The EMBO Journal vol.12 no.11 pp.4053-4062, 1993

Molecular genetic analysis of a locus required for resistance to antimicrobial peptides in *Salmonella typhimurium*

Carlos Parra-Lopez, Michael T.Baer and Eduardo A.Groisman

Department of Molecular Microbiology, Washington University School of Medicine, Box 8230, St Louis, MO 63110, USA

Communicated by S.Normark

The innate immunity of vertebrates and invertebrates to microbial infection is mediated in part by small cationic peptides with antimicrobial activity. Successful pathogens have evolved mechanisms to withstand the antibiotic activity of these molecules. We have isolated a set of genes from *Salmonella typhimurium* which are required for virulence and resistance to the antimicrobial peptides melittin and protamine. Sequence analysis of a 5.7 kb segment from the wild-type plasmid conferring resistance to protamine contained five open reading frames: *sapA*, *sapB*, *sapC*, *sapD* and *sapF*, organized in an operon structure and transcribed as a 5.3 kb mRNA. SapD and SapF exhibited similarity with the 'ATP binding cassette' family of transporters including the bacterial Opp and SpoOK, involved in the uptake of oligopeptides; the yeast STE6, necessary for the export of a peptide pheromone; and the mammalian *mdr*, which mediates resistance to chemotherapeutic agents in cancer cells. SapA showed identity with other periplasmic solute binding proteins involved in peptide transport. The SapABCDF system constitutes a novel transporter for enteric bacteria and the first one harboring a periplasmic component with a role in virulence.

Key words: antimicrobial peptide/ATP binding cassette transporter/melittin/*Salmonella*/virulence

Introduction

Animals and plants are continuously exposed to a variety of microorganisms with the capacity to cause disease. An effective mechanism of defense against invading pathogens is mediated by a family of small cationic peptide antibiotics with the capacity to adopt amphipathic α-helical structures and to form voltage-gated channels in membranes (Boman, 1991; Lehrer *et al.*, 1991; Zasloff, 1992). Although the production of such compounds had originally been associated with species lacking T cells or antibodies, the purification of cationic peptide homologs from both mammals and insects suggests a conservation in the evolution of what seems to be an effective host defense strategy (Du Pasquier, 1992). Some of the best characterized antimicrobial peptides include the magainins from the skin (Zasloff, 1987) and stomach (Moore *et al.*, 1991) of the frog, the cecropins from the *Cecropia* moth hemolymph (Steiner *et al.*, 1981) and pig intestine (Lee *et al.*, 1989) and the defensins, which have been purified from phagocytic (Ganz *et al.*, 1990; Selsted *et al.*, 1993) and epithelial (Diamond *et al.*, 1991; Eisenhauer *et al.*, 1992; Ouellette *et al.*, 1992a,b) cells of several mammals. Defensin homologs have also been found in insects (Lambert *et al.*, 1989).

Facultative intracellular pathogens are organisms which can survive within phagocytic cells because they have evolved mechanisms to circumvent or withstand the microbicidal compounds presented by the host cell (Mims, 1987). For example, the Gram-negative bacterium *Salmonella typhimurium* can replicate within macrophages and resist the battery of cationic peptides which are normally found within the lysosomal granules. The requirement of resistance to microbicidal peptides for *Salmonella* pathogenesis was first shown with mutants exhibiting hypersensitivity to defensins (Fields *et al.*, 1989). These strains were highly attenuated for virulence, exhibited hypersusceptibility to magainins, cecropins, melittin and mastoparan (Groisman *et al.*, 1992b), and harbored mutations in the *phoP* locus. PhoP encodes a transcription factor that belongs to the family of two-component regulatory systems (Groisman *et al.*, 1989; Miller *et al.*, 1989). While this observation suggested that resistance to host defense peptides is transcriptionally regulated, none of the five PhoP-regulated genes identified so far plays a role in defensin resistance (Fields *et al.*, 1989; Groisman *et al.*, 1989; Miller *et al.*, 1990).

Gram-negative bacteria have two membranes of distinct composition and function. Peptide-mediated killing is associated with permeability changes affecting the integrity of the inner membrane (Lehrer *et al.*, 1989). The lipopolysaccharide (LPS) present in the outer membrane is one of the bacterial factors with a demonstrated role in resistance to magainin 2 (Macias *et al.*, 1990; Rana *et al.*, 1991) and other antimicrobial compounds (Groisman *et al.*, 1992b). To identify additional determinants involved in resistance to antimicrobial peptides, we screened *Salmonella* mutants for hypersusceptibility to the antimicrobial peptide protamine, and recovered 12 strains that were hypersensitive to one or more peptides from a group of six tested (Groisman *et al.*, 1992b). The distinct patterns of susceptibility to these peptides and the distribution of resistance loci in the chromosome demonstrated that *Salmonella* possesses several mechanisms of peptide resistance (Groisman *et al.*, 1992b). Moreover, these mutants had reduced virulence, providing evidence for a direct role of resistance to cationic peptides in *Salmonella* pathogenesis.

Three of these *sap* (sensitive to antimicrobial peptides) mutants were of particular interest because they exhibited hypersensitivity to melittin and to crude extracts from human neutrophil granules but were still resistant to defensins. While each of these mutants harbored transposon insertions mapping near *pyrF* at 33 min in the chromosome (Groisman *et al.*, 1992b), one was more attenuated with respect to peptide susceptibility and virulence. This suggested that the transposon insertions affected at least two different genes. In this report, we describe the molecular genetic

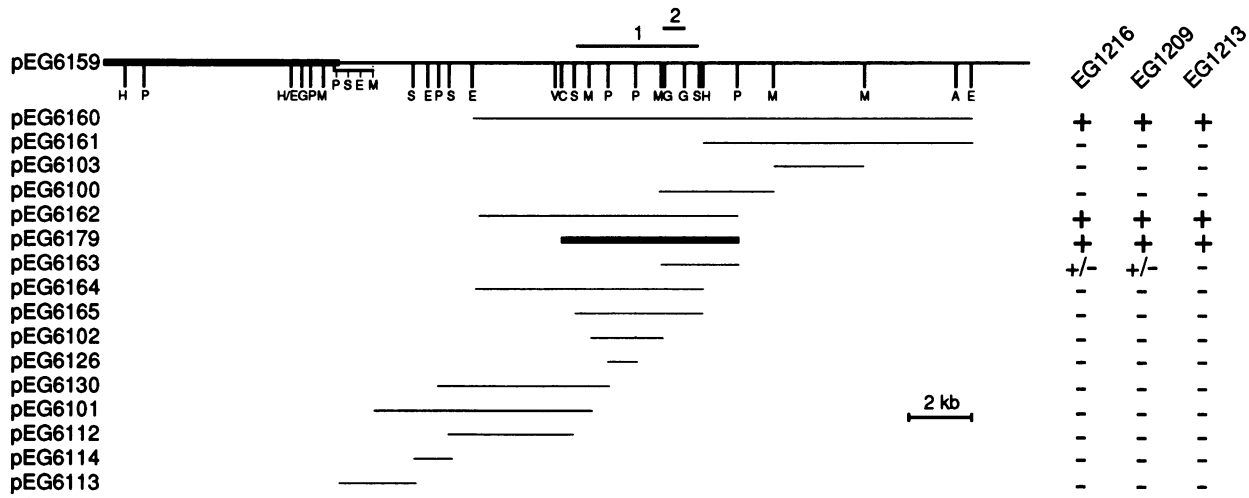

Fig. 1. Physical maps and complementation data for *sap*-containing plasmids. The restriction map of the *sapABCDF*-containing plasmid pEG6159 is shown at the top; the dark bar corresponds to Mud5005 vector sequences (Groisman and Casadaban, 1986). DNA present in subclones of pEG6159 is illustrated as horizontal lines beneath the restriction map, and the ability of these subclones to complement protamine-sensitive mutants EG1209, EG1213 and EG1216 for growth in the presence of protamine is indicated on the right hand side of the figure: +, growth; −, no growth; +/−, partial growth. The bar representing plasmid pEG6179 shows the smallest amount of DNA that can rescue all three mutants. Restriction fragments 1 and 2 denote DNA segments utilized as probes for Southern (Figure 6) and colony hybridization experiments (see text), respectively. All occurrences for restriction sites are indicated with the exception of *Msc*I and *Eco*RV. Restriction sites are abbreviated as follows: A, *Sac*I; C, *Msc*I; E, *Eco*RI; G, *Bgl*II; H, *Hin*dIII; M, *Sma*I; P, *Pst*I; S, *Sal*I and V, *Eco*RV.

characterization of this resistance locus uncovering a system that exhibits sequence similarity with prokaryotic and eukaryotic transporters that have been implicated in resistance to drugs and the transport of different solutes.

Results

Cloning of the sap locus

To characterize further the resistance determinants mapping to the 33 min region, we cloned the wild-type genes and localized the site of transposon insertion for each of the protamine-sensitive mutants. First, we isolated the MudJ−*Salmonella* junction fragments by preparing genomic libraries using the restriction endonuclease *Sal*I, whose single recognition site in MudJ is present downstream of the *kan* gene. By selecting for kanamycin- and ampicillin (vector-encoded)-resistant transformants, we obtained plasmid clones that harbored fragments consisting of the *Salmonella* DNA adjacent to the left end of MudJ. We then constructed restriction maps for the three plasmid clones, purified a 900 bp *Bgl*II *Salmonella*-specific fragment present in two of them, and used it to screen a plasmid library for the wild-type gene. One candidate clone—pEG6159 (Figure 1)—had a 21.5 kb insert with several restriction fragments in common with the plasmids harboring the MudJ−*Salmonella* joints. This clone rescued the protamine-sensitive phenotype of CP1216, a kanamycin-sensitive, ampicillin-resistant derivative of mutant EG1216. This strain was used because the wild-type library was generated with the vector Mud5005 (Groisman and Casadaban, 1986), which encodes resistance to kanamycin like the MudJ element in strain EG1216. To localize the wild-type gene within pEG6159, and to determine the minimal DNA fragment which could rescue the protamine susceptibility phenotype, we generated and tested several subclones for their ability to complement the protamine-sensitive strains. The only plasmids that fully complemented all three mutants were pEG6160, pEG6162 and pEG6179, delimiting the region of interest to the 6.3 kb segment between the rightmost *Pst*I site and the shown *Msc*I site (Figure 1). Southern hybridization experiments using chromosomal DNA from wild-type and mutant strains probed with *Salmonella* or MudJ-specific DNA established that the three MudJ insertions were present within a 1.2 kb region (data not shown): this is consistent with their similar genetic linkage to *pyrF* (Groisman *et al.*, 1992b).

Sequence analysis of the sap locus

The nucleotide sequence of a 5.7 kb segment from plasmid pEG6162 was analyzed to identify the determinants responsible for peptide resistance. There were five open reading frames in one strand, designated *sapA*, *sapB*, *sapC*, *sapD* and *sapF*, which encoded products with predicted mol. wts of 61 622, 36 082, 31 547, 37 611 and 30 611, respectively (Figure 2). Assignment of genes to these open reading frames is based on the following: first, except for these five genes, all remaining open reading frames within this segment are shorter than 101 amino acids. Second, the codon usage within these reading frames is typical of other highly expressed *S.typhimurium* genes. Third, there is extensive sequence similarity of the *sap* gene products with other proteins in the database. And fourth, there are stop codons in all three reading frames downstream of *sapF*. All open reading frames except for *sapA* harbored potential ribosome binding sites at the appropriate distance from their putative initiation codons. The assignment of the start codon for *sapA* was based on the agreement between the predicted and experimentally determined sizes of the protein, the presence of a potential signal sequence and the similarity of SapA with known proteins in the database.

That *sapABCDF* might be transcribed as a single multicistronic operon was suggested by the absence of intergenic regions and the presence of overlapping stop and start codons. Downstream of *sapF* there is a region that could form a stem−loop structure and favor *rho* independent termination of transcription (Figure 2). We detected a single transcript of ~5.3 kb by Northern analysis with RNA

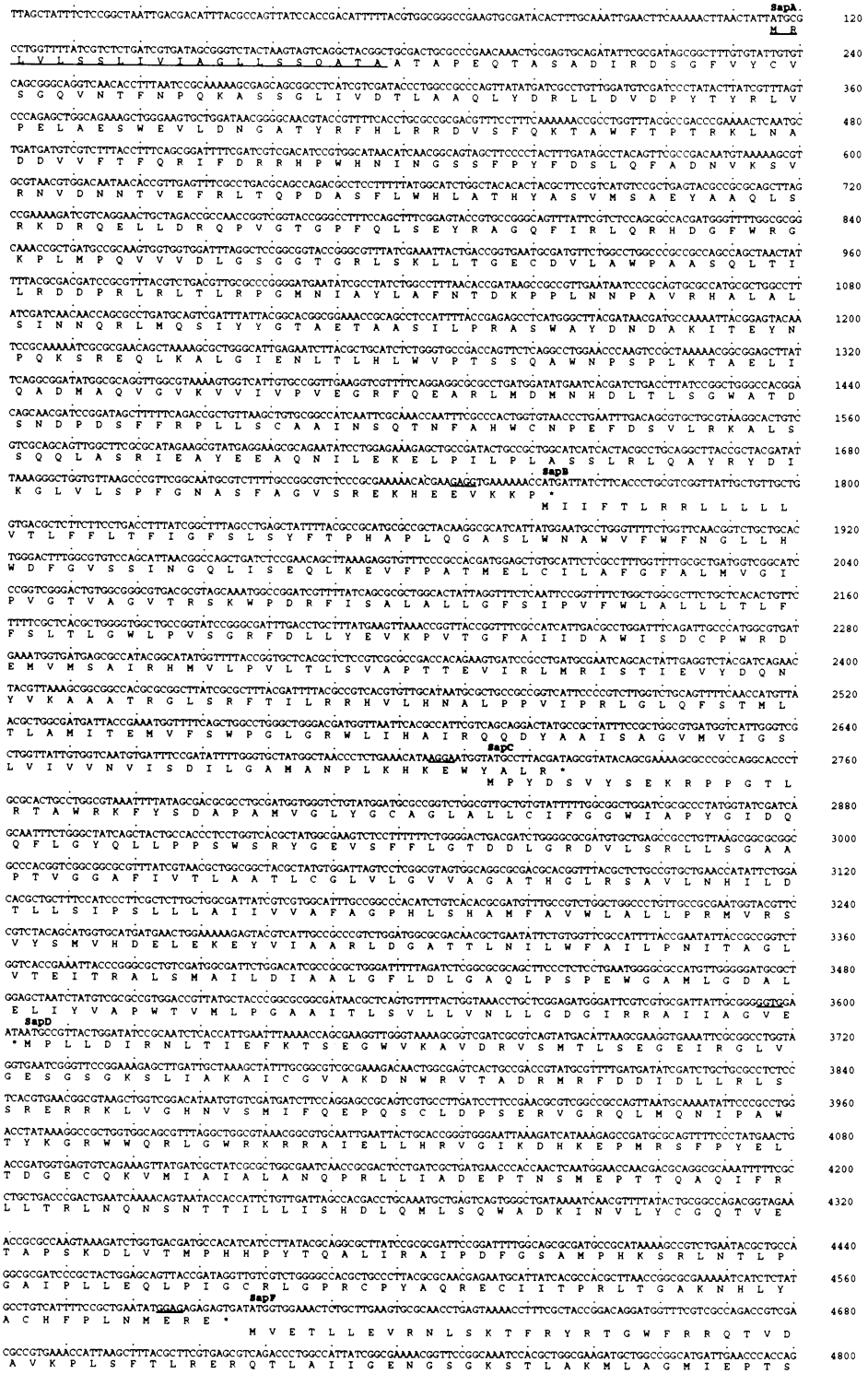

Fig. 2. Complete nucleotide sequence of the *sapABCDF* operon and deduced amino acid sequences of the encoded products. Potential ribosome binding sites and potential signal sequence in SapA are underlined. An inverted repeat downstream of *sapF* is in bold and also underlined. The nucleotide sequence reported in this paper has been deposited in the EMBL Data Library under accession number X74212.

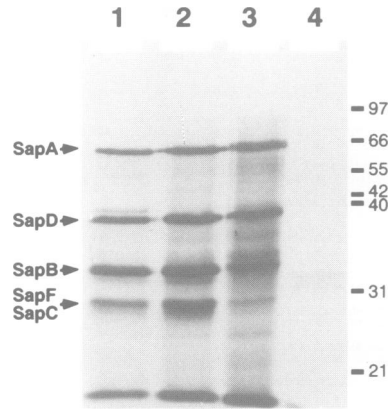

Fig. 3. Expression of *sapABCDF*-encoded products. The proteins encoded by plasmids pEG6162, pEG6179 and pEG6188 were expressed as described in Materials and methods and a fluorogram corresponding to a 12 h exposure at −70°C is shown. Numbers indicate mol. wt in kDa. Lysates were prepared from HB101 cells harboring: 1, pEG6162, pGP1-4; 2, pEG6179, pGP1-4; 3, pEG6188, pGP1-4 and 4, pT7-5, pGP1-4 as control.

prepared from mid-log phase wild-type cells and probes corresponding to the coding region of either *sapA*, *sapB* or *sapF* (data not shown). The right end of the MudJ transposon contains a promoterless *lac* operon and upon insertion in the proper orientation, transcriptional fusions are created (Castilho *et al.*, 1984). Mutants EG1209 and EG1213, which expressed similarly low levels of β-galactosidase activity in M63 glucose X-Gal plates, harbored the MudJ transposon with its right end closer to the putative promoter region of *sapABCDF*. Mutant EG1216, with the MudJ inserted in the opposite orientation, did not produce β-galactosidase.

Expression of Sap proteins

To identify the products encoded by the *sapABCDF* operon, we used the phage T7 promoter system to express the genes encoded by plasmids pEG6162 and pEG6179, which harbor the complete operon in their 9.3 kb and 6.3 kb inserts (Figure 1), respectively; and pEG6188, a derivative of pEG6179 with only 36 amino acids of the *sapF* coding region. Induction of the T7 polymerase was followed by addition of rifampicin to inhibit host transcription and labeled polypeptides were visualized using ^{3}H amino acids (Figure 3). We identified polypeptides of about 61, 38 and 36 kDa in size, in agreement with the predicted molecular weights for SapA, SapD and SapB, respectively; and a doublet of ~30 kDa corresponding to SapC and SapF. Assignment of the visualized bands to the predicted gene products of the *sapABCDF* operon is based on a series of T7 expression and *in vitro* transcription/translation experiments using plasmid pEG6162 derivatives pEG6163, pEG6164 and pEG6165 (Figure 1), which carry different subsets of the *sap* genes (data not shown).

Sap belongs to the 'ATP binding cassette' family of proteins

We searched the sequence databases (GenBank release 73) with the predicted amino acid sequences of each of the five open reading frames. We found that SapD and SapF have a high degree of sequence similarity with several members of the 'ATP binding cassette' family of transporters (Figure 4). This group of proteins includes the eukaryotic *mdr* which has been implicated in resistance to chemotherapeutic agents in mice and humans (Chen *et al.*, 1986), the CFTR, involved in the transport of Cl^- anions (Riordan *et al.*, 1989), and the yeast STE6, required for the release of a peptide pheromone (McGrath and Varshavsky, 1989). Bacterial homologs include the Mal and His systems involved in the uptake of maltose and histidine, respectively, as well as the Hly and Clv necessary for the export of hemolysin and colicin in *Escherichia coli* (see Higgins, 1992, for review). The region of similarity is located within a stretch of ~200 amino acids and extends beyond the two ATP binding sites known as Walker motifs (Walker *et al.*, 1982) which are present in several ATP binding proteins (Figure 4). The highest degree of identity (Table I) was found with homologs in the Opp (Hiles *et al.*, 1987) and SpoOK (Perego *et al.*, 1991; Rudner *et al.*, 1991) systems, which participate in the uptake of oligopeptides in *S.typhimurium* and *Bacillus subtilis*, respectively. SapD and SapF are 40% identical (61% similar) to the *Salmonella* OppD and OppF; and all four of these proteins show frequent regions of hydrophilicity with no long hydrophobic stretches along the sequence (data not shown). Furthermore, like the OppD/OppF pair, SapD and SapF are similar to each other.

SapA contains a signal sequence and its predicted periplasmic location suggests that it may be responsible for binding the ligand to be transported into the cytoplasm of the cell. This protein exhibited the highest degree of sequence identity (36%) with DppA from *E.coli* (Abouhamad *et al.*, 1991), which is involved in the transport of dipeptides and the *Salmonella* OppA (26% identity), required for oligopeptide uptake. Analysis of the SapB and SapC sequences predicted several stretches of hydrophobic amino acids in each of these proteins which could correspond to transmembrane domains (data not shown). These proteins also showed sequence identity with each other and with the membrane components of the Opp and SpoOK transport systems: SapB and SapC were 33% and 26% identical to OppB and OppC, respectively (Table II). The similarity to these proteins extends to their predicted secondary structures regarding the distribution of hydrophilic and hydrophobic regions. Beside the similarities exhibited between the individual proteins of the Sap, Opp and SpoOK systems, there are interesting differences between these transporters. While both systems harbor two distinct ATP binding proteins, SapD and SapF are very different in size (331 and 269 residues, respectively) and show only 28% identity, relative to the homologous pairs OppD/OppF and SpoOKD/SpoOKE which are of similar size and display higher levels of sequence identity (41% and 42%, respectively). The order of the genes is conserved in the three operons but their arrangement is not identical: the intergenic region present between the first and second genes in the *opp* (120 bp; Hiles *et al.*, 1987) and *spoOK* (106 bp; Perego *et al.*, 1991; Rudner *et al.*, 1991) operons is absent from *sap*.

Functional requirement of the different components of the Sap transporter

Sequence analysis revealed that mutant EG1213 harbored a MudJ in *sapC* (nucleotide 3605), and that both EG1209 and EG1216 harbored insertions in *sapD* (nucleotides 4308 and 4785, respectively; see Figure 5). Given the similarities between the Opp and Sap systems, one might predict that all five Sap proteins are required for activity of this system.

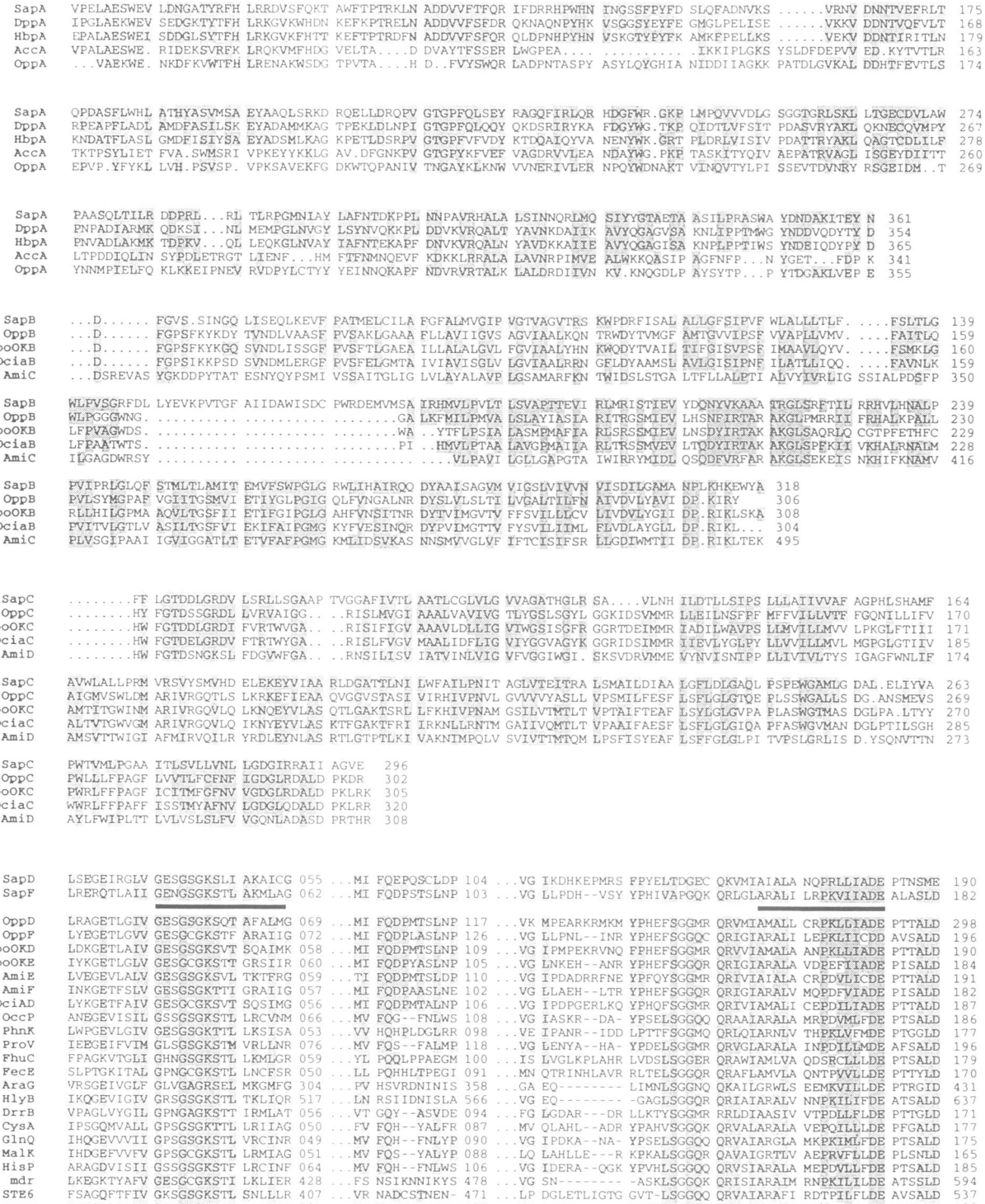

Fig. 4. Comparison of Sap proteins with other transporters. Alignment of the amino acid sequences of Sap proteins with similar proteins in the database (GenBank release 73) was performed using a combination of the programs TFASTA (average match, 0.54; average mismatch, −0.396) and PILEUP (gap weight: 3.0 and gap length weight: 0.1) from GCG. Identical or similar amino acids within each comparison are highlighted and grouped as follows: AST, DEQ, ILMV, RK, FYW. Dark bars indicate the positions of the consensus sequence for the ATP binding site motif (Walker *et al.*, 1982). Opp and SpoOK, oligopeptide transporters of *S.typhimurium* and *B.subtilis;* DppA, dipeptide binding protein of *E.coli*; HbpA, heme binding lipoprotein of *Haemophilus influenzae*; AccA, potential opine binding protein of *A.tumefaciens*; Ami, aminopterin transporter of *Streptococcus pneumoniae*; Dcia, dipeptide transporter of *B.subtilis;* OccP, octopine permease of *Agrobacterium tumefaciens*; PhnK, carbon-phosphorus lyase of *E.coli*; ProV, betaine transporter of *E.coli*; FhuC, iron (III) hydroxamate transporter of *E.coli*; FecE, iron (III) dicitrate transporter of *E.coli*; AraG, arabinose transporter of *E.coli*; HlyB, haemolysin exporter of *E.coli*; DrrB, daunorubicin exporter of *Streptomyces peucetius*; CysA, sulfate permease of *E.coli*; GlnQ, glutamine transporter of *Bacillus stearothermophilus;* MalK, maltose transporter of *S.typhimurium;* His, histidine permease of *S.typhimurium;* mdr, multiple drug resistance P-glycoprotein of *Plasmodium falciparum*; STE6, peptide pheromone transporter of *Saccharomyces cerevisiae*; Brown and White, pigment transporters of *Drosophila melanogaster*; Ham1, peptide transporter of mouse endoplasmic reticulum; CFTR, cystic fibrosis transmembrane conductance regulator of humans.

Table I. Sequence identitiy between cytoplasmic components of 'ABC' transporters[a]

	SapF[b]	OppD	OppF	SpoOKD	SpoOKE	AmiE	AmiF	TAP1	HisP	DciAD
SapD	27(56)[c]	40(61)	32(59)	36(60)	35(55)	33(59)	31(54)	24(55)	24(53)	34(57)
SapF		35(55)	40(61)	35(60)	38(61)	34(57)	34(63)	25(51)	27(56)	33(58)
OppD			41(62)	51(69)	39(58)	44(63)	36(55)	26(48)	31(55)	50(70)
OppF				41(62)	53(69)	37(57)	47(68)	26(55)	31(59)	43(63)
SpoOKD					42(60)	54(71)	39(60)	28(54)	31(54)	58(72)
SpoOKE						35(58)	54(74)	25(51)	30(62)	43(68)
AmiE							37(57)	27(47)	31(53)	45(64)
AmiF								25(53)	34(56)	40(61)
TAP1									28(53)	24(50)
HisP										32(56)

[a]Alignments were produced using the GCG program BESTFIT (gap weight of 3.0 and gap length weight of 0.1).
[b]The proteins correspond to those described in the legend to Figure 4.
[c]Values correspond to percentage identity. Numbers in parentheses correspond to percentage similarity.

Table II. Sequence identity between transmembrane components of different transport systems[a]

	SapC[b]	OppB	OppC	SpoOKB	SpoOKC	AmiC	AmiD	HisM	HisQ	DciAB	DciAC
SapB	22(50)[c]	33(58)	18(50)	22(50)	23(54)	19(49)	23(52)	21(51)	18(49)	32(59)	21(49)
SapC		22(56)	26(54)	23(53)	30(58)	22(48)	22(48)	23(46)	17(51)	26(52)	32(57)
OppB			27(51)	47(73)	24(51)	23(58)	20(44)	23(49)	20(48)	30(42)	26(53)
OppC				19(51)	43(67)	21(50)	30(61)	18(47)	22(50)	22(49)	41(67)
SpoOKB					19(51)	25(51)	18(46)	18(48)	17(44)	51(71)	20(52)
SpoOKC						22(48)	37(62)	21(47)	21(49)	19(49)	57(77)
AmiC							21(46)	20(44)	22(47)	30(60)	18(48)
AmiF								19(46)	20(46)	19(48)	32(60)
HisM									30(57)	19(46)	22(50)
HisQ										21(46)	20(52)
DciAB											23(52)

[a]Alignments were produced using the GCG program BESTFIT (gap weight of 3.0 and gap length weight of 0.1).
[b]The proteins correspond to those described in the legend to Figure 4.
[c]Values correspond to percentage identity. Numbers in parentheses correspond to percentage similarity.

However, the *sapD* mutants EG1209 and EG1216 had a low level of residual activity: their median lethal dose was two orders of magnitude lower and they were less susceptible to protamine than the *sapC* mutant EG1213 (Groisman *et al.*, 1992b). While it has been suggested that the interactions between the membrane and the ATP binding components are transporter specific (Higgins, 1992), our results raise the possibility of molecular 'cross-talk' between homologous systems whereas the membrane components of the Sap system might interact with ATP binding cassette proteins homologous to SapD and SapF. The more attenuated phenotype of EG1213 could be due to disruption of *sapC* coupled to polar effects on *sapD* and *sapF*. To test whether mutations of other *sap* genes led to further attenuation, we constructed a strain with a deletion that removed the complete *sapABCDF* coding region except for the nucleotide sequences corresponding to the 34 N-terminal amino acids of SapA and the 44 C-terminal amino acids of SapF (Figure 5). The resulting Δ*sap* strain EG6501 behaved like the *sapC* mutant EG1213 when tested for protamine resistance, and its defect could be complemented by plasmid pEG6179 (Figure 1).

We investigated whether SapF was required for resistance to antimicrobial peptides given (i) the 56% similarity between SapD and SapF, (ii) that certain homologous transporters, such as the His and Mal systems (Higgins, 1992) function with two copies of the same ATP binding protein, and (iii) that the homologous SpoOK system of *Bacillus* harbors two distinct, albeit homologous, ATP binding proteins, but the SapF homolog is not required for peptide transport (Rudner *et al.*, 1991). Moreover, the MudJ in EG1216 is close to the 3′ end of the *sapD* gene—affecting the last 15 residues of the 331 amino acid SapD—and it is possible that the protamine hypersusceptibility phenotype is due to polarity resulting in a *sapF* defect. Therefore, we constructed a strain (EG6545) harboring a 1.6 kb *kan* insertion at the *Hin*dIII site within the *sapF* coding region (Figure 5) that left only the 36 N-terminal amino acids of SapF. This strain was as sensitive to protamine as the *sapD* mutants EG1216 and EG1209 demonstrating that *sapF* is required for activity. Furthermore, complementation of these mutants was achieved by a plasmid containing *sapDF* (pEG6163) but not by one (pEG6164) lacking *sapF* but carrying the *sapABCD* genes (Figure 1).

In bacteria, transporters that harbor a periplasmic solute binding protein usually import solutes while those lacking such a component are involved in export (Higgins, 1992). However, this requirement might not be absolute because mutants can be isolated in the maltose (Shuman *et al.*, 1982) and histidine (Speiser and Ames, 1991) transporters that function in the absence of the periplasmic solute binding proteins. To establish whether SapA—the putative periplasmic component—was required for peptide resistance, we constructed a strain harboring a non-polar in-frame deletion of *sapA* which removed 261 residues of the 549

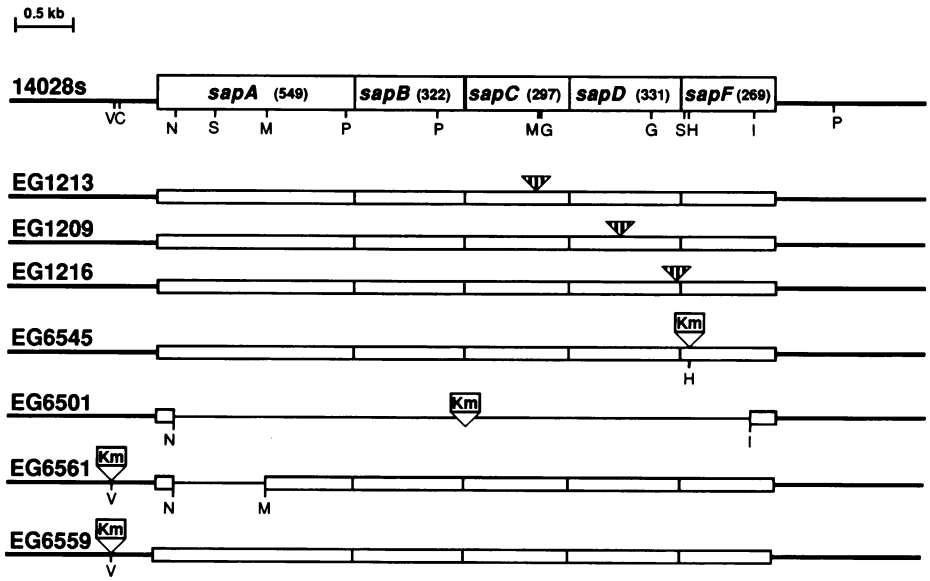

Fig. 5. Physical and genetic maps of the *S.typhimurium sapABCDF* chromosomal region for wild-type and *sap* mutant strains. Genetic organization of the *sap* region in the following strains: 14028s, wild-type; EG1213, *sapC*::MudJ; EG1209, *sapD*::MudJ; EG1216, *sapD*::MudJ; EG6545, *sapF*::*kan*; EG6501, Δ*sapABCDF*::*kan*; EG6561, *kan* Δ*sapA*; EG6559, *kan* is shown. Numbers in parentheses correspond to the length of each Sap protein. Striped triangles indicate the position of MudJ insertions in mutants EG1209, EG1213 and EG1216; and Km indicates the position of kanamycin resistance cassettes in the remaining mutants. The thin lines in strains EG6501 and EG6561 correspond to the portion of the *S.typhimurium* chromosome deleted in these strains. Restriction sites were abbreviated as described in Figure 1; I, *Nsi*I; N, *Nru*I. All occurrences for restriction sites are indicated with the exception of *Msc*I, *Nru*I, *Nsi*I and *Eco*RV.

amino acid protein; this strain also harbored a *kan* insertion 477 bp upstream of *sapA* required for the transfer of the mutation from the plasmid back to the chromosome (Figure 5). The Δ*sapA* mutant was more susceptible to protamine than the isogenic *sapA*$^+$ strain but could still grow on LB agar protamine plates at concentrations that completely obliterated the growth of strains with mutations in *sapC* or *sapD*.

***sap* homologs are present in other Enterobacteriaceae**

Because bacterial species vary in their susceptibility to antimicrobial peptides (Zasloff, 1987; Groisman *et al.*, 1992a), we investigated whether sequences homologous to *sap* were present in other species. Southern hybridization analysis using the 4.1 kb *Sal*I fragment internal to the *sapABCDF* operon as a probe (Figure 1) showed that homologous sequences were present in the genomes of several bacterial species. Under high stringency conditions, *sap* hybridizing sequences could be detected in eight out of 15 species tested (Figure 6).

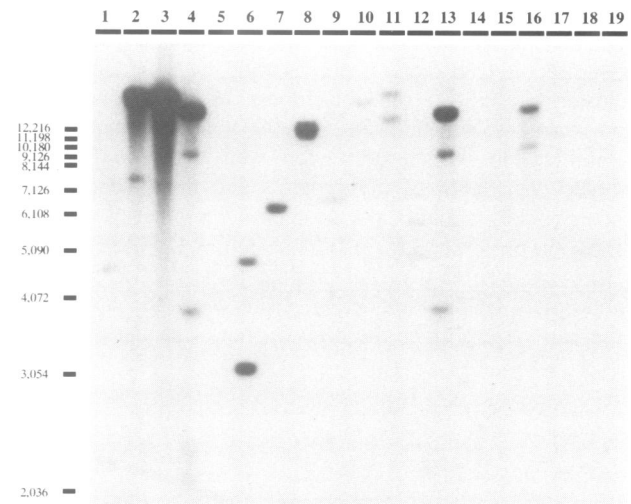

Fig. 6. Phylogenetic distribution of *sap* sequences. Southern hybridization analysis was performed using total chromosomal DNA digested with *Eco*RI and probed with the labeled 4.1 kb *Sal*I fragment internal to the *sap* operon (Figure 1) as described in Materials and methods. 1, Mol. wt markers; 2, *S.typhimurium;* 3, *S.typhimurium*; 4, *Shigella flexneri*; 5, *Enterobacter cloacae*; 6, *Citrobacter freundii;* 7, *Erwinia herbicola*; 8, *Enterobacter aerogenes*; 9, *Erwinia caratovora*; 10, *E.cloacae*; 11, *Klebsiella pneumoniae*; 12, *Proteus vulgaris*; 13, *S.flexneri*; 14, *B.subtilis*; 15, *Serratia marcescens*; 16, *E.coli*; 17, *Providencia stuartii*; 18, *Mycobacterium tuberculosis*; 19, *Streptococcus pyogenes*. Size of standards, shown in bp, are given on the left.

Discussion

Microorganisms are exposed to a variety of stress conditions both within and outside animal hosts. Pathogens that require growth within the host to cause disease have evolved distinct strategies to survive the inhibitory environments faced during the different stages of infection. For example, *Salmonella* is in contact with the acidic pH contents of the stomach, bile salts and osmolarity of the small intestine, and the various antimicrobial compounds present in the phagolysosomal compartments of phagocytic cells. We have described a system from *S.typhimurium* that is required for resistance to antibacterial peptides *in vitro* and for virulence *in vivo*. This complex is composed of five proteins exhibiting sequence identity with transporters which use energy derived from ATP to import or export a variety of solutes (Higgins, 1992). Apart from the oligopeptide permeases of bacteria (Hiles *et al.*, 1987; Alloing *et al.*, 1990; Perego *et al.*, 1991; Rudner *et al.*, 1991), this protein family includes eukaryotic

peptide transporters such as the yeast STE6 (McGrath and Varshavsky, 1989)—required for export of a peptide pheromone—and the TAP1−TAP2 complex (Monaco *et al.*, 1990; Deverson *et al.*, 1990; Trowsdale *et al.*, 1990; Spies *et al.*, 1990; Parham, 1992), which delivers cytosolic peptides across the endoplasmic reticulum for antigen presentation by the major histocompatibility complex class I molecules. Based on the presence of a predicted periplasmic component and the identity with peptide transporters, the Sap system is likely to mediate resistance to antimicrobial peptides by transporting them into the cytoplasm (and away from their putative membrane targets) where they could either be degraded by peptidases or initiate a regulatory cascade resulting in the activation of the resistance deteminants (Figure 7).

In prokaryotes, the presence of periplasmic solute binding proteins usually indicates that a transporter is involved in import rather than export of solutes. Removal of the *sapA* gene, whose product exhibits identity with other solute binding proteins, resulted in peptide susceptibility, suggesting that the SapABCDF system may be involved in uptake. This constitutes a novel drug resistance mechanism in bacteria, which often capture and pump toxic compounds in ways that are reminiscent of the mammalian Multiple drug resistance (Mdr) transporter. For example, the DrrAB system of *Streptomyces*, which exhibits sequence identity with Mdr, confers resistance to daunorubicin and doxorubicin in strains that do not manufacture these antibiotics (Patrick and Hutchinson, 1991). Periplasmic solute binding proteins of Gram-negative bacteria and homologous lipoproteins of Gram-positive bacteria were recently classified into eight clusters based on the relatedness of their amino acid sequences (Tam and Saier, 1993). According to this classification, which correlated well with the predicted molecular weight of the proteins and the type of solute bound, SapA would be placed in cluster 5 with other proteins showing preference for peptides such as OppA, SpoOKA, DppA as well as a nickel binding protein of *E.coli* (Tam and Saier, 1993). Peptides are likely to be the substrate of the Sap system given the high degree of sequence identity between other Sap proteins with components of the Opp (Hiles *et al.*, 1987) and SpoOK (Perego *et al.*, 1991; Rudner *et al.*, 1991) permeases of *Salmonella* and *Bacillus*, respectively.

Three distinct peptide transporters have already been described in *Salmonella*: Dpp (Abouhamad *et al.*, 1991) and Tpp (Gibson *et al.*, 1984), which are responsible for the uptake of dipeptides and tripeptides, respectively; and Opp (Hiles *et al.*, 1987), which is involved in recycling cell wall peptides and uptake of peptides of up to five amino acids (Payne and Gilvarg, 1968; Goodell and Higgins, 1987). Because a mutant strain defective in these three transporters shows no residual peptide uptake, it has been suggested that these permeases provide the only broad specificity pathway for peptide uptake in enteric species (cited in Hiles and Higgins, 1986). However, the SapABCDF system may also be involved in the uptake of peptides albeit of different length, composition, with different kinetic properties or under different environmental conditions. The various transporters are likely to have distinct roles in bacterial physiology. In contrast to *sap* mutants, strains with mutations in *opp*, *dpp* or *tpp* or a strain defective in all three permeases are fully resistant to protamine (E.A.Groisman and C.Parra-Lopez, unpublished results); and *opp* and *tpp* mutants are virulent for mice (Dorman *et al.*, 1989; Benjamin *et al.*, 1991). Furthermore, anaerobiosis controls expression of the *opp* system in *E.coli* (Andrews and Short, 1986) but does not modulate transcription of the *sapABCDF* (C.Parra-Lopez and E.A.Groisman, unpublished results) or *opp* operons in *Salmonella* (Jamieson and Higgins, 1984).

There are other possible mechanisms by which SapABCDF could mediate resistance to host defense peptides. By detecting the presence of toxic compounds, Sap could initiate a regulatory cascade resulting in the activation of the relevant peptide resistance determinants (i.e. peptidase). Such a system might not require peptide transport and activation might be achieved just by binding of the SapA−ligand complex to the membrane components (Figure 7). For example, ATP hydrolysis in the maltose system can be achieved even in the absence of ligand

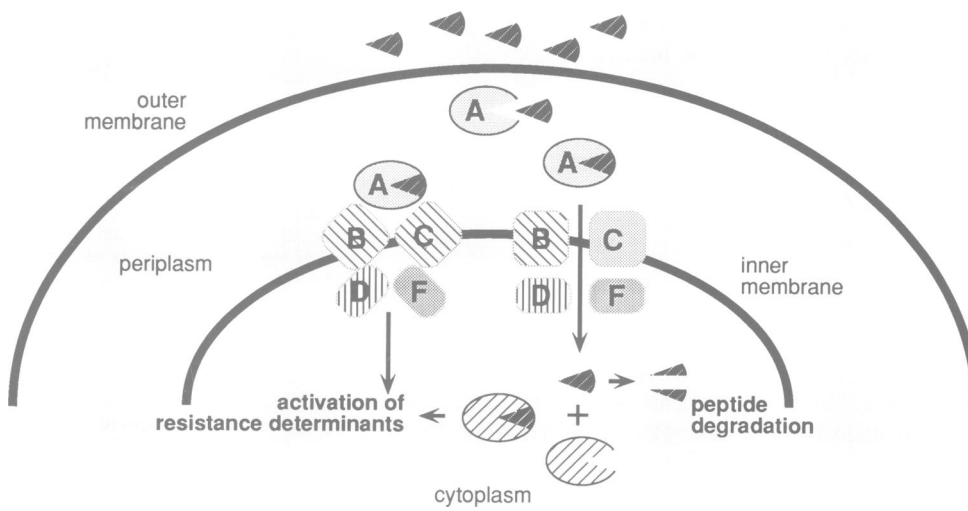

Fig. 7. Potential mechanisms of action of the Sap system: right, peptides (triangles) bind SapA in the periplasm and are transported to the cytoplasm via the remaining components of the Sap system. Once in the cytoplasm peptides can either be degraded by proteases to amino acids that could be used as building blocks of new polypeptides, or bind a regulatory molecule to activate expression of resistance determinants. Left, SapA−peptide interaction induces a conformational change in the subunits of the Sap system which is transmitted to other proteins resulting in the modulation of expression of resistance genes.

transport whereas the maltose binding protein transmits a signal across the membrane (Davidson *et al.*, 1992). Moreover, the SpoOK system of *B.subtilis* senses sporulation signals to initiate a regulatory cascade which involves activation of proteins that belong to the two-component regulatory systems (Rudner *et al.*, 1991). *Salmonella* mutants defective in one such system, PhoP/PhoQ, are avirulent and exhibit peptide hypersensitivity (Fields *et al.*, 1989; Groisman *et al.*, 1992b). While this raises the possibility of a regulatory interaction between the Sap and the PhoP systems, our preliminary data provide no evidence for such an interaction (C.Parra-Lopez and E.A.Groisman, unpublished results).

Resistance to antimicrobial cationic peptides has traditionally been associated with alterations in outer membrane integrity (Vaara, 1992): *pmrA* mutants of *Salmonella* exhibit increased resistance to protamine due to higher content of ethanolamine and amino arabinose (Vaara, 1981) and *E.coli* strains defficient in the OmpC protein show hypersensitivity to cecropin D (Sidén and Boman, 1983). The *sap* mutants did not exhibit outer membrane permeability defects because RNase release and susceptibility to lysozyme and a variety of antibiotics was identical to that of the wild-type strain (C.Parra-Lopez and E.A.Groisman, unpublished results). Sap-mediated resistance may involve the removal of the cationic peptide away from its target site in the bacterial membrane. This system constitutes a novel virulence determinant of *S.typhimurium* (Groisman *et al.*, 1992b) required for survival within macrophages *in vitro* (C.Parra-Lopez and E.A.Groisman, unpublished results) and resistance to microbicidal compounds of human neutrophil granules (Groisman *et al.*, 1992b). Granule components tested to date [defensin NP-1 (Groisman *et al.*, 1992b), lactoferrin and lysozyme (C.Parra-Lopez and E.A.Groisman, unpublished results)] did not exhibit preferential antimicrobial activity upon *sap* mutants. These strains should provide an assay for the identification of novel antimicrobial compounds of phagocytic cells. Candidate compounds include bactericidal peptide fragments derived from *in vitro* proteolysis of cathepsin G (Bangalore *et al.*, 1990), CAP37 (Pereira *et al.*, 1993) and lactoferrin (Yamauchi *et al.*, 1993). Finally, the broad phylogenetic distribution of the *sapABCDF* genes in enteric bacteria (Figure 6), suggests that this system plays a central physiological role, such as in nutrient uptake or recycling of murein peptides (Park, 1993), and that *Salmonella* evolved to utilize it for resistance to host defense compounds.

Materials and methods

Bacterial strains, plasmids and growth media

Mutant strains were isogenic derivatives of the wild-type peptide-resistant *S.typhimurium* strain 14028s. These mutants are designated as follows: EG1209, *sapD*::MudJ; EG1213, *sapC*::MudJ; EG1216, *sapD*::MudJ; EG6545, *sapF*::*kan*; EG6501, Δ*sapABCDF*::*kan*; EG6561, *kan* Δ*sapA* and EG6559, *kan*, a *sapA*$^+$ derivative of EG6561. The chromosomal mutations in strains EG6545, EG6501, EG6559 and EG6561 were made as described (Groisman *et al.*, 1993) using plasmids harboring *kan* insertions in the *sapABCDF* operon. Their structures were verified by Southern hybridization (data not shown). The position of the *kan* insertions and the segments deleted in the different mutants are illustrated in Figure 5. The *kan* cassettes were from plasmid pUC4-K (Pharmacia) in strain EG6501 and from pUC4-KIXX (Pharmacia) in the three remaining strains. *E.coli* strain POI1734 (Castilho *et al.*, 1984) was used as a source of transposon MudJ for the mutagenesis of plasmids containing the *sapABCDF* operon. Plasmid pIBI25 (IBI) was used for subcloning, pT7-5 and pGP1-4 (Tabor and Richardson, 1985) for expression studies, and pEG5005 (Groisman and Casadaban, 1986) for the construction of genomic libraries by the *in vivo* cloning procedure (Groisman, 1991). Plasmid pEG6188 was constructed by linearizing pEG6179 (Figure 1) at the unique *Hin*dIII site, filling in with Klenow fragment of DNA polymerase and ligating; as expected this plasmid lost its *Hin*dIII site. LB, M9 and M63 media were prepared as described (Miller, 1972). Kanamycin and ampicillin were used at 40 μg/ml and 50 μg/ml, respectively. Protamine (Calbiochem)-containing LB agar plates were prepared freshly to final concentrations of 0.5−1.5 mg/ml.

DNA biochemistry and molecular biological techniques

Restriction endonucleases and T4 DNA ligase were purchased from Bethesda Research Laboratories, Inc., Boehringer Mannheim Biochemicals or New England BioLabs, Inc., and were used according to the manufacturers' specifications. The wild-type operon was cloned from a genomic library prepared by the *in vivo* cloning technique using the Mud5005 mini-Mu replicon as described (Groisman, 1991). DNA was purified from host cells using reagents and midi-prep columns from QIAGEN Inc. The nucleotide sequence was determined by dideoxynucleotide chain-termination method using Sequenase Version 2.0 (USB), [α-^{35}S]dATP (Amersham), templates corresponding to different plasmid subclones, some of which harbored MudJ insertions and primers complementary to the ends of phage Mu. Additional primers were synthesized using the phosphoramidite method in a Cyclone Plus DNA synthesizer (Millipore) based on the sequence obtained. The DNA sequence presented in this paper was determined completely on both strands. Isolation of MudJ insertions in plasmid clones was performed as described (Groisman, 1991). Computer analyses were performed using the software packages GCG (University of Wisconsin Biotechnology Center, Madison, WI) and GeneWorks (Intelligenetics). We searched the sequence databases for protein similarities using the programs TFASTA and BLAST. PILEUP and PPLOT were used for optimal alignment of protein sequences and the determination of average hydrophobicity/hydrophilicity graphics, respectively.

Expression of sap genes

Plasmids harboring different segments of the *sapABCDF* chromosomal region cloned into the T7 promoter-containing vector pT7-5 were cotransformed with plasmid pGP1-4 into strain HB101 by electroporation. Expression of *sap*-encoded products was achieved following a modification of a previously described protocol (Tabor and Richardson, 1985). HB101 cells containing pGP1-4 and the pT7-5 derivatives harboring *sap* were grown in LB broth with 100 μg/ml of ampicillin and 50 μg/ml of kanamycin at 30°C. When cells reached an OD_{590} of 0.5, they were centrifuged, washed with 5 ml of M9 medium without glucose and resuspended in 1 ml of M9 medium + 0.2% glucose. Cells were then grown at 30°C for 60−180 min and shifted to 42°C for 20 min. Rifampicin (from a 20 mg/ml stock in methanol) was added to a final concentration of 200 μg/ml, and cells were incubated at 42°C for 10 min and then transferred to 30°C for 20 min. Samples were labeled with 10 μCi of ^{3}H amino acid mixture (Amersham) for 5 min at 30°C, centrifuged in a bench top centrifuge at 13 000 r.p.m. for 20 s and the pellet subjected to 10% SDS−PAGE after being resuspended in loading buffer. The gel was fixed and stained in methanol (30%)/acetic acid (10%)/Coomassie brilliant blue (0.2%) at 65°C for 5 min and destained in methanol (30%)/acetic acid (10%) at 65°C for 15 min. It was then incubated in EN^{3}HANCE (DuPont) for 1 h, in water for 30 min and dried at 85°C. Reagents for protein electrophoresis analysis were from Bio-Rad.

Acknowledgments

We thank A.Miczak for plasmids, and H.Ochman, J.Woods, R.Kranz and R.Munson for comments on the manuscript. This work was supported by National Science Foundation Grant DCB-8916403 and by a grant from the Monsanto/Searle−Washington University School of Medicine Biomedical Agreement.

References

Abouhamad,W.N., Manson,M., Gibson,M.M. and Higgins,C.F. (1991) *Mol. Microbiol.*, **5**, 1035−1047.

Alloing,G., Trombe,M.-C. and Claverys,J.-P. (1990) *Mol. Microbiol.*, **4**, 633−644.

Andrews,J.C. and Short,S.A. (1986) *J. Bacteriol.*, **165**, 434−442.

Bangalore,N., Travis,J., Onunka,V.L., Pohl,J. and Shafer,W.M. (1990) *J. Biol. Chem.*, **265**, 13584−13588.

Benjamin,W.H.J., Hall,P. and Briles,D.E. (1991) *Microb. Pathog.*, **11**, 289−295.

Boman,H.G. (1991) *Cell*, **65**, 205–207.
Castilho,B.A., Olfson,P. and Casadaban,M. (1984) *J. Bacteriol.*, **158**, 488–495.
Chen,C., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B. (1986) *Cell*, **47**, 381–389.
Davidson,A.L., Shuman,H.A. and Nikaido,H. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2360–2364.
Deverson,E.V., Gow,I.R., Coadwell,W.J., Monaco,J.J., Butcher,G.W. and Howard,J.C. (1990) *Nature*, **348**, 738–741.
Diamond,G., Zasloff,M., Eck,H., Brasseur,M., Maloy,W.L. and Bevins,C.L. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 3952–3956.
Dorman,C.J., Chatfield,S., Higgins,C.F., Hayward,C. and Dougan,G. (1989) *Infect. Immun.*, **57**, 2136–2140.
Du Pasquier,L. (1992) *APMIS*, **100**, 383–392.
Eisenhauer,P.B., Harwig,S.S. and Lehrer,R.I. (1992) *Infect. Immun.*, **60**, 3556–3565.
Fields,P.I., Groisman,E.A. and Heffron,F. (1989) *Science*, **243**, 1059–1062.
Ganz,T., Selsted,M.E. and Lehrer,R.I. (1990) *Eur. J. Haematol.*, **44**, 1–8.
Gibson,M.M., Price,M. and Higgins,C.F. (1984) *J. Bacteriol.*, **160**, 122–130.
Goodell,E.W. and Higgins,C.F. (1987) *J. Bacteriol.*, **169**, 3861–3865.
Groisman,E.A. (1991) *Methods Enzymol.*, **204**, 180–212.
Groisman,E.A. and Casadaban,M.J. (1986) *J. Bacteriol.*, **168**, 357–364.
Groisman,E.A., Chiao,E., Lipps,C.J. and Heffron,F. (1989a) *Proc. Natl Acad. Sci. USA*, **86**, 7077–7081.
Groisman,E.A., Heffron,F. and Solomon,F. (1992b) *J. Bacteriol.*, **174**, 486–491.
Groisman,E.A., Parra,C.A., Salcedo,M., Lipps,C.J. and Heffron,F. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11939–11943.
Groisman,E.A., Sturmoski,M.A., Solomon,F., Lin,R. and Ochman,H. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 1033–1037.
Higgins,C.F. (1992) *Annu. Rev. Cell Biol.*, **8**, 67–113.
Hiles,I.D. and Higgins,C.F. (1986) *Eur. J. Biochem.*, **158**, 561–567.
Hiles,I.D., Gallagher,M.P., Jamieson,D.J. and Higgins,C.F. (1987) *J. Mol. Biol.*, **195**, 125–142.
Jamieson,D.J. and Higgins,C.F. (1984) *J. Bacteriol.*, **160**, 131–136.
Lambert,J. *et al.* (1989) *Proc. Natl Acad. Sci. USA*, **86**, 262–266.
Lee,J.Y., Boman,A., Sun,C.X., Anderson,M., Jornvall,H., Mutt,V. and Boman,H.G. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9159–9162.
Lehrer,R.I., Barton,A., Daher,K.A., Harwig,S.S.L., Ganz,T. and Selsted,M.E. (1989) *J. Clin. Invest.*, **84**, 553–561.
Lehrer,R.I., Ganz,T. and Selsted,M.E. (1991) *Cell*, **64**, 229–230.
Macias,E.A., Rana,F., Blazyk,J. and Modrzakowski,M.C. (1990) *Can. J. Microbiol.*, **36**, 582–584.
McGrath,J.P. and Varshavsky,A. (1989) *Nature*, **340**, 400–404.
Miller,J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Miller,S.I., Kukral,A.M. and Mekalanos,J.J. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5054–5058.
Miller,S.I., Pulkkinen,W.S., Selsted,M.E. and Mekalanos,J.J. (1990) *Infect. Immun.*, **58**, 3706–3710.
Mims,C.A. (1987) *The Pathogenesis of Infectious Disease.* Academic Press, Inc., Orlando, FL.
Monaco,J.J., Cho,S. and Attaya,M. (1990) *Science*, **250**, 1723–1726.
Moore,K.S., Bevins,C.L., Brasseur,M.M., Tomassini,N., Turner,K., Eck,H. and Zasloff,M. (1991) *J. Biol. Chem.*, **266**, 19851–19857.
Ouellette,A.J., Miller,S.I., Henschen,A.H. and Selsted,M.E. (1992a) *FEBS Lett.*, **304**, 146–148.
Ouellette,A.J., Selsted,M.E., Miller,S.I. and Henschen,A.H. (1992b) *J. Cell Biol.*, **118**, 929–936.
Parham,P. (1992) *Nature*, **357**, 193–194.
Park,J.T. (1993) *J. Bacteriol.*, **175**, 7–11.
Patrick,G.G. and Hutchinson,C.R. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8553–8557.
Payne,W.J. and Gilvarg,C. (1968) *J. Biol. Chem.*, **243**, 6291–6299.
Perego,M., Higgins,C.F., Pearce,S.R., Gallagher,M.P. and Hoch,J.A. (1991) *Mol. Microbiol.*, **5**, 173–185.
Pereira,H.A., Erdem,I., Pohl,J. and Spitznagel,J.K. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4733–4737.
Rana,F.R., Macias,E.A., Sultany,C.M., Modzrakowski,M.C. and Blazyk,J. (1991) *Biochemistry*, **30**, 5858–5866.
Riordan,J.R. *et al.* (1989) *Science*, **253**, 1066–1073.
Rudner,D.Z., LeDeaux,J.R., Ireton,K. and Grossman,A.D. (1991) *J. Bacteriol.*, **173**, 1388–1393.
Selsted,M.E., Tang,Y.-Q., Morris,W.L., McGuire,P.A., Novotny,M.J., Smith,W., Henschen,A.H. and Cullor,J.S. (1993) *J. Biol. Chem.*, **268**, 6641–6648.
Shuman,H.A. (1982) *J. Biol. Chem.*, **257**, 5455–5461.
Sidén,I. and Boman,H.G. (1983) *J. Bacteriol.*, **154**, 170–176.
Speiser,D.M. and Ames,G.F.-L. (1991) *J. Bacteriol.*, **173**, 1444–1451.
Spies,T., Bresnahan,M., Bahram,S., Arnold,D., Blanck,G., Mellins,E., Pious,D. and DeMars,R. (1990) *Nature*, **348**, 744–747.
Steiner,H., Hultmark,D., Engström,Å., Bennich,H. and Boman,H.G. (1981) *Nature*, **292**, 246–248.
Tabor,S. and Richardson,C.C. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 1074–1078.
Tam,R. and Saier,M.H. (1993) *Microbiol. Rev.*, **57**, 320–346.
Trowsdale,J., Hanson,I., Mockridge,I., Beck,S., Townsend,A. and Kelly,A. (1990) *Nature*, **348**, 741–744.
Vaara,M. (1981) *J. Bacteriol.*, **148**, 426–434.
Vaara,M. (1992) *Microbiol. Rev.*, **56**, 395–411.
Walker,J.E., Saraste,M., Runswick,M.J. and Gay,N.J. (1982) *EMBO J.*, **1**, 945–951.
Yamauchi,K., Tomita,M., Giehl,T.J. and Ellison,R.T.,III (1993) *Infect. Immun.*, **61**, 719–728.
Zasloff,M. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 5449–5453.
Zasloff,M. (1992) *Curr. Opin. Immunol.*, **4**, 3–7.

Received on June 2, 1993; revised on July 13, 1993