## Identification of a pathogenicity island required for *Salmonella* survival in host cells

Howard Ochman\*, Fernando C. Soncini<sup>†</sup>, Felix Solomon<sup>†</sup>, and Eduardo A. Groisman<sup>†‡</sup>

\*Department of Biology, University of Rochester, Rochester, NY 14627; and <sup>†</sup>Department of Molecular Microbiology, Washington University School of Medicine, Campus Box 8230, St. Louis, MO 63110

Communicated by Robert K. Selander, Pennsylvania State University, University Park, PA, April 1, 1996 (received for review October 26, 1995)

ABSTRACT We have identified a region unique to the Salmonella typhimurium chromosome that is essential for virulence in mice. This region harbors at least three genes: two (spiA and spiB) encode products that are similar to proteins found in type III secretion systems, and a third (*spiR*) encodes a putative regulator. A strain with a mutation in spiA was unable to survive within macrophages but displayed wild-type levels of epithelial cell invasion. The culture supernatants of the spi mutants lacked a modified form of flagellin, which was present in the supernatant of the wild-type strain. This suggests that the Spi secretory apparatus exports a protease, or a protein that can alter the activity of a secreted protease. The "pathogenicity island" harboring the spi genes may encode the virulence determinants that set Salmonella apart from other enteric pathogens.

Although virulence genes map to numerous chromosomal locations (1), several large regions-termed "pathogenicity islands"-often define virulence characteristics in enteric bacteria. Pathogenicity islands are found at several positions in the chromosome, and the genes within a given island often determine the specific disease condition that results from infection. For example, uropathogenic (2) and enteropathogenic (3) strains of Escherichia coli harbor different pathogenicity islands at the same chromosomal location (the selC locus). Although the pathogenicity island specific to uropathogenic strains encodes a hemolysin (2), the island in enteropathogenic E. coli encodes a type III secretion system that exports proteins responsible for the attachment and effacing lesions of intestinal cells (4). Therefore, the specific virulence properties of two types of pathogenic E. coli are largely determined by the "cassette" present at the selC locus.

Through the systematic screening of a Salmonella typhimurium DNA library, Fitts (5) recovered numerous clones containing sequences that were apparently unique to the salmonellae. Three of these clones appeared to have been acquired by horizontal gene transfer since they had base compositions much lower than the overall G+C content of 52-54% of the Salmonella genome (6). The restricted phylogenetic distributions of these sequences suggested that they encode biochemical or cellular functions that set Salmonella apart from other enteric species (7). One of these three clones contained a gene cluster-designated spa-that enables Salmonella to invade epithelial cells (8). The spa gene cluster is part of a 40-kb pathogenicity island that encodes a variety of determinants that mediate the entry of Salmonella into nonphagocytic cells (9). These determinants include a type III secretion system and its substrates, which are homologous to an antigen export apparatus on the virulence plasmid of Shigella (10). Another of these clones specified a member of the LysR family of transcriptional regulators—SinR—and has no role in virulence (6).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

In a survey of enteric species, only one of the clones— RF333—was strictly confined to *Salmonella* (6). Here we report the molecular genetic characterization of this clone and the virulence properties of strains harboring mutations in the corresponding region of the chromosome. These analyses define a new pathogenicity island (*spi*) within *Salmonella* encoding a regulator and a type III secretion system essential for virulence in mice and survival in macrophages. The *spi* island is distinct from the invasion region containing the *spa* locus, suggesting that multiple events of horizontal transfer have been responsible for the acquisition of pathogenic properties in *Salmonella*.

## MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions. All S. typhimurium strains used in this study were derived from the mouse-virulent wild-type strain 14028s. Strains with mutations in the RF333 chromosomal region were constructed from the pMS333 derivatives pEG7186 and pEG7200, which harbor kan insertions in spiA and spiR, respectively (6). The resulting strains-EG5793 and EG5799-contain the SmaI 1.3-kb kan fragment from plasmid pUC4-KIXX (Pharmacia) inserted at the PmeI and NaeI sites of RF333, respectively (Fig. 1). The kan gene is in the same transcriptional orientation as spiA in EG5793 and as spiR in EG5799, and the kan promoter is predicted to transcribe the genes located downstream of spiA and spiR. [Similar mutations within the spa region generated by inserting the same kan cassette were not polar on downstream genes (8)]. The structure of the spi locus in the mutant strains was verified by PCR with RF333-specific primers and by Southern hybridization analysis with both the 5.7-kb BamHI fragment of RF333 and kan-specific probes. Plasmid pMS333 is a pUC19 derivative with the BamHI fragment from RF333 inserted at the BamHI site. Ampicillin was used at 50  $\mu$ g/ml and kanamycin at 40  $\mu$ g/ml.

**DNA Sequencing, Analysis, and Other Molecular Biological Manipulations.** Sequencing was carried out on both strands of plasmid pMS333 by the dideoxy chain termination method using the Sequenase kit (United States Biochemical) with <sup>35</sup>S-labeled dATP. Additional primers were synthesized as the partial sequences were obtained. Sequence analyses were performed with both GeneWorks (IntelliGenetics) and the GCG package (University of Wisconsin). Restriction endonucleases and phage T4 ligase were purchased from Bethesda Research Laboratories, Boehringer Mannheim, or New England Biolabs, and used according to the supplier's specifications. Other protocols were taken from Maniatis *et al.* (11).

**Virulence Assays and Preparation of Culture Supernatants.** Macrophage survival assays were conducted with the macrophage-like cell line J774 as described (12). Adhesion to and invasion of human intestinal Henle-407 cells were investigated

Abbreviation: ORF, open reading frame.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. U51927). \*To whom reprint requests should be addressed.

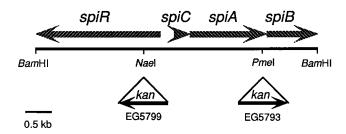


FIG. 1. Genetic and physical maps of the *S. typhimurium* RF333 (*spi*) region. The arrows indicate the size and direction of transcription of the four largest open reading frames (ORF) contained within RF333. The positions of the *kan* gene insertions in the mutants EG5793 and EG5799 are indicated.

as described (13) except that dilutions of bacterial cells were performed in PBS instead of Luria–Bertani broth. Virulence assays were performed with 7–8-week-old female BALB/c mice inoculated intraperitoneally with 100  $\mu$ l of bacteria diluted in PBS. Viability was recorded for at least 30 days in 10 mice per mutant at each dose. Culture supernatants were prepared and analyzed as described (14).

## RESULTS

Molecular Genetic Analysis of a Salmonella-Specific Chromosome Segment. We determined the nucleotide sequence of the 5.7-kb BamHI insert of clone RF333 and identified four ORFs, which were designated as spiA, spiB, spiC, and spiR (spi for <u>Salmonella</u> pathogenicity island). The spiA, spiB, and spiC genes are organized in an operon that is divergently transcribed from *spiR* (Fig. 1). The *spiA* gene encodes a protein of 497 amino acids, and the *spiB* ORF is >323 amino acids and extends beyond the BamHI site at the end of the insert. The spiA gene is preceded by the spiC gene, which encodes a protein of 127 or 133 amino acids, depending on the translational start site. (Presumably, the 127-amino acid ORF is translated because the sequence of its putative ribosome binding site—AGGAG—corresponds to the 3' end of the 16s rRNA.) The spiA initiation codon is 1 bp downstream of the stop codon for *spiC*, which suggests that these proteins are translationally coupled. The initiation codon of spiC is separated by 400 bp from that of the spiR gene, which encodes a protein of > 822 amino acids.

SpiR Is Similar to Proteins of the Two-Component Family. Two-component systems generally consist of a sensor protein that, in response to environmental cues, modifies the ability of a regulatory protein to affect transcription of particular genes (15). Several such systems have been implicated in the regulation of virulence functions in the salmonellae (16). The deduced amino acid sequence of the SpiR protein exhibits homology to the subgroup of two-component systems that have both conserved domains-the histidine kinase domain of sensors and the receiver domain of response regulators-in a single molecule (Fig. 2A). The SpiR protein sequence is related to: (i) E. coli BarA, which is encoded by a gene that was isolated as a multicopy number suppressor of an envZ mutation (17); (ii) RscC, which controls capsule synthesis in E. coli (18); (iii) BvgS, a major regulator of virulence determinants in Bordetella (19); as well as two-component systems of plants and yeast (20).

SpiA and SpiB Are Similar to Proteins of Type III Secretion Systems. Type III secretion systems export proteins by a mechanism distinct from either the classical *sec*-dependent pathway or that responsible for secretion of hemolysin in *E. coli* (21). These systems mediate secretion of virulence proteins in both animal and plant pathogens: Homologs include the *Salmonella* Inv/Spa and the *Shigella* Mxi/Spa complexes, which are necessary for host cell invasion (22); the *Yersinia* Ysc/Lcr proteins, which secrete a tyrosine phosphatase and a cytotoxin (23); and proteins in the plant pathogens *Erwinia*, *Pseudomonas*, and *Xanthomonas*, which are involved in the hypersensitive response (24, 25).

As shown in Fig. 2B, SpiA displays the highest degree of sequence similarity to SepC of enteropathogenic E. coli (4), YscC of Yersinia (26), and InvG of Salmonella (27); and SpiB exhibited a low level of similarity to Yersinia YscD (26) and to an unreported ORF adjacent to eaeA in enteropathogenic E. coli (4). The gene order of spiA and spiB on the Salmonella chromosome is the same as that of yscC and yscD on the virulence plasmid of Yersinia (26). But spiC has no homologs in the sequence databases and apparently is a component specific to the spi locus.

**Virulence Properties of** *spi* **Mutants.** To determine the function of the genes present in RF333, we constructed mutations in the *spiA* and *spiR* ORFs (Fig. 1; see *Materials and Methods*). When these mutations were transferred back to the *S. typhimurium* chromosome, viable colonies were obtained, indicating that the inactivated genes are not essential for growth in complex laboratory media. These mutants behaved like the wild-type parent in their ability to grow at different temperatures and in defined media.

Given the prevalence of type III secretion systems among animal and plant pathogens (21, 22), we hypothesized that strains with mutations in *spi* genes might be unable to export certain virulence determinants and, consequently, would be attenuated for virulence. Indeed, neither *spiR* nor *spiA* mutants killed BALB/c mice when inoculated intraperitoneally at >10,000 times the median lethal dose of the wild-type parent. This finding is in contrast to the phenotypes of the *inv/spa* mutants of *S. typhimurium* that are attenuated only when the bacteria are administered orally (28).

To determine the particular stage of infection in which the RF333-encoded determinants are required, we tested the ability of the *spiA* mutant to invade cultured epithelial cells and to survive within macrophages *in vitro*. The *spiA* mutant displayed wild-type levels of epithelial cell invasion but was defective for intramacrophage survival (Fig. 3), a result that further distinguishes the role of the *spi* genes from that of the genes in the *inv/spa* locus. Taken together with the mice virulence data, these experiments establish that the Spi proteins are essential for stages of infection beyond the initial interaction with intestinal cells.

Identification of Proteins Exported by the Spi Secretion System. Because spi apparently encodes components of a type III secretion system, we investigated whether spi mutants were defective in the export of proteins to the growth media. Culture supernatants of wild-type, spiA, and spiR mutant strains were prepared and their protein profiles were compared. The supernatant of the wild-type strain had a 45-kDa protein that was absent from the supernatant of the mutant strains; instead, the mutant supernatant harbored a 47-kDa protein that was absent from the wild-type cultures (Fig. 4). We isolated both the 45- and 47-kDa proteins and determined the amino acid sequences of their first 15 residues. The two sequences were identical to one another and to the S. typhimurium flagellin, and it is likely that the 45-kDa protein is a processed form of the 47-kDa protein. The profile of secreted proteins for the spiR mutant was identical to that displayed by the spiA mutant strain, suggesting that SpiR controls transcription of the *spiCAB* operon.

## DISCUSSION

The virulence phenotypes of several enteric pathogens have often been attributed to the presence of DNA segments that are absent from the genomes of nonpathogenic strains. These pathogenicity islands are largely responsible for the virulence properties of enteropathogenic and uropathogenic strains of *E. coli* (2, 3). In *Salmonella*, many of the invasion determinants

Α											
SpiR	371LE	NKV	A	ERTQALNE	AKKRAERANK	RKSIHLTVIS	* HELRTPMNGV	LGAIELLQ	TTPLNIEQQG	LADTARNCTL	442
BarA	247LAAYH	EEMQHNIDQA	TSDLRETLEQ	MEIQNVELDL	AKKRAQEAAR	IKSEFLANMS	HELRTPLNGV	IGFTRLTL	KTELTPTORD	HLNTIERSAN	339
RcsC		VHSRYRNENV									500
LemA		EELQLSIDQA									321
BvgS Etrl		SLGELKGIIG ILEE									767 390
Etri	307. LSHAA	IPPE	SMRARDL	TWEGNAUDT	AKKEAETAIK	ARNDFLAVMN	HEMRTPMHAI	TAL. SSEDQ	ETELTPEQKL	MVETILKSSN	390
SpiR		DFSRIESGHF									537
BarA		DFSKLEAGKL									439
RcsC		DFSKIESEQL									595
LemA BvgS		DFSKIEAGKL DIAKIEAGKF									421 866
Etrl		DLSRLEDGSL									490
SpiR		SDSGKGIEIQ									629
BarA		RDTGIGIPER									538
RcsC LemA		RDTGVGIPAK								LNIP.LGGLR	692 519
BvgS											947
Etrl		KDSGAGINPQ									565
a. in						*					700
SpiR BarA		VDDADINRDI VDDNPANLKL									780 757
ResC		VDDHPANLKL									896
LemA		VDDNPANLLL									749
BvgS		VDDHKPNLML									1067
Etrl	593LKVLV	MDENGVSRMV	TKGLLVHLGC	EVTTVSSNEE	CLRV.VSHEH	KVVFMDVCMP	GVENYQIALR	IHEKFTKQ	RHQRPLLVAL	SGNTDKSTKE	684
Cnip	DOVENETURY	* ITKPVTLATL	NEWTOTAN //	808							
SpiR BarA		LAKPIEEERL		785							
ResC		LSKPVTLDVI									
LemA		LTKPISERQL		777							
BvgS	ACRAAGMDDC	LFKPIGVDAL	RQRLNEAV//	1095							
Etrl	KCMSFGLDGV	LLKPVSLDNI	RDVLSDLL//	712							
В											
	MUUN		INTAKOD	FIGHECHDER	I VARONDI NE	UT UT I CENVD		M POCKIDD	CREVITINAL	ANOVINTIMUP	9.0
SpiA		KRLILILLFI SFFIFTALFC									90 93
	MKKI	KRLILILLFI SFFIFTALFC ARVLACAALV	CSAQAAPSSL	EKRLGKNEYF	IITKSSPVRA	ILNDFAANYS	IPVFISSSVN	DD.FSGEIKN	EKPVKVLEKL	SKLYHLTWYY	90 93 94
SpiA SepC	MKKI MKTHILL	SFFIFTALFC	CSAQAAPSSL LVTPGYSSEK	EKRLGKNEYF IPVTGSG	IITKSSPVRA FVAKDDSLRT	ILNDFAANYS FFDAMALQLK	IPVFISSSVN EPVIVSKMAA	DD.FSGEIKN RKKITGNFEF	EKPVKVLEKL HDPNALLEKL	SKLYHLTWYY SLQLGLIWYF	93
SpiA SepC InvG	MKKI MKTHILL MKKFNIK	SFFIFTALFC ARVLACAALV	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS	EKRLGKNEYF IPVTGSG HLLEQNDIAK	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS	ILNDFAANYS FFDAMALQLK FFERFSALLN	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA	DD.FSGEIKN RKKITGNFEF KKRISGEFDL	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK	93 94
SpiA SepC InvG MxiD YscC	MKKI MKTHILL MKKFNIK MAFPLHSFFK	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY	93 94 97 96
SpiA SepC InvG MxiD YscC SpiA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH SVLDNALI	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYF ASLYNLVWYY IYTLKYATAM	93 94 97 96
SpiA SepC InvG MxiD YscC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYKT	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF NEISRSIITP	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH SVLDNALI ESLDKEAQ	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT	93 94 97 96 188 191
SpiA SepC InvG MxiD YscC SpiA SepC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYKT DGQAIYIYDA	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG	93 94 97 96
SpiA SepC InvG MxiD YscC SpiA SepC nvG	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYKT DGQAIYIYDA DGNALYIYDS	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG	SKLYHLTWYY SLQLGLIWYP TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS	93 94 97 96 188 191 191
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYKT DGQAIYIYDA DGNALYIYDS DGNVLYIFKN	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP NRLVYVSGPP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA	DD.FSGEIKN RKKITGNFEF KKRISGEPDL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS	93 94 97 96 188 191 191 194 195
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC øiA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYKT DGQAIYIYDA DGNALYIYDS DGNVLYIFKN DTQYQYRDQS	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q	DD.FSGEIKN RKKITGNFEF KKRISGEFDL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DFRQNAVIVR	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR	93 94 97 96 188 191 191 194 195 256
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYYPA DGNLYIYDA DGNALYIYDS DGNVLYIFKN DTQYQYRDOS DITYKYRQQN	SFFIFTALFC ARVLACAALV SLTLLIVILP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEPNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK PLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE	IITKSSPVRA FVAKDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP DKTFYISGPP NRLVYVSGPP	IPVFISSSVN EPVIVSKAAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q	DD.FSGEIKN RKKITGMFEF KKRISGEFDL DK.VSGOFEH SVLDNALI ESLDKEAQ TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ	93 94 97 96 188 191 191 194 195
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC	MKKI MKTHILL MKKFNIK MAFPLHSPFK DGSMLYVYPA DENILYIYTD DGNALYIYDS DGNALYIYDS DGNUYIFKN DTQYQYRDQS DITYKYRQN DTTYLRQK	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL	CSAQAAPSSL LVTPGYSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP KTMASNGSLP	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP	IITKSSPVRA FVARDDSLRT YVAQSDTVGS PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSSQEALKQN	DD.FSGEIKN RKKIISGEPEL KKRISGEPEL SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK YPDTNSLUK	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYF ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE	93 94 97 96 188 191 191 194 195 256 263
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG	MKKI MKTHILL MKKPIK DGSMLYVPA DENILYVTA DGNLYYVA DGNUYIFKN DTVQYRDOS DITYKYRQON DRTYNLROOK DRTYNMRGED	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL	CSAQAAPSSL LVTFQYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEEQP ERLLN	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEP .GKALSNRQA	IITKSSPVRA FVARDDSLRT YVAQSDTUGS YVARGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPPFNI	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA ATVVVSKIAN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q N MSLQEALEAN FSF.SSVTNS	DD.FSGEIKN RKKITGNFEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK YPDTNSLLVK	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQUQIIR	93 94 97 96 188 191 191 194 195 256 263 291
spiA SepC InvG MxiD SepC SpiA SepC nvG xiD scC xiD scC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYUYPA DGNALYIYDS DGNVLYIFKN DTQYQYRQGS DITYKYRQON DRTYNLRDQK DRTYNLRDQK DRTYNMRGED	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSUL VVVPGVVSVL VVVPGVASVL IVIPGIATAI IVIPGVATVV VAAPGVATL	CSAQAAPSSL LVTPGYSEK LIVNANNIDS SSYSWAQ NILSTGRPIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEEQP ERLLNN QRVLSDATIQ	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STCKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP	IITKSSPVRA FVARDDSLRT VVAGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPPFNI QAA	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP KGKAANYAGG TQKVSEDSND	IPVFISSSVN EPVIVSKAAA ATVVVSKQAA ATVVVSKQAA ECIKYITSLS VYVDMVVNAA ALVELVANTA ALVELVANTA NSLQEALKQN FSF.SSVTNS T	DD.FSGEIKN RKKITGRFEPL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQOTQIR ALPMFAA AFONYTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA	EKPVKVLEKL HDPNALLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK YPDTNSLLVK	SKLYHLTWYY SLQLGIIWYF ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ	93 94 97 96 188 191 191 194 256 263 291 288 269
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVYPA DENILYIYTT DGQAIYIYDA DGNALYIYDS DGNVLYIFKN DTYYLXFRQN DTYYLXFQQ DRTYLNRQC DRTYNMRGED DRTHYRDDE KLITELDQQQ	SFFIFTALFC ARVLACAALV SLTLLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVATVL VAPGVATIL QMIEISVKII	CSAQAAPSSL LVTFQYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVNAGDINQL	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL	IITKSSPVRA FVARDDSLRT YVAQSDTUGS VPAGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP KGKAANYAGG TQKVSEDSND	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALKQN FSF.SSVTNS T STVIS	DD.FSGEIKN RKKITGNFEF KKRISGEPGL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTDKVNFG DPRQNAVIVR DPRLNAVVVK YPDYNSLLVK YPDYNSLLVK DPSLNAIIVR ALEKSSQAYV	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYF ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN	93 94 97 96 188 191 194 195 256 263 291 288 269 350
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC piA epC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVPA DGNIYIYAD DGNIYIYAD DGNIYIYAD DTYYRADON DRTYNLRDQK DRTYMRGED DRTIHYRDDE KLITELDQRQ QLISELDIEQ	SFFIFTALFC ARVLACAALV SLTLLIVILP RVLTGTLLLL SLLKHQVTFF NEISRSITFP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSIL MVIPGIATAI UVIPGVATVV VAAPGVATVL QMIEISVKII RQIEISVSII	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEEQP ERLLQGEEQP ERLLN QRVLSDATIQ DVNAGDINQL DVDANDLQQL	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWCTAVSL GVNWSGTLNA	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVARGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPFNI QAA GGK.KIAFNT GQG.TIAFNS	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KKGTFYVSGPP DKTFYISGPP NRLVYVSGPP KGKAANYAGG TQKVSEDSND GLNDGGASGF STAQVNIS.	IPVFISSSVN EPVIVSKMAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q N MSLQEALKQN MSLQEALKQN SSVIS	DD.FSGEIKN RKKITGNPEF KKRISGEPDL DK.VSGOPEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFONSVTISA AAAGNIKIVA RASAQARVEA DTSNFMVRLN NASNFMIRVN	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK DPSLNAIIVR ALEKSSQAYV ALQQNSKAKI	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIL DSPERMPMYQ LSQPSVVTLN LSQPSIITLN	93 94 97 96 188 191 191 194 195 256 263 291 288 269 355
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA	MKKI MKTHILL MKKFNIK MAFPLHSPFK DGSMLYUYAA DENILYIYKT DGNALYIYDS DGNVLYIFKN DTQYQYRDQS DITYKYRDQN DRTYNLRDQK DRTYNLRDQK DRTIHYRDDE KLITELDQRQ MLVKALDVAK	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVVSUL VVVPGVSVL VVVPGVSVL VVVPGVSIL MVIPGIATAI IVIPGVATVV VAPGVATIL QMIEISVKII RQIEISVSII RHVELSLWIV	CSAQAAPSSL LVTPGYSEK SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP ERLLQGEEQP ERLLQGEEQP ERLLQGETQ QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNXGDIRL	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE S.STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDQRIP GIDWGTAVSL GIDWGTAVSL	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVAKGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PAT AMPAFSANGE QNDMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN ILTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP NRLVYVSGPP NRLVYVSGPP KGKAANYAGG TQKVSEDSND 	IPVFISSSVN EPVIVSKAAA YPIVVSKQAA ATVVVSKQIA SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA ALVELVANTA NSLQEALKQN FSF.SSVTNS T STVIS STVIS STL	DD.FSGEIKN RKKITGRFEPL DK.VSGQFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQOTQIR ALPMFAA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFWVRLN NASNFMIRVN NGSRFIAAVN	EKPVKVLEKL HDPNALLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK ADPLNAIVVR ALQONSKAKI ALGENKQATV	SKLYHLTWYY SLQLGIIWYF ASLYNLWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVITLN VSRPVLTQE	93 94 97 96 188 191 194 195 256 263 291 288 269 350
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG piA epC nvG	MKKI MKTHILL MKKFNIK MAFPLHSPFK DGSMLYVYPA DENILYIYTT DGQAIYIYDS DGNUYIFKN DTYVLYFKN DTYVLYFKN DTYVLYFKN DRTYNNRGED DRTIHYRDDE KLITELDQRQ QLISELDIEQ MLVKALDVAK DITYQLDVAK	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL QMIEISVAII QMIEISVKII RQIEISVSII RHVELSLWIV	CSAQAAPSSL LVTPGYSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DLNKSDLERL DLDKSDLENNL	EKRLGKNEYF IPVTGSG HLLEQNDTAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWGGTLNA GTSWSGTINA	IITKSSPVRA FVARDDSLRT YVAQSDTVGS PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGRWPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV	ILNDFAANYS FFDAMALQLK FFERFSALLN ILTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP  KGKAANYAGG TQKVSEDSND  GLNDGGASGF STAQVNIS SLQSSI SFNMSSSASI	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALKQN FSF.SSVTNS T STVIS SSVIS STL STL	DD.FSGEIKN RKKITGNFEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSNFIASVM	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK ALEKSSQAYV ALQONSKAKI ALEEKKQATV	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIL DSPERMPMYQ LSQPSVVTLN LSQPSIITLN	93 94 97 96 188 191 191 194 195 256 263 291 288 269 350 355 377
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYVPA DGQAIYIYDA DGQAIYIYDA DGQAIYIYDA DGQAIYIYDA DGVLYIFKN DTQYQYRDQS DITYNYRDQN DRTYMRGED DRTYMRGED DRTIHYRDDE KLITELDQRQ QLISELDIEQ MUVKALDVAK RLIHALDKPS	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVTFF NEISRSITFP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSIL MVIPGIATAI IVIPGVATVV VAAPGVATVL QMIEISVKII RQIEISVSII RHVELSLWIV RHIELSLWII ARIEVALSIV	CSAQAAPSSL LVTPGYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEEQP ERLLQGEEQP ERLLN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNASDLERL DINSDLERL DINSDLERL	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP GIDWGTAVSL GVNWSGTLNA GTSWSGTITI GVNWQGTASF GVDWRVGIRT	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVARGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNNHQVVIKT	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP KGKAANYAGG TQKVSEDSND GLNDGGASGF STAQVNIS SLNQSSI SFNMSSSASI TGDQSNIASN	IPVFISSSVN EPVIVSKMAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q N MSLQEALKQN MSLQEALKQN STVISSTVIS STLSTL GALGSLVDAR	DD.FSGEIKN RKKITGNPEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIKVN DGSRFIAAVN DGNKFIASVM GLDYLLARVN	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK DPSLNAIIVR ALEKSQAYV ALQONSKAKI ALEEKQATV ALNQKKANV LLENEGSAQV	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIL DSPERMPMYQ LSQPSVVTLN LSQPSVITLN VSRPVLLTQE VSRPVLLTQE	93 94 97 96 188 191 191 194 195 256 263 291 288 269 350 355 377 377 369
spiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYUYAD DGNALYIYDS DGNULYIFKD DTQYQYRQOS DTTYKNRQON DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTIHYRDDE KLITELDQRQ QLISELDIEQ MLVKALDVAK RLIHALDVAK	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVVSLL WVPGVVSLL WVPGVVSLL WVPGVSVL UVPGVATVL VAAPGVATIL QMIEISVKII RQIEISVSII ARIEVALSIV TFYTKLQGEK	CSAQAAPSSL LVTFQYSEK LIVNANNIDS SSYSWAQ NILSTGRPIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVDAGDIQL DVDAGDIQL DVDAGDIQL DUNASDLFRL DIDKSELNNL DINADQLTEL VAKLESITTG	EKRLGKNEYF IPYTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS UGNIVSEPP .GKALSNRQA QVTVDNQRIP GIDWCTAVSL GVNWGGTINI GVNWGGTASF GVNWQGTASF GVDWRVGIRT	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVAQSDTVGS YVAQSDTVGS NSCNURKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKIGV GDSFGA GNNHQVVIKT	ILNDFAANYS FFDAMALQLK FFERFSALLN TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q M MSLQEALKQN FSF.SSVTNS T STVIS STL GALGSLVDAR IQDGQQSDTQ	DD.FSGEIKN RKKITGNFEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSRFIAAVN DGSRFIAAVN GLDYLLARVN SETD	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL SNPEEMLEKL DNPQDFJQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK ALEKSQAYV ALQONSKAKI ALEKKSQAYV ALQONSKAKI ALEREGSAQV PLPEVQNS	SKLYHLTWYY SLQLGLIWYF ALLVGLIWYK ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVUTLN LSQPSVUTLN LSQPSIITLN VSRPVLLTQE VSRPVLLTQE EIASQATLLA	93 94 97 96 188 191 191 194 195 256 263 291 288 269 350 355 377 377 369 435
spiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA	MKKI MKTHILL MKKTHIK MATPLHSFK DGSMLYVYPA DGNALYIYDS DGNVLYIFKN DGNVLYIFKN DTYYLXPQN DRTYHLRDQK DRTYHNRGED DRTYHYRDDE KLITELDQRQ QLISELDIEQ MLVKALDVAK RLIHALDKPS NIQAVLDKNI	SFFIFTALFC ARVLACAALV SLTLLIVILP RVLTGTLLLL SLLKHQVTFF MEISRSITFP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVATVL VVVPGVATVL VAPGVATIL QMIEISVKII RVELSLWIV ARIEVALSIV TFYTKLQGEK TFYTKLQGEK	CSAQAAPSSL LVTFQYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNKSDLERL DINKSDLERL DINKSDLERL DINKSDLERL DINKSDLERL VAKLESITTG VASLRSTSG	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWSGTLNA GTSWSGTITI SLLRVTPRLL TLLEVTPRLL	IITKSSPVRA FVARDDSLRT YVAQSDTUGS VGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGRWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKFGA GNNHQVVIKT NDNGT DDSSNSLTGK	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALKQN FSF.SSVTNS T STVIS STLI GALGSLVDAR IQDGQQSDTQ IQDGNQSTNQ	DD.FSGEIKN RKKITGNFEF KKRISGEPGL SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSRFIAAVM GLDYLLARVN SETD SNAQDA	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRLNAVVVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK ALEKSQAYV ALQONSKAKI ALEKKQATV LLENEGSAQV PLPEVQNS SSTLPEVQNS	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYF ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN LSQPSVITLN VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA	93 94 97 96 188 191 191 194 195 256 263 291 288 269 355 377 369 435
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA	MKKI MKTHILL MKKPIKK MAFPLHSFFK DGSMLYVPA DGNLYIYDA DGNLYIYDA DGNUYIFKN DTYYRAQM DRTYNLROK DRTYNRCED DRTYHYRDDE KLITELDQRQ QLISELDIEQ MUKALDVAK DITYLDDAK NIQAVLDKNI NIQAVLDKNI	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVTFF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVSTL NVIPGLATAI IVIPGVATVV VAAPGVATVL VAAPGVATVL VAAPGVATVL TPYTKLOGEK TFYTKLOGEK TFYTKLIGER	CSAQAAPSSL LVTFQYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEEQP ERLLQGEEQP ERLLN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNASDLERL DINADQLTEL VAKLESITTG VASLRSITSG NVALEHVTYG	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEPP GIDWGTAVSL GVNWSGTLNA GTSWSGTITI GVNWGGTASF GVDWVGIRT SLLRVTPRIL TLLRVTPRIL TMIRVLPRFS	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVARGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNNHQVVIKT NDNGT DDSSNSLTGK ADG	ILNDFAANYS FFDAMALQLK FFPERFSALLN LLTDFGANYD TKAVEVSGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP  KGKAANYAGG TQKVSEDSND  SLNGGASGF STAQVNIS SLNGSSI SFNMSSSASI TGDQSNIASN QKIMLNIN RRERVRLLD O[EMSLD	IPVFISSSVN EPVIVSKNAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q N MSLQEALKQN MSLQEALKQN STLSTLSTLSTLG GALGSLVDAR IQDGQQSDTQ IQDGQQSTQQ IEDGNDKTPQ	DD.FSGEIKN RKKITGNPEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGNKFIASVM GLDYLLARVN SETD SNAQDA SDTTTS	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK VPDTNSLLVK YPDTNSLLVK YPDTNSLLVK YPDTNSLLVK VPDTNSLLVK DPSLNAIIVR ALEKSQAYV ALQQNSKAKI ALEKKQATV ALQQNSKAKI ALEKKQATV LLENEGSAQV PLPEVQNS SSTLPEVQNS	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GMDQIQIIR DSPERMPMYQ LSOPSVVTLN LSOPSVITLN VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA	93 94 97 96 188 191 194 195 256 263 291 288 269 355 377 377 369 435 451 463
spiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYUYA DGNALYIYDS DGNULYIFXN DTQYQYRQGS DTTYKNRQQN DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK RLIHELDQRQ QLISELDIEQ MLVKALDVAK RLIHALDVAK RLIHALDVAS NIQAVLDKNI NVQAILDNNV NYPAIFDNNR	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVSVL VVVPGVSVL MVIPGIATAI IVIPGVATVV VAAPGVATIL QMIEISVKII RQIEISVSII ARIEVALSIV TFYTKLQEK TFYTKLGEK TFYTKLGEK	CSAQAAPSSL LVTPGYSEK LIVNANNIDS SSYSWAQ NILSTGRPIH TYLDIDSLLK RNVSLNEPNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP ERLLQGEEQP ERLLQGEEQP ERLLNN QRVLSDATIQ DVDANDLQQL DVDANDLQQL DINKSDLFRL DINKSDLFRL DINADQLTEL VAKLESITTG VASLESITTG VASLENTYG	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS U.GNLSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWSGTINA GVNWSGTINA GVNWSGTINT SLLRVTPRLL TLLRVTPRLL TLINVIPRFS	IITKSSPVRA FVARDDSLRT YVAQSDTVGS YVAQSSLRD PGCEVKEITG NSCNVRKITT KNYPLRGONR HRYPLRGONR HRYPLRGNIS PRFGWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDSFGA GNNHQVVIKT NDNGT DDSSNSLTGK ADG SRG	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKQIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA 	DD.FSGEIKN RKKITGRFEF KKRISGEFDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN GLDYLLARVN SETD SNAQDA SDTTS	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLLVK YPDTNSLVVK YPDTNSLVK Y	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYF ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN LSQPSVITLN VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA	93 94 97 96 188 191 191 194 195 256 263 291 288 269 355 377 369 435
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA ScC NvG xiD ScC NvG NvG NvG NvG NvG NvG NvG NvG NvG NvG	MKKI MKTHILL MKKPIK DGSMLYVPA DGSMLYVPA DGNILYIYT DGQAIYIYDA DGNVLYIFKN DTYVYRQQN DRTYNLRQCK DRTYNMRGED DRTIHYRDDE KLITELDQRQ QLISELDIEQ MLVKALDVAK DIITQLDVAK NIQAVLDKNI NNQAILDKNV NVPAIPDNNR NIQAVLDKNI	SFFIFTALFC ARVLACAALV SLTLLIVLLP RVLTGTLLLL SLLKHQVTFP MEISRSITFP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVSVSIL MVIPGLATAI IVIPGVATVV VAAPGVATUL QMIEISVKII RQIEISVSII RHVELSLMIV ARIEVALSIV TFYTKLQEK TFYTKLQEK TFYTKLGER TFYTSLVGER TYYVKVGK	CSAQAAPSSL LVTFQYSSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLN QRVLSDATIQ DVNAGDINQL DUNAGDINQL DUNAGDINQL DUNAGDINQL DUNAGDINQL DINKSDERL DINADQLTEL VAKLESITTG VASLRSITSG NVALEHVTYG NSSLEHVTYG VAELKGITYG	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS STGKGAVE LGNIVSSEP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWSGTLNA GTSWSGTINA GTSWSGTINA GTSWSGTINA SLLRVTPRLL TLLRVTPRLL TMIRVLPRFS TLLNVIPRFS TLINVIPRFS	IITKSSPVRA FVARDDSLRT YVAQSDTUGS YVARGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GDSFGA DDSSNSLTGK ADG SRG TQGDKS	ILNDFAANYS FFDAMALQLK FFERFSALLN LLTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP 	IPVFISSSVN EPVIVSKNAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALVEQTA SVISS STLS STLSS STLS STLS TLS TLS IQDGQQSDTQ IQDGQQSTQ IEDGNDKTPQ IEDGNDKTPQ IEDGNQKPNS	DD.FSGEIKN RKKITGRFEF KKRISGEFDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN GLDYLLARVN SETD SNAQDA SDTTS	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLLVK YPDTNSLVVK YPDTNSLVK Y	SKLYHLTWYY SLQLGIIWYF ALLYGLIWYK ASLYNLWWY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVTLM LSQPSVTLIN VSRPVLLTQE VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPQ	93 94 97 96 188 191 191 194 195 256 263 291 288 269 355 377 377 369 435 463 465 454
spiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA SepC NvG xiD scC SpiA SepC SpiA SepC SpiA SepC SepC SepC SepC SepC SepC SepC SepC	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYUYFA DGNALYIYDS DGNVLYIFKN DTQYQYRQS DTTYKYRQN DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTYNLRDQC MLYKALDVAK DITYLLFLDQRQ MLVKALDVAK DIITQLDVAK RLIHALDVAS NIQAVLDKNI NNQALDKNY NNPAIFDNNR NAQAVIDHSE GQSLLLGGFK	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVVSVL VVVPGVVSVL VVVPGVSVL MVIPGIATAI IVIPGVATVV VAAPGVATIL QMIEISVKII RVELSLWIV RHIELSLWIV RHIELSLWIV RHIELSLWIV TFYTKLQGEK TFYTKLGER TFYTVLGER TFYTVLGER TFYTVLGEK	CSAQAAPSSL LVTFQYSEK LVTRANNIDS SSYSWAQ NILSTGRPIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ QESEAAELKQ ERLLQGEEQP ERLLQGEEQP ERLLNN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNASDLERL DINSDERL DINSDERL DINADQLTEL VAKLESITTG VASLESITTG VASLENTYG VAELKGITYG IPLLGDIFVV	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE .STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GIDWGTAVSL GUDWGTASF GVDWRVGIRT SLLRVTPRLL TLLRVTPRLL TLLRVTPRLL TLLRVTPRLS TLLNVIPRFS TMLRWTPRVL GHLFRNDTTQ	IITKSSPVRA FVARDDSLRT VVAGESLRD PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT AMPAFSANGE QNDMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GGKLGV GDSFGA GNNHQVVIKT NDNGT DDSSNSLTGK ADG SRG TQGDKS	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN FNSIEVRGVP KGTFYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGP STAQVNIS SINQSSI SFNMSSSASI TGDQSNIASN QKIMLNIN RREKVRLLD QIEMSLD QIEMSLT EISLNLH KASVVNNGIS	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKQIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVATA NSLQEALKQN FSF.SSVTNS T STVIS STL GALGSLVDAR IQDGQQSDTQ IQDGQQSDTQ IEDGNDKTPQ IEDGNDKTPQ IEDGNDKTPS IEDGNQKPNS	DD.FSGEIKN RKKITGRFEF KKRISGEFDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN GLDYLLARVN SETD SNAQDA SDTTS	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKAKNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLLVK YPDTNSLVVK YPDTNSLVK Y	SKLYHLTWYY SLQLGIIWYF ALLYGLIWYK ASLYNLWWY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVTLM LSQPSVTLIN VSRPVLLTQE VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPQ	93 94 97 96 188 191 191 194 195 256 263 291 288 269 350 355 355 357 377 369 435 451 465 454
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA SepC nvG xiD ScC SpiA	MKKI MKTHILL MKKFNIK MAFPLHSFFK DGSMLYUYPA DGNLYIYTD DGNLYIYDS DGNUYIFKN DTYVKYRQN DTYVLYFKN DRTYNNRGED DRTYHYRDDE KLITELDQRQ OLISELDIEQ MUVKALDVAK DIITQLDVAK RLIHALDKPS NIQAVLDKNI NHQAILDKNI NHQAILDKNI NHQAILDKNI NHQAILDKNI NHQAILDKNI NHQAILDKNI SCALLIGGFK GQSLLLGGFK	SFFIFTALFC ARVLACALV SLTLLIVLP RVLTGTLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSL VVVPGVVSL VVVPGVVSL VVVPGVVSL QVVFGVSVL QMIEISKVII RQIEISVSII RQIEISVSII RUFLSLVIV TFYTKLQEK TFYTKLQEK TFYTKLGEK TFYVKLGEK TFYVKLGEK CMICLON	CSAQAAPSSL LVTPGYSEK LIVNANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DVNAGDINQL DUNKSDLRRL DINKSDLRRL DINKSELNNL DINADQLTEL VAKLESITTG VASLRSITSG NVALEHTYG VASLKGITYG IPLEGDIPVV IPLLSDIPVI	EKRLGKNEYF IPYTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKO.ANLYD AL.QRSGIWE TSSTNNGS UGNUSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWGGTASF GVDWRVGIRT SLLRVTPRLL TLLRVTPRIL TLINVTPRFS TLINVIPRFS TLINVIPRFS	IITKSSPVRA FVARDDSLRT YVAQSDTVGS PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNNHQVVIKT NDNGT DDSSNSLTGK ADG TQGDKS YHSVIRLFLI	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGPP NRLVYVSGPP  KGKAANYAGG TQKVSEDSND  GLNDGGASGF STAQVNIS SINQSSI SFNMSSSASI TGDQSNIASN QKIMLNLN RRERVRLLD OIEMSLD OIEMSLD EISLNLH KASVVNNGIS KATPIKSAS	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALKQN FSF.SSVTNS STL GALGSLVDAR IQDGQQSDTQ IEDGTQSTNQ IEDGRNDKTPQ IEDGRNDKTPQ IEDGRNDKTPQ IEDGRNDKTPQ IEDGRNQKPNS HG E	DD.FSGEIKN RKKITGNFEF KKRISGEPDL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFIIAVN DGSKFIASVM GLDYLLARVN SETD SNAQDA SDTTTS NYNYNKENS SGIE.	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL SNPEEMLEKL DNPQDFJQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK YPDTNSLUK ALEKSQAYV ALQONSKAKI ALEEKQATV ALQONSKAKI ALEEKQATV LLENEGSAQV PLPEVQNS SSTLPEVQNS VDALPEVQRT GIPTISRT	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVUTLN LSQPSVUTLN USRPVLLTQE VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPH VVDTVARVGH	93 94 97 96 188 191 191 194 256 263 291 288 269 350 355 377 377 369 435 451 463 465 454 498 512
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA SepC InvG SepC InvG SepC SpiA	MKKI MKTHILL MKKPIK MAFPLHSFFK DGSMLYVYPA DGNALYIYDS DGNVLYIFKN DGNVLYIFKN DTYYLRDQN DRTYNLRDQK DRTYNMRGED DRTYNMRGED DRTYNMRGED DRTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNMRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNNRGED RTYNRGED RTYNNRGED RTY	SFFIFTALFC ARVLACALV SLTLLIVLLP SLTLLIVLLP RVLTGTLLLL SLLKHQVTF NEISRSITTP SEMRNAVVSL GELISKVILL SEVASRLIRL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVVSVL VVVPGVATVL VVPGVSVSL QVVFGVATVL VAPGVATVL VAPGVATVL VAPGVATVL TFYTKLGEK TFYTKLGEK TFYTKLGER TFYVSLVGER TYYVKVTGKE QGKQIHSQNK KIKKVPSKDG RDANTDTVQS	CSAQAAPSSL LVTFQYSSEK LIVTANNIDS SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ REM.SKTSVP KTMASNGSLP ERLLQGEQP ERLLNN QRVLSDATIQ DVNAGDINQL DVNAGDINQL DUNAGDINQL DUNAGDINQL DUNAGDINQL DUNAGDINQL VALKESITTG VASLRSITSG NVALEHVTYG NSSLEHVTYG VAELKGITYG IPLLGDIPVI IPLLGDIPVI	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE TSSTNNGS CSTGKGAVE LGNIVSSEP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GVNWSGTLNA GTSWSGTIN GVNWGGTASF GVDWRVGIRT SLLRVTPRLL TMLRWTPRLL TMLRWTPRL GHLFRNDTTQ GSLFSSTVKQ GSLFSSTVKQ	IITKSSPVRA FVARDDSLRT YVAQSDTUGS PGCEVKEITG NSCNVRKITT KNYPLRGDNR HRYPIRGNIS PRFGWRPDAS PAT RSG AMPAFSANGE QNDPMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNNHQVVIKT NDNGT DDSSNSLTGK ADG SRG YGQGKS	ILNDFAANYS FFDAMALQLK FFERFSALLM ILTDFGANYD TKAVEVSGVP FNSIEVRGVP KGTFYVSGPP DKTFYISGP NRLVYVSGPP 	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSDKIN SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVANTA RYLELVEQTA Q Q MSLQEALKQN FSF.SSVTNS T STLIS GALGSLVDAR IQDGQQSDTQ IQDGQQSDTQ IEDGRDKTPQ IEDGRDKTPQ IEDGRDKTPQ IEDGRDKTPS IEDGRDKTPS	DD.FSGEIKN RKKITGNFEF KKRISGEPEL SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQQTQIR ALPMFAA LFDNSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFMVRLN NASNFMIRVN DGSNFIAAVN DGSNFIAAVN GLDYLLARVN SETD SNAQDA SDTTTS NYNYNNENTS SGIE.	EKPVKVLEKL HDPNALLEKL SNPEEMLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG SEKTGALAIE DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK YPDTNSLVK ALQONSKAKI ALEKSQAYV PLPEVQNS SSTPEVQNS VDALPEVGRT GIPTISRT	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN LSQPSVITIN VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPH KISTIARVPH KISTIARVPH VSTVARVGH	93 94 97 96 188 191 191 194 195 256 263 291 288 269 355 377 369 435 451 463 465 454 498 512 552
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA SepC nvG xiD ScC SpiA	MKKI MKTHILL MKKPNIK MAFPLHSPFK DGSMLYUYPA DENILYIYKT DGALYIYDA DGNALYIYDS DGNVLYIFKN DTYYKROON DRTYNLRDQK DTYYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK NITYLELDQQK QLISELDIQQ MLVKALDVAK DIITQLDVAK RLHALDVAS NIQAVLDKNI NMQAVLDKNI NNQALDKNY NYPAIFDNNR NAQAVIDISE GQSLLLGGFK GKSLLUGGFT GKSLLUGGYT	SFFIFTALFC ARVLACALV SLITLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVVSLL VVVPGVVSLL MVIPGIATAI IVIPGVATVV VAPGVATIL QMIEISVKII RVELSLWIV RHIELSLWIV RHIELSLWIV RHIELSLWIV RHIELSLWIV RTFYTKLQGEK TFYTKLIGER TFYTKLIGER TFYVSLVGEK QGKQIHSQNK KIKKVPSKDG RDANTDTVQS	CSAQAAPSSL LVTFQYSEK LVTFQYSEK SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ QESEAAELKQ QRUSDATIQ DVNAGSLNU DVNAGDINQL DVNAGDINQL DVNAGDINQL DINX5DLERL DINX5LENIL DINADQLTEL VASLESITTG VASLESITSG VASL	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE ATSSTNNGS STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GIDWGTAVSL GIDWGTAVSL TLLRVTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRL	IITKSSPVRA FVARDDSLRT YVAGSSLRD PGCEVKEITG NSCNVRKITT KNYPLRGNR HRYPIRGNS PAT AMPAFSANGE QNDMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNHQVVIKT NDNGT SRG SRG YQGKS VHSVIRLFLI KSNVVRVFLI ISNIVRVFLI SNIVRVFLI	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN FFERFSALLN FNSIEVRGVP KGTFYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGP NRLVYVSGP STAQVNIS. STAVNIS. STAV	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKQIA SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVATA NSLQEALKQN FSF.SSVTNS T STVIS STL GALGSLVDAR IQDGQQSDTQ IQDGNQSTNQ IEDGNDKTPQ IEDGNDKTPQ IEDGNDKTPA IEDGNDKTPA IEDGNDKTPA IEDGNDKTPA	DD.FSGEIKN RKKITGRFEP KKRISGEFPL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQOTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFWVRLN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN SUVNILKQSG SEREIQKTTQ	EKPVKVLEKL HDPNALLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS	SKLYHLTWYY SLQLGIIWYF ALUGLIWYF ASLYNLVWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN LSQPSVVTLN USRPVLLTQE VSRPVLLTQE VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPH KISTIARVPH KISTIARVPA VVDTVARVGH	93 94 97 96 188 191 191 194 256 263 291 288 269 350 355 377 377 369 435 451 463 465 454 498 512
SpiA SepC InvG MxiD YscC SpiA SepC nvG xiD scC piA epC nvG xiD scC piA epC nvG xiD scC SpiA SepC InvG MxiD YscC	MKKI MKTHILL MKKPNIK MAFPLHSPFK DGSMLYUYPA DENILYIYKT DGALYIYDA DGNALYIYDS DGNVLYIFKN DTYYKROON DRTYNLRDQK DTYYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK DRTYNLRDQK NITYLELDQQK QLISELDIQQ MLVKALDVAK DIITQLDVAK RLHALDVAS NIQAVLDKNI NMQAVLDKNI NNQALDKNY NYPAIFDNNR NAQAVIDISE GQSLLLGGFK GKSLLUGGFT GKSLLUGGYT	SFFIFTALFC ARVLACALV SLITLIVLP RVLTGTLLLL SLLKHQVITF NEISRSIITP SEMRNAVVSL GELISKVILL VVVPGVVSLL VVVPGVVSLL MVIPGIATAI IVIPGVATVV VAPGVATIL QMIEISVKII RVELSLWIV RHIELSLWIV RHIELSLWIV RHIELSLWIV RHIELSLWIV RTFYTKLQGEK TFYTKLIGER TFYTKLIGER TFYVSLVGEK QGKQIHSQNK KIKKVPSKDG RDANTDTVQS	CSAQAAPSSL LVTFQYSEK LVTFQYSEK SSYSWAQ NILSTGRFIH TYLDIDSLLK RNVSLNEFNN ENISLNYLIQ QESEAAELKQ QESEAAELKQ QRUSDATIQ DVNAGSLNU DVNAGDINQL DVNAGDINQL DVNAGDINQL DINX5DLERL DINX5LENL DINADQLTEL VAKLESITTG VASLESITTG VASLESITTG VASLENTYG VAELKGITYG IPLLGDIPVV IPLLSDIPVI IPFLGSLPVI	EKRLGKNEYF IPVTGSG HLLEQNDIAK ELDWLPIPYV YLRSQNILSS YLSDTISVNK FLKR.SGLYN YLKD.ANLYD AL.QRSGIWE ATSSTNNGS STGKGAVE LGNIVSSEPP .GKALSNRQA QVTVDNQRIP GIDWGTAVSL GIDWGTAVSL GIDWGTAVSL TLLRVTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRLTPRL	IITKSSPVRA FVARDDSLRT YVAGSSLRD PGCEVKEITG NSCNVRKITT KNYPLRGNR HRYPIRGNS PAT AMPAFSANGE QNDMPPFNI QAA GGK.KIAFNT GQG.TIAFNS GDKLGV GDSFGA GNHQVVIKT NDNGT SRG SRG YQGKS VHSVIRLFLI KSNVVRVFLI ISNIVRVFLI SNIVRVFLI	ILNDFAANYS FFDAMALQLK FFERFSALLN FFERFSALLN FFERFSALLN FNSIEVRGVP KGTFYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGPP NRLVYVSGP SILVGSNI SFNGSSASI TGDQSNIASN QIEMSLD QIEMSLD QIEMSLT EISLNLH KASVVNNGIS KATPIKSASS	IPVFISSSVN EPVIVSKMAA YPIVVSKQAA ATVVVSKQIA SCLTRISQLA ECIKYITSLS VYVDMVVNAA ALVELVATA NSLQEALKQN FSF.SSVTNS T STVIS STL GALGSLVDAR IQDGQQSDTQ IQDGNQSTNQ IEDGNDKTPQ IEDGNDKTPQ IEDGNDKTPA IEDGNDKTPA IEDGNDKTPA IEDGNDKTPA	DD.FSGEIKN RKKITGRFEP KKRISGEFPL DK.VSGOFEH SVLDNALI ESLDKEAQ TMMDKQND TLLDKQVS AALEQOTQIR ALPMFAA LFONSVTISA AAAGNIKIVA SILEDVSLIA RASAQARVEA DTSNFWVRLN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN DGSRFIAAVN SUVNILKQSG SEREIQKTTQ	EKPVKVLEKL HDPNALLEKL DNPQDFLQHI KRKDSAVSVS SKARNKDVVK GIELGRQKIG SIGTDKVNFG DPRQNAVIVR DPRLNAVVVK YPDTNSLLVK SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS SSTLPEVQNS	SKLYHLTWYY SLQLGLIWYF TLLVGLIWYK ASLYNLWWYY IYTLKYATAM VFKLNYASAT VMRLNNTFVG VIKLKNTFVS IFPLKYASAS DYAANMAGYR DREITMDIYQ GTAEQVHFIE GNDQQIQIIR DSPERMPMYQ LSQPSVVTLN LSQPSVITIN VSRPVLLTQE EIASQATLLA EMTTEATLSA LISTIARVPH KISTIARVPH KISTIARVPH VSTVARVGH	93 94 97 96 188 191 191 194 195 256 263 291 288 269 350 355 377 377 369 435 451 463 465 454 498 512 252 562

FIG. 2. Comparison of *Salmonella* Spi with related proteins. The deduced amino acid sequences of proteins encoded by the *S. typhimurium spi* island are shown in the single-letter code. (*A*) Comparison of SpiR to members of the two-component family. (*B*) Comparison of SpiA to components of type III secretion systems. Alignment was conducted using the PILEUP program (GCG). Asterisks indicate residues conserved in all members of this family. For any given comparison, residues present in the majority of related proteins are highlighted.

have been localized to a 40-kb segment of the chromosome specifying at least 25 genes encoding a secretion system, its effectors, and the regulators controlling their expression (9). Because these genes are ancestral to all eight subgroups of *Salmonella enterica*, it has been hypothesized that the acquisition of this segment was an essential step in the evolution of invasion by *Salmonella* (14). However, the presence of structurally and functionally equivalent genes in the *Shigella flexneri* virulence plasmid (22) indicates that this invasion island does not confer the specific properties that distinguish *Salmonella* from other enteric pathogens.

We have identified a new gene cluster essential for *Salmo-nella* virulence that includes the *spiCAB* operon, encoding components of a putative type III secretion apparatus, and a regulatory gene, *spiR*. Because the SpiA and SpiB proteins were similar to components of secretion apparatuses, we examined the profiles of proteins exported by *spi* and wild-type strains. The most conspicuous difference was the presence of

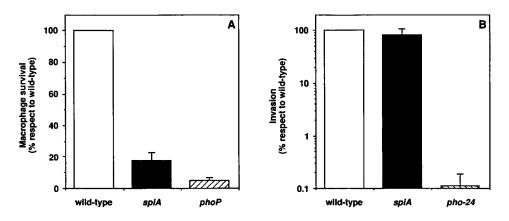


FIG. 3. Intramacrophage survival (A) and invasion (B) properties of an *spiA* mutant of S. *typhimurium*. Values represent the mean of triplicate samples  $\pm$  SD. Properties of *phoP* and *pho-24* strains are shown for comparative purposes.

two distinct forms of flagellin: 45 kDa in wild-type and 47 kDa in the *spi* mutants (Fig. 4). Thus, Spi may export a protease or a protein that modifies the activity of a secreted protease. However, flagellin is probably not the physiologically relevant substrate of this protease because flagellin is not essential for virulence and *spi* mutants are motile.

Although the Spi secretion system is reminiscent of that encoded within the *spa* invasion island of *Salmonella*, strains with mutations in the *spi* genes displayed wild-type levels of invasion (Fig. 3). Unlike *invA* mutants, which are only attenuated upon oral inoculation (28), the *spi* mutants failed to cause lethal infections in mice even when inoculated intraperitoneally. The *spi* genes were required for intramacrophage survival—a trait fundamental to *Salmonella* pathogenesis (29, 30)—and not surprisingly, homologous sequences have not been detected in other enteric species (6). Thus, Spi constitutes a novel type III secretion system that exports proteins that permit the survival of *Salmonella* within phagocytic cells, perhaps by modifying host factors required for phagosome– lysosome fusion (31) or phagosome acidification (32).

The *spi* gene cluster has an anomalous base composition of only 42.1% G+C, which is much lower than the overall G+C content of 52–54% of the *Salmonella* genome (6), and suggests that it was acquired by horizontal transfer. Regions acquired through horizontal processes often harbor the genes encoding both structural and regulatory elements in a contiguous DNA segment (9, 33); and within the 5.7-kb region analyzed in this study, we identified genes encoding products similar to components of a secretion apparatus (*spiA* and *spiB*) and a transcriptional regulator (*spiR*). SpiR exhibits broad-scale similarity to proteins of the two-component regulatory family and is likely to control the expression of adjacent loci in the *spi* pathogenicity island. Indeed, the *spiR* mutant displayed a

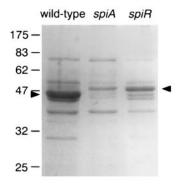


FIG. 4. Protein profile of supernatants prepared from wild-type, *spiA*, and *spiR* strains of *S. typhimurium*.

profile of exported proteins identical to that of the *spiA* strain (Fig. 4), implying that SpiR regulates the *spiCAB* operon and/or genes encoding the substrates of the Spi secretion system.

The *spi* locus is situated between 30.5' and 33.5', a region of the *S. typhimurium* chromosome that is relatively devoid of mapped genes (34). Only two other *Salmonella*-specific genes have been mapped to this region: *sly* and *ompD*. Despite the role of *sly* in virulence (35), it is not likely to be contained within the *spi* pathogenicity island due to its base composition—50.1% vs. 42.1% G+C for the *spi* locus—and its presence in *Shigella* and enteropathogenic *E. coli*. The *spi* genes do not correspond to any previously described macrophagesurvival loci (36). However, it is presently unknown whether *spiA* is allelic with a virulence gene that has been reported to be homologous to *invG* because neither the DNA sequence nor the map position of this gene was provided (37).

What is the mechanism by which horizontally acquired DNA segments are incorporated into the bacterial genome? In *E. coli*, a site adjacent to the *selC* tRNA gene is the target of insertion of two different pathogenicity islands (2, 3) and of a retronphage (38). Moreover, the second pathogenicity island present in certain uropathogenic strains of *E. coli* is located next to *leuX*, which, like *selC*, encodes a tRNA gene for rarely used codons (2). On the other hand, a 100-kb virulence region in the *Yersinia pestis* chromosome is flanked by repeated sequences (39). Thus, different mechanisms may mediate the incorporation of foreign DNA sequences. Future analysis of the entire *spi* island should reveal the manner in which the *spi* region was obtained by ancestral *Salmonella*.

The enteric pathogens *Salmonella* and *Shigella* employ similar secretion systems to promote entry into host cells (22). The differences in disease pathology and host range displayed by these pathogens most likely result from genes that are species-specific. The *spi* pathogenicity island may harbor such genes because: (*i*) it is present in all *Salmonella* serotypes investigated (5); (*ii*) *spi*-hybridizing sequences have not been detected in the genomes of other enteric species (6); and (*iii*) this region is essential for virulence functions beyond host cell invasion.

We thank Eric Mayer for technical assistance, Christine Smith (Searle/Monsanto) for protein sequencing, and anonymous reviewers for suggestions on a previous version of this manuscript. This research was supported by grants from the National Institutes of Health (GM-54900 and GM-48407). E.A.G. is a recipient of a Junior Faculty Research Award from the American Cancer Society. F.C.S. is a Career Investigator of the Consejo de Investigaciones de la Universidad Nacional de Rosario (CIUNR) (Rosario, Argentina) and is also supported by the Pew Charitable Trusts.

1. Groisman, E. A. & Ochman, H. (1994) Trends Microbiol. 2, 289–294.

- Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschäpe, H. & Hacker, J. (1994) *Infect. Immun.* 62, 606–614.
- McDaniel, T. K., Jarvis, K. G., Donnenberg, M. S. & Kaper, J. B. (1995) Proc. Natl. Acad. Sci. USA 92, 1664–1668.
- Jarvis, K. G., Girón, J. A., Jerse, A. E., McDaniel, T. K., Donnenberg, M. S. & Kaper, J. B. (1995) *Proc. Natl. Acad. Sci. USA* 92, 7996–8000.
- 5. Fitts, R. (1985) Food Technol. 39, 95-102.
- Groisman, E. A., Sturmoski, M. A., Solomon, F. R., Lin, R. & Ochman, H. (1993) Proc. Natl. Acad. Sci. USA 90, 1033–1037.
- Ochman, H. & Groisman, E. A. (1994) in *The Origin and Evolution of Species Differences in Escherichia coli and Salmonella typhimurium*, eds. Schierwater, B., Streit, B., Wagner, G. P. & DeSalle, R. (Birkhauser, New York), pp. 479–493.
- 8. Groisman, E. A. & Ochman, H. (1993) EMBO J. 12, 3779-3787.
- 9. Mills, D. M., Bajaj, V. & Lee, C. A. (1995) *Mol. Microbiol.* 15, 749–759.
- 10. Parsot, C. (1994) Curr. Top. Microbiol. Immunol. 192, 217-241.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- 12. Buchmeier, N. A. & Heffron, F. (1989) Infect. Immun. 57, 1-7.
- Lee, C. A., Jones, B. D. & Falkow, S. (1992) Proc. Natl. Acad. Sci. USA 89, 1847–1851.
- Li, J., Ochman, H., Groisman, E. A., Boyd, E. F., Solomon, F., Nelson, K. & Selander, R. K. (1995) *Proc. Natl. Acad. Sci. USA* 92, 7252–7256.
- Stock, J. B., Surette, M. G., Levit, M. & Park, P. (1995) in Two-Component Signal Transduction Systems: Structure-Function Relationships and Mechanisms of Catalysis, eds. Hoch, J. A. & Silhavy, T. J. (Am. Soc. Microbiol., Washington, DC), pp. 25–51.
- Groisman, E. A. & Heffron, F. (1995) in *Regulation of Salmonella* Virulence by Two-Component Regulatory Systems, eds. Hoch, J. A. & Silhavy, T. J. (Am. Soc. Microbiol., Washington, DC), pp. 319–332.
- Nagasawa, S., Tokishita, S., Aiba, H. & Mizuno, T. (1992) Mol. Microbiol. 6, 799–807.
- Gottesman, S. (1995) in Regulation of Capsule Synthesis: Modification of the Two-Component Paradigm by an Accessory Unstable Regulator, eds. Hoch, J. A. & Silhavy, T. J. (Am. Soc. Microbiol., Washington, DC), pp. 253–262.
- Uhl, A. A. & Miller, J. F. (1995) in *Bordetella pertussis BvgAS* Virulence Control System, eds. Hoch, J. A. & Silhavy, T. J. (Am. Soc. Microbiol., Washington, DC), pp. 333–349.

- Chang, C. & Meyerowitz, E. M. (1994) Res. Microbiol. 145, 481–486.
- 21. Salmond, G. P. C. & Reeves, P. J. (1993) *Trends Biochem. Sci.* 18, 7–12.
- 22. Ochman, H. & Groisman, E. A. (1995) Can. J. Microbiol. 41, 555–561.
- Cornelis, G. R. (1994) Curr. Top. Microbiol. Immunol. 192, 243–263.
- Van Gijsegem, F., Gough, C., Zischek, C., Niqueux, E., Arlat, M., Genin, S., Barberis, P., German, S., Castello, P. & Boucher, C. (1995) *Mol. Microbiol.* 15, 1095–1114.
- 25. Wei, Z.-M. & Beer, S. V. (1993) J. Bacteriol. 175, 7958-7967.
- Michiels, T., Vanooteghem, J.-C., Lambert de Rouvroit, C., China, B., Gustin, A., Boudry, P. & Cornelis, G. R. (1991) J. Bacteriol. 173, 4994–5009.
- Kaniga, K., Bossio, J. C. & Galán, J. E. (1994) Mol. Microbiol. 13, 555–568.
- Galán, J. & Curtiss, R. (1989) Proc. Natl. Acad. Sci. USA 86, 6383–6387.
- 29. Finlay, B. B. (1994) Curr. Top. Microbiol. Immunol. 192, 163-185.
- Fields, P. I., Swanson, R. V., Haidaris, C. G. & Heffron, F. (1986) Proc. Natl. Acad. Sci. USA 83, 5189–5193.
- García-del Portillo, F. & Finlay, B. B. (1995) J. Cell Biol. 129, 81–97.
- Alpuche-Aranda, C. M., Swanson, J. A., Loomis, W. P. & Miller, S. I. (1992) Proc. Natl. Acad. Sci. USA 89, 10079–10083.
- Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S. & Church, G. M. (1993) J. Bacteriol. 175, 3303–3316.
- Sanderson, K. E., Hessel, A. & Rudd, K. E. (1995) *Microbiol. Rev.* 59, 241–303.
- Libby, S. J., Goebel, W., Ludwig, A., Buchmeier, N., Bowe, F., Fang, F. C., Guiney, D. G., Songer, J. G. & Heffron, F. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 489–493.
- Bäumler, A. J., Kusters, J. G., Stojiljkovic, I. & Heffron, F. (1994) Infect. Immun. 62, 1623–1630.
- Hensel, M., Shea, J. E., Gleeson, C., Jones, M. D., Dalton, E. & Holden, D. W. (1995) *Science* 269, 400–403.
- Inouye, S., Sunshine, M. G., Six, E. W. & Inouye, M. (1991) Science 252, 969–971.
- Fetherston, J. D., Schuetze, P. & Perry, R. D. (1992) Mol. Microbiol. 6, 2693–2704.