middle part of the gene and protein, respectively. The amino termini differed considerably, and the carboxy termini could not be aligned because of the difference in molecular weight. No sequences homologous to ORF2 or ORF3 were detected during the homology searches.



FIG. 3. Assay of sphingomyelinase activity of various recombinant DNA clones to localize the region on pHL2-B3 encoding enzyme activity. Sonicated *E. coli* cells containing the plasmids indicated at the top were tested for sphingomyelinase activity; the reaction products were analyzed by thin-layer chromatography. Purified sphingomyelinase C from *S. aureus* and *E. coli* containing the pBS vector were used as positive and negative controls, respectively. The sphingomyelin substrate and ceramide degradation products are indicated on the left by SP and C, respectively.

Codon usage. The codons used in L. interrogans for the expression of the ORFs of pHL2-B3 are shown in Table 1 and compared with the codon usage in L. biflexa trpE and trpG genes (42), E. coli (average values for 407 genes calculated from the data from Aota et al. [1]), S. aureus β -hemolysin (32), and *B*. cereus sphingomyelinase C genes (average values compiled from three published sequences [14, 22, 41]). Codon usage in the L. interrogans genes was quite different from the codon usage in E. coli genes for the amino acids arginine, asparagine, cysteine, glutamine, glycine, proline, and leucine. The codon usage was more similar to that of L. biflexa, although there were differences for asparagine, arginine, isoleucine, and proline. Comparison of the codon usage in the different sphingomyelinase genes revealed the frequent use of the arginine codon AGA and the proline codon CCC in the leptospiral sphingomyelinase gene, whereas these codons were hardly or not at all used in the other four genes. The frequencies of bases in the third position of a codon were calculated and were similar to what has been shown for L. biflexa (42), reflecting the overall base composition of the organism.

Sphingomyelinase activities of different Leptospira strains. Total culture, pellet, and supernatant fractions of four L. interrogans strains belonging to different serogroups were compared for the ability to degrade sphingomyelin (Fig. 5). The results indicated that all four strains contained sphingomyelinase activity. In strains Sponselee and Mus 127, sphingomyelinase activity was associated with the cellular fraction, whereas the sphingomyelinase activity from strains



FIG. 4. Alignment of the amino acid sequences from sphingomyelinases from *L. interrogans* (LI) Sponselee (this work), *S. aureus* β -hemolysin (SA) (32), and *B. cereus* (BC) SE1 (22), IAM1208 (41), and GP4 (14). Identical residues are marked with an asterisk, and conservative mutations are marked with a dot, according to the log-odds amino acid similarity matrix of Dayhoff (9). Strokes indicate a gap introduced into the sequence for alignment purposes.

Pomona and Hond Utrecht IV seemed to be secreted. When sonication was omitted, the cellular fraction of strain Sponselee still had sphingomyelinase activity (data not shown but identical to those in Fig. 5). Therefore the enzyme is probably located in the outer envelope. No sphingomyelinase activity could be demonstrated in strains Wijnberg and M20 (belonging to serogroup Icterohaemorrhagiae) and the apathogenic PatocI strain (L. biflexa).

DISCUSSION

The nucleotide sequence analysis of a 3,987-bp DNA fragment of L. *interrogans*, encoding sphingomyelinase, revealed the presence of three ORFs. These are the first L. *interrogans* protein-encoding genes for which the nucleotide sequence has been determined. The molecular mass of the product of ORF1 (63,268 Da), corresponds to the apparent

molecular mass of 64 kDa from a hemolysin cloned from L. interrogans serovar pomona (8; A. A. Dain, M. N. Rozinov, and Y. G. Chernukha, VI Joint Meeting of Leptospira Workers, abstr. no. 7.7, 1988). It has previously been shown that from the DNA insert of pHL2-B3, only one smaller protein of 39.2 kDa (41.6 kDa, including the 2.4-kDa signal peptide) is expressed in E. coli (10). Indeed, ORF1 has a putative signal peptide with a calculated molecular mass of 3,175 Da. Moreover, ORF1 is the only reading frame large enough to encode such a protein and has been shown here to code for hemolysin and sphingomyelinase activities. We therefore conclude that the 39.2-kDa protein is the mature expression product of ORF1 in E. coli minicells and probably has sphingomyelinase activity. However, we cannot exclude the possibility that sphingomyelinase activity is expressed as a larger, short-lived protein, and that the

2182 SEGERS ET AL.

| TADLE 1 | Coder were in T | | | 1. 0 | | л <i>а</i> |
|----------|-------------------|----------------|---------------|---------|---------------------|------------|
| IADLE I. | Couon usage in L. | interrogans, i | L. вілеха, Е. | cou, s. | <i>aureus</i> , and | B. cereus" |

| | | Frequency of codon usage (%) in: | | | | | | | | | | |
|------------|--------------|----------------------------------|----------------|------------------|--------------------|--------|------------------|-------------------|--|--|--|--|
| Amino acid | Codon | | L. interrogans | | I. hiflera | F coli | S aurous | R carava | | | | |
| | | ORF1 | ORF2 | ORF3 | г. однеха | E. COU | s. aureus | B. cereus | | | | |
| Ala | GCT | 8 (28) | 1 (50) | 1 (25) | 16 (45) | (20) | 3 (18) | 14 (21) | | | | |
| Ala | GCC | 4 (13) | 0 (0) | 0 (0) | 3 (9) | (23) | 3 (18) | 4 (6) | | | | |
| Ala | GCA | 9 (31) | 0 (0) | 3 (75) | 10 (29) | (22) | 8 (46) | 37 (54) | | | | |
| Ala | GCG | 8 (28) | 1 (50) | 0 (0) | 6 (17) | (35) | 3 (18) | 13 (19) | | | | |
| Arg | CGT | 4 (16) | 1 (13) | 1 (20) | 7 (25) | (59) | 2 (40) | 16 (100) | | | | |
| Arg | CGC | 0 (0) | 2 (25) | 0 (0) | 2 (7) | (37) | 1 (20) | 0 (0) | | | | |
| Arg | CGA | 3 (12) | 1 (13) | 1 (20) | 10 (36) | (4) | 1 (20) | 0 (0) | | | | |
| Arg | CGG | 1 (4) | 1 (13) | 1 (20) | 2 (7) | (6) | 0 (0) | 0 (0) | | | | |
| Arg | AGA | 13 (52) | 3 (37) | 2 (40) | 5 (18) | (2) | 1 (20) | 0 (0) | | | | |
| Arg | AGG | 4 (16) | 0 (0) | 0 (0) | 2 (7) | (2) | 0 (0) | 0 (0) | | | | |
| Asn | AAT | 20 (51) | 13 (62) | 13 (76) | 24 (89) | (37) | 21 (81) | 63 (70) | | | | |
| Asn | AAC | 19 (49) | 8 (38) | 4 (24) | 2 (11) | (63) | 5 (19) | 27 (30) | | | | |
| Asp | GAT | 32 (86) | 7 (58) | 5 (71) | 24 (83) | (57) | 19 (73) | 57 (92) | | | | |
| Asp | GAC | 5 (14) | 5 (42) | 2 (29) | 5 (17) | (43) | 7 (27) | 5 (8) | | | | |
| Cvs | тст | 7 (99) | 1 (50) | 0 (0) | 2 (60) | (42) | 2 (100) | 1 (17) | | | | |
| Cys | TGC | 1 (12) | 1 (50) | 0 (0) | 2 (40) | (42) | 2 (100) 0 (0) | 5 (83) | | | | |
| | | - (, | - () | | - (, | (| 0 (0) | - () | | | | |
| Gln | CAA | 15 (83) | 8 (100) | 14 (93) | 19 (90) | (30) | 9 (90) | 27 (75) | | | | |
| Gin | CAG | 3 (17) | 0(0) | 1(/) | 2 (10) | (70) | 1 (10) | 9 (25) | | | | |
| Glu | GAA | 15 (75) | 18 (78) | 19 (76) | 49 (86) | (71) | 13 (87) | 25 (60) | | | | |
| Glu | GAG | 5 (25) | 5 (22) | 6 (24) | 8 (14) | (29) | 2 (13) | 17 (40) | | | | |
| Glv | GGT | 10 (25) | 1 (4) | 1 (13) | 20 (32) | (41) | 11 (52) | 22 (33) | | | | |
| Glv | GGC | 4 (10) | 6 (25) | 0(0) | 2(3) | (41) | 5 (24) | 5 (7) | | | | |
| Glv | GGA | 22 (55) | 14 (58) | 6 (74) | 31 (50) | (7) | 3 (14) | 27 (40) | | | | |
| Gly | GGG | 4 (10) | 3 (13) | 1 (13) | 9 (15) | (11) | 2 (10) | 13 (20) | | | | |
| His | САТ | 7 (70) | 3 (75) | 1 (100) | 12 (75) | (49) | 8 (89) | 13 (76) | | | | |
| His | CAC | 3 (30) | 1 (25) | 0 (0) | 4 (25) | (51) | 1 (11) | 4 (24) | | | | |
| Ца | ۸ T T | 16 (40) | 12 (49) | 5 (29) | 22 (64) | (14) | 5 (22) | 27 (62) | | | | |
| Ile | ATC | 10 (40) | 13 (40) | J (20) 3 (22) | 52 (04) 12 (24) | (44) | 5 (33) | 37 (02) 7 (12) | | | | |
| Ile | ATA | 6 (15) | 5 (19) | 7 (39) | 6 (12) | (51) | 4 (27) | 16 (26) | | | | |
| | | 0 (10) | | | 0 (11) | | . (=.) | | | | | |
| Leu | TTA | 10 (24) | 7 (22) | 2 (10) | 18 (29) | (10) | 11 (52) | 43 (61) | | | | |
| Leu | TTG | 10 (24) | 6 (19) | 9 (43) | 14 (22) | (11) | 3 (14) | 14 (20) | | | | |
| Leu | CTT | 8 (20) | 8 (25) | 3 (14) | 17 (28) | (9) | 4 (19) | 3 (6) | | | | |
| Leu | CTC | 6 (15) | 2 (6) | 3 (14) | 9 (13) | (9) | 1 (5) | 1(1) | | | | |
| Leu | CTA | 4 (10) | 5 (16) | 4 (29) | 5 (8) | (3) | 2 (10) | 8 (11) | | | | |
| Leu | CIG | 3 (7) | 4 (12) | 0 (0) | 0(0) | (58) | 0(0) | 1 (1) | | | | |
| Lys | AAA | 35 (80) | 19 (83) | 20 (83) | 39 (81) | (76) | 36 (88) | 61 (74) | | | | |
| Lys | AAG | 9 (20) | 4 (17) | 4 (17) | 9 (19) | (24) | 5 (12) | 21 (26) | | | | |
| Met | ATG | 6 (100) | 6 (100) | 1 (100) | 15 (100) | (100) | 4 (100) | 19 (100) | | | | |
| Phe | TTT | 16 (64) | 18 (78) | 3 (75) | 32 (84) | (47) | 5 (50) | 17 (63) | | | | |
| Phe | TTC | 9 (36) | 5 (22) | 1 (25) | 6 (16) | (53) | 5 (50) | 10 (37) | | | | |
| Pro | ССТ | 4 (20) | 3 (33) | 3 (50) | 11 (31) | (14) | 7 (50) | 9 (23) | | | | |
| Pro | | 8 (40) | 0(0) | 1 (17) | 8 (22) | (8) | 0 (0) | 0 (0) | | | | |
| Pro | CCA | 3 (15) | | 0 (0) | 13 (36) | (18) | 7 (50) | 25 (62) | | | | |
| Pro | CCG | 5 (25) | 6 (67) | 2 (33) | 4 (11) | (60) | 0 (0) | 6 (15) | | | | |
| Ser | тст | 15 (22) | 0 (28) | 5 (20) | 13 (78) | (21) | 5 (19) | 25 (28) | | | | |
| Ser | TCC | 11 (24) | 5 (21) | 1 (5) | 7 (14) | (18) | 1 (4) | 1 (1) | | | | |
| Ser | TCA | 6 (13) | 3 (13) | 3 (15) | 8 (16) | (10) | 9 (35) | 25 (28) | | | | |
| Ser | TCG | 4 (9) | 1 (4) | 6 (30) | 9 (18) | (13) | 1 (4) | 6 (7) | | | | |
| Ser | AGT | 8 (17) | 3 (12) | 2 (10) | 11 (22) | (11) | 6 (23) | 20 (22) | | | | |
| Ser | AGC | 2(4) | 3 (12) | 3 (15) | 1(2) | (27) | 4 (15) | 13 (14) | | | | |
| ~~. | | - () | - () | - (10) | - (-) | () | / | - (= -) | | | | |

Continued on following page

| Amino acid | | Frequency of codon usage (%) in: | | | | | | | | | | |
|------------|-------|----------------------------------|----------------|---------|------------|--------|-----------|-----------|--|--|--|--|
| | Codon | | L. interrogans | | I hifterna | E coli | | R caraus | | | | |
| | | ORF1 | ORF2 | ORF3 | L. Dipexa | E. cou | S. aureus | B. cereus | | | | |
| Thr | ACT | 9 (32) | 7 (35) | 3 (25) | 5 (25) | (21) | 6 (38) | 23 (34) | | | | |
| Thr | ACC | 4 (14) | 3 (15) | 4 (33) | 5 (25) | (46) | 1 (6) | 0 (0) | | | | |
| Thr | ACA | 5 (18) | 3 (15) | 2 (17) | 7 (35) | (10) | 8 (50) | 24 (35) | | | | |
| Thr | ACG | 10 (36) | 7 (35) | 3 (25) | 3 (15) | (22) | 1 (6) | 21 (31) | | | | |
| Trp | TGG | 15 (100) | 5 (100) | 1 (100) | 2 (100) | (100) | 4 (100) | 18 (100) | | | | |
| Tvr | TAT | 20 (67) | 9 (64) | 3 (75) | 15 (60) | (50) | 15 (71) | 45 (87) | | | | |
| Tyr | TAC | 10 (33) | 5 (36) | 1 (25) | 10 (40) | (50) | 6 (29) | 7 (13) | | | | |
| Val | GTT | 11 (31) | 4 (25) | 0 (0) | 12 (31) | (31) | 14 (52) | 19 (28) | | | | |
| Val | GTC | 6 (17) | 5 (31) | 1 (33) | 4 (11) | (18) | 3 (11) | 1 (2) | | | | |
| Val | GTA | 9 (26) | 5 (31) | 1 (33) | 13 (33) | (18) | 7 (26) | 26 (38) | | | | |
| Val | GTG | 9 (26) | 2 (13) | 1 (33) | 10 (25) | (33) | 3 (11) | 22 (32) | | | | |

TABLE 1—Continued

^a Codons used in the three ORFs of *L. interrogans* were compared with those of *L. biflexa trpE* and trpG genes (41), *E. coli* (average values for 407 genes calculated from the data from Aota et al. [1]), *S. aureus* β -hemolysin (32), and *B. cereus* sphingomyelinase C genes (average values compiled from three published sequences [14, 22, 41]). The frequency of codon usage is indicated, followed by a percentage, representing the number of times this codon is used to encode its amino acid.

39.2-kDa protein, generated by degradation or processing, has no enzymatic activity. The molecular mass of 39.2 kDa is comparable to the observed molecular mass of the four homologous sphingomyelinases and corresponds very well to the calculated molecular mass of 38,660 Da of the Nterminal part of the protein, starting after the signal sequence, up to the point where homology with the other sphingomyelinases ends (Fig. 4). How do we explain the discrepancy between the size of ORF1 and the experimentally detected product? Inverted repeats IR2 and IR3, which could be involved in transcription or translation termination, are located immediately downstream of the DNA region coding for this part of the protein (Fig. 1). Premature transcription or translation termination, however, is unlikely to occur, since the whole coding region of ORF1 is needed for optimal sphingomyelinase activity in E. coli. More likely, the 39.2-kDa protein would be the result of posttranslational processing of the complete 63-kDa expression product of ORF1. Since in E. coli minicells processing occurs before cleavage of the amino-terminal signal sequence, the former processing can only take place at the carboxy terminus of the protein (10). This is supported by the homology data presented in Fig. 4. Similar posttranslational processing at the carboxy terminus of a protein has previously been reported for the immunoglobulin A protease from Neisseria gonorrhoeae (31), serine protease from Serratia marcescens (29), and activation of aerolysin from Aeromonas hydrophila (21).

The codon usage in the L. interrogans sphingomyelinase gene is quite different from what is normally observed in E. coli for genes with an average expression level. Since the expression rate of genes in E. coli is known to be related to their codon usage (15, 34), the leptospiral genes will probably have a low level of expression in E. coli. Indeed, no difference could be detected between Coomassie-stained polyacrylamide gels containing lysates from E. coli containing pBS or pHL2-B3 (data not shown). Alternatively, the low expression level of the enzyme could be the result of weak promoter activity. On both the DNA and protein levels, the middle part of the leptospiral sphingomyelinases shares a high degree of similarity with sphingomyelinases from the distantly related bacteria S. aureus and B. cereus. Obviously this part is important for enzyme activity. The differences in the amino termini of the proteins could reflect differences in transport, since *B. cereus* and *S. aureus* sphingomyelinase are extracellular enzymes, whereas the sphingomyelinase C activity of *L. interrogans* Sponselee seems to be cell bound. ORF2 and ORF3 are not necessary for sphingomyelinase activity (Fig. 3); no similar nucleotide or amino acid sequences were found in the data banks, and the functions of the proteins encoded by ORF2 and ORF3 remain unknown. Unlike the case in *B. cereus* GP-4 and IAM1208 (14, 41), in which a phospholipase C gene is located directly downstream of the sphingomyelinase C gene, no such gene was found downstream of the leptospiral sphingomyelinase gene. Since it has previously been shown that five different clones, containing over 10 kilobases of genomic



FIG. 5. Sphingomyelinase activity of four different L. interrogans strains. Bacterial cultures were harvested in the logarithmic growth phase, and a total (T) culture sample was separated into pellet (P) and supernatant (S) fractions by centrifugation for 10 min at $6,000 \times g$. After mild sonication, the samples were tested for their ability to degrade sphingomyelin, which was monitored by thin-layer chromatography. The sphingomyelin substrate and ceramide degradation products are indicated on the left by SP and C, respectively. Purified sphingomyelinase C from S. aureus and E. coli containing the pBS vector were used as positive and negative controls, respectively. The following L. interrogans strains were tested; Sponselee (serogroup Sejroe, serovar hardjo); Mus127 (serogroup Ballum, serovar canicola); Pomona (serogroup Pomona, serovar pomona).

DNA on which the sphingomyelinase gene is located, were negative in a phospholipase C assay (10), it is unlikely that such a gene is located nearby on the L. interrogans genome.

Among strains of L. interrogans, various sphingomyelinases are produced. Strain Mus127 (serovar ballum) seems to contain a gene similar to that present in pHL2-B3 (10). Strains Hond Utrecht IV and Pomona, however, which do not cross-hybridize with the sphingomyelinase gene from strain Sponselee under stringent conditions (10), do degrade sphingomyelin. Moreover, the sphingomyelinase activities of strains Hond Utrecht IV and Pomona are predominantly found in the supernatant, whereas those of strain Sponselee and Mus127 are cell associated. Possibly, the sphingomyelinase is a contact hemolysin in strains Mus127 and Sponselee and an excreted hemolysin in strains Pomona and Hond Utrecht IV. Strains belonging to the pomona and ballum serovars were also reported to have different hemolytic properties (37). Bovine erythrocytes are preferentially lysed by pomona strains, and hamster erythrocytes are preferentially lysed by ballum strains. It is not known whether this difference is caused by a difference in substrate specificity of the sphingomyelinase. Contrary to the Dutch field strain Sponselee, L. interrogans strains of the serovar hardjo, isolated in New Zealand, do not lyse ovine erythrocytes (18). Therefore genetic variation seems to occur within the serovar hardjo. This is supported by the presence of different genotypes within the serovar hardjo, based on DNA restriction endonuclease analysis (12), and differences in the G+Cpercentage of the genomic DNA (27). Although the presence of multiple sphingomyelinase genes within the L. interrogans species indicates the importance of this enzyme for the bacterium, the involvement of the sphingomyelinase in pathogenesis is not known. However, several speculations can be made. The homologous β -hemolysin from S. aureus significantly increases recovery of bacteria from experimentally infected mice, compared with that of the β -hemolysinnegative mutant, and therefore contributes to virulence in vivo (5). Second, leptospires cannot grow without iron (13) and use free fatty acids as the main carbon and energy source (23). The sphingomyelinase could therefore play a role in obtaining iron and fatty acids from lysed erythrocytes. Third, sphingomyelinase may be an important factor in pathogenesis without lysing erythrocytes; sphingolipids like sphingomyelin and their degradation products affect many pharmalogical responses, growth factor action, receptor functions, and phorbolester-induced responses and have been implicated as second messengers (17).

ACKNOWLEDGMENTS

R.P.A.M.S. thanks Intervet International B.V. (Boxmeer, The Netherlands) for a research grant. Use of the services and facilities of the Dutch National NWO/SURF Expertise Center CAOS/ CAMM, under grant numbers SON 326-052 and STW NCH99.1751, is gratefully acknowledged.

LITERATURE CITED

- Aota, S., S. Gojobori, F. Ishibashi, T. Maruyama, and T. Ikemura. 1988. Codon usage tabulated from the Genbank genetic sequence data. Nucleic Acids Res. 16:r315-r402.
- Bauer, D. C., L. N. Eames, S. D. Sleight, and L. C. Ferguson. 1961. The significance of leptospiral hemolysin in the pathogenesis of *Leptospira pomona* infections. J. Infect. Dis. 108:229-236.
- 3. Bernheimer, A. W., and R. F. Bey. 1986. Copurification of Leptospira interrogans serovar pomona hemolysin and sphingomyelinase C. Infect. Immun. 54:262-264.
- 4. Berry, A. M., J. Yother, D. E. Briles, D. Hansman, and J. C. Paton. 1989. Reduced virulence of a defined pneumolysin-

negative mutant of *Streptococcus pneumoniae*. Infect. Immun. 57:2037–2042.

- Bramley, A. J., A. H. Patel, M. O'Reilly, R. Foster, and T. J. Foster. 1989. Roles of alpha-toxin and beta-toxin in virulence of *Staphylococcus aureus* for the mouse mammary gland. Infect. Immun. 57:2489-2494.
- 6. Chorvath, B. 1974. Some properties of leptospiral phospholipase A and hemolysin. Folia Fac. Med. Univ. Comenianae Bratisl. Suppl. 12:43-56.
- Cossart, P., M. F. Vincente, J. Mengaud, F. Baquero, J. C. Perez-Diaz, and P. Berche. 1989. Listeriolysin O is essential for virulence of *Listeria monocytogenes*: direct evidence obtained by gene complementation. Infect. Immun. 57:3629–3636.
- Dain, A. A., M. N. Rozinov, T. A. Holzmayer, V. N. Gershanovich, and G. Chernukha. 1985. Cloning and expression of *Leptospira pomona* hemolysin gene in *Escherichia coli*. Zh. Mikrobiol. Epidemiol. Immunol. 7:7-10.
- Dayhoff, M. O. 1987. A model of evolutionary change in proteins. Matrices for detecting distant relationships, p. 345– 358. In M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5, suppl. 3. National Biomedical Research Foundation, Washington, D.C.
- Del Real, G., R. P. A. M. Segers, B. A. M. Van der Zeijst, and W. Gaastra. 1989. Cloning of a hemolysin gene from *Leptospira interrogans* serovar hardjo. Infect. Immun. 57:2588-2590.
- 11. Ellis, W. A., J. J. O'Brien, D. G. Brynson, and D. P. Mackie. 1985. Bovine leptospirosis: some clinical features of serovar *hardjo* infection. Vet. Rec. 117:101-104.
- Ellis, W. A., A. B. Thiermann, J. Montgomery, A. Handsacker, P. J. Winter, and R. B. Marshall. 1988. Restriction endonuclease analysis of *Leptospira interrogans* serovar *hardjo* isolates from cattle. Res. Vet. Sci. 44:375–379.
- 13. Faine, S. 1959. Iron as a growth requirement for pathogenic leptospires. J. Gen. Microbiol. 20:246–251.
- Gilmore, M. S., A. L. Cruz-Rodz, M. Leimeister-Wachter, J. Kreft, and W. Goebel. 1989. A *Bacillus cereus* cytolytic determinant cereolysin AB which comprises the phospholipase C and Sphingomyelinase genes: nucleotide sequence and genetic linkage. J. Bacteriol. 171:744–753.
- Grantham, R., C. Gautier, M. Gouy, and R. Mercier. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. 9:r43-r74.
- Hacker, J., H. Hof, L. Emody, and W. Goebel. 1986. Influence of cloned *Escherichia coli* hemolysin genes, S-fimbriae and serum resistance on pathogenicity in different animal models. Microb. Pathogen. 1:533-547.
- Hannun, Y. A., and R. M. Bell. 1989. Functions of sphingolipids and sphingolipid breakdown products in cellular regulation. Science 243:500-507.
- Hathaway, S. C., and R. B. Marchall. 1980. Haemolysis as a means of distinguishing between *Leptospira interrogans* serovars *balcanica* and *hardjo*. J. Med. Microbiol. 13:477–481.
- 19. Higgins, D. G., and P. M. Sharp. 1988. CLUSTAL: a package for performing multiple alignment on a microcomputer. Gene 73:237-244.
- Higgins, D. G., and P. M. Sharp. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS Commun. 5:151-153.
- Howard, S. P., and J. T. Buckley. 1985. Activation of the hole-forming toxin aerolysin by extracellular processing. J. Bacteriol. 163:336-340.
- Johansen, T., F. B. Haugli, H. Ikezawa, and C. Little. 1988. Bacillus cereus strain SE-1: nucleotide sequence of the sphingomyelinase C gene. Nucleic Acids Res. 16:10370.
- 23. Johnson, R. C., and S. Faine. 1984. Genus I. Leptospira Noguchi 1917, p. 62. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology. The Williams & Wilkins Co., Baltimore.
- 24. Kojima, T., Y. Yanagihara, and I. Mifuchi. 1984. Characterization of inhibitor to leptospiral hemolysin present in bovine serum. Microbiol. Immunol. 28:291-302.
- 25. Kozaki, S., K. Kato, T. Asao, Y. Kamata, and G. Sakaguchi. 1987. Activities of *Aeromonas hydrophila* hemolysins and their

interaction with erythrocyte membranes. Infect. Immun. 55: 1594-1599.

- Kuhn, M., S. Kathariou, and W. Goebel. 1988. Hemolysin supports survival but not entry of the intracellular bacterium Listeria monocytogenes. Infect. Immun. 56:79-82.
- LeFebvre, R. B., A. B. Thiermann, and J. Foley. 1987. Genetic and antigenic differences of serologically indistinguishable leptospires of serovar hardjo. J. Clin. Microbiol. 25:2094–2097.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 29. Miyazaki, H., N. Yanagida, S. Horinouchi, and T. Beppu. 1989. Characterization of the precursor of *Serratia marcescens* serine protease and COOH-terminal processing of the precursor during its excretion through the outer membrane of *Escherichia coli*. J. Bacteriol. 171:6566–6572.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.
- Pohler, J., R. Halter, K. Beyreuther, and T. F. Meyer. 1987. Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature (London) 325:458–462.
- 32. Projan, S. J., J. Kornlum, B. Kreiswirth, S. L. Moghazeh, W. Eisner, and R. P. Novick. 1989. Nucleotide sequence: the Beta-hemolysin gene of *Staphylococcus aureus*. Nucleic Acids Res. 17:3305.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 34. Sharp, P. M., E. Cowe, D. G. Higgins, D. C. Shields, K. H.

Wolfe, and F. Wright. 1988. Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster, and Homo sapiens; a review of the considerable within-species diversity. Nucleic Acids Res. 16:8207–8211.

- 35. Stalheim, O. H. V. 1971. Virulent and avirulent leptospires: biochemical activities and survival in blood. Am. J. Vet. Res. 32:843-849.
- Thiermann, A. B. 1984. Leptospirosis: current developments and trends. J. Am. Vet. Med. Assoc. 184:722-725.
- Thompson, J. C., and R. B. Marshall. 1986. In vitro studies of haemolysis by *Leptospira interrogans* serovars *pomona* and *ballum*. Vet. Microbiol. 11:285-292.
- Trowbridge, A. A., J. B. Green, J. D. Bonnet, S. B. Shohet, B. D. Ponnappa, and W. B. McCombs. 1981. Hemolytic anemia associated with leptospirosis. Morphologic and lipid studies. Am. J. Clin. Pathol. 76:493–498.
- Volina, E. G., L. F. Levina, and G. L. Soboleva. 1986. Phospholipase activity and virulence of pathogenic *Leptospirae*. J. Hyg. Epidem. Microbiol. Immunol. 30:163–169.
- 40. Von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683-4690.
- Yamada, A., N. Tsukagoshi, S. Udaka, S. Makino, S. Nakamura, C. Little, M. Tomita, and H. Ikezawa. 1988. Nucleotide sequence and expression in *Escherichia coli* of the gene coding for sphingomyelinase of *Bacillus cereus*. Eur. J. Biochem. 175: 213-220.
- Yelton, D. B., and S. L. Peng. 1989. Identification and nucleotide sequence of the *Leptospira biflexa* serovar patoc *trpE* and *trpG* genes. J. Bacteriol. 171:2038–2089.