# **ABC** Transporters: Bacterial Exporters

MICHAEL J. FATH<sup>†</sup> AND ROBERTO KOLTER<sup>\*</sup>

Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115

INTRODUCTION	996
Requirement for an ATP-Binding Domain	996
OVERVIEW OF ABC TRANSPORTERS	996
Bacterial ABC Importers (Periplasmic Permeases)	996
Eukaryotic ABC Transporters	997
Bacterial ABC Exporters	997
Accessory factors	998
List of bacterial export systems	999
PROTOTYPE BACTERIAL ABC EXPORTERS	999
Protein Transport Systems	999
E. coli alpha-hemolysin. (i) Sequence information	999
(ii) HlyA secretion signal	999
(iii) Structures of HlyB and HlyD	.1000
(iv) Functional complementation	.1000
(v) HlyA secretion pathway	.1000
Erwinia chrysanthemi proteases	1000
Peptide Transporters	.1001
E. coli colicin V	1001
Nonprotein Substrates	1002
Agrobacterium tumefaciens ChvA and Rhizobium meliloti NdvA	1002
Capsular polysaccharide exporter— <i>E. coli</i> KpsMT	1002
Incomplete Systems	1002
SEQUENCE AND PHYLOGENETIC ANALYSIS OF BACTERIAL ABC EXPORTERS	1003
INSIGHTS FROM BACTERIAL ABC EXPORTERS	1003
How Do ABC Transporters Recognize Such a Wide Range of Substrates?	1003
Why Are Dedicated Exporters Required?	1005
What Does the Complex Look Like?	1005
How Does the Export Complex Interact with Substrate?	1006
CONCLUSIONS	1006
THE SURVEY	1006
Proteus vulgaris and Morganella morganii Hemolysins	1006
Actinobacillus Hemolysins and Leukotoxins	1007
Pasteurella haemolytica Leukotoxin	1007
Bordetella pertussis Cyclolysin	1007
Pseudomonas aeruginosa Alkaline Protease	1007
Serratia marcescens Zinc Metalloprotease	1008
Rhizobium leguminosarum NodO	1008
Neisseria meningitidis FrpA and FrpC	1008
Enterococcus faecalis Bacteriocin/Hemolysin	1008
Bacillus subtilis Subtilin	.1008
Lactococcus lactis Nisin	1009
Staphylococcus epidermidis Epidermin	
Staphylococcus gallidermidis Gallidermin	1009
Streptococcus pneumonae Competence Factor	1009
Pediococcus acidilactici Pediocin PA-1	1009
Lactococcus lactis Lactococcus A and G	1009
Escnericnia coli Microcin B17	1003
Streptomyces peucetius UTTAB	1010
NAME AND ADDRESS AND A MARKED AND A M	1010
naemonias injuenzue DexAD	1010
Neisseria meningitidis CtrCD	1010 1010 1010
Neisseria meningitidis CtrCD Bradyrhizobium japonicum CycVW	1010 1010 1010 1010
Neisseria meningitidis CtrCD Bradyrhizobium japonicum CycVW Rhodobacter capsulatus HelABC Arabaras Sa. Staria BCC 7120 Hand	1010 1010 1010 1010 1010 1010

<sup>\*</sup> Corresponding author. † Present address: Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, IL 60637.

Escherichia coli SurB	
Escherichia coli MsbA	
Escherichia coli FtsE	
Streptomyces griseus AmfA and AmfB	
Rhizobium leguminosarum and Bradyrhizobium japonicum NodLJ	
Rhizobium meliloti ORF1	
Staphylococcus epidermidis MsrA	
Streptomyces TlrC, SrmB, and CarA	
ACKNOWLEDGMENTS	
ADDENDUM IN PROOF	
REFERENCES	

# INTRODUCTION

Proteins, peptides, polysaccharides, and many other molecules that are synthesized in the bacterial cytoplasm must often cross one or more membranes to reach their final destination. Many bacterial proteins are transported across the cytoplasmic membrane via the Sec pathway (139, 151, 191). Sec-dependent secretion requires that the secreted product be a protein with an N-terminal signal sequence, severely limiting the type of molecule that can be transported across membranes by using this system. Nonproteinaceous secreted products as well as extracellular proteins from gram-negative bacteria (which must cross both inner and outer membranes) cannot use the Sec pathway. In addition, there appear to be structural features of some proteins that are inherently incompatible with use of the Sec pathway (100, 116). Therefore many molecules must find another way to leave the cytoplasm. It has become apparent from results obtained in the last few years that this problem is often solved by the existence of dedicated export systems that facilitate membrane translocation with a large degree substrate specificity (80, 121, 140, 184).

Two major groups of dedicated export systems have been identified recently from gram-negative bacteria: the ABC transporters and the pullulanase-like family of transporters (80, 139, 150). This review will focus on the bacterial ABC transporters and will include information on transporters from both gram-negative and gram-positive systems.

#### **Requirement for an ATP-Binding Domain**

A feature common to dedicated exporter systems is that they, like the Sec system, require energy to facilitate translocation. Although the precise mechanism of secretion in these systems is still largely unknown, many of them appear to utilize ATP hydrolysis as a source of energy. Systems that use ATP require a component that binds ATP. Structural domains that can efficiently bind ATP and facilitate its hydrolysis have been highly conserved throughout evolution, and therefore it is not surprising that proteins making up the dedicated export systems display a highly conserved ATP-binding motif. This motif is known by several names including the Rossman fold (148), the Walker motif (182), and the Doolittle motif (28) and consists of two conserved sites (A and B) that form an ATP-binding pocket (148) (see Fig. 3). The ATP-binding site occurs at the end of an  $\alpha$ -helix: the residues GXGKST form a turn, bringing the lysine residue in close proximity to the phosphates in the Mg<sup>2+</sup>-ATP. The aspartic acid residue in the B site is in close proximity in space to the A site, and its negative charge may interact with the Mg<sup>2+</sup> molecule (182). The consensus sequence was defined by comparing distantly related sequences in the alpha and beta subunits of ATP synthase,

myosin, and many kinases. This consensus in the amino acid sequence involves a very small number of residues and is likely to be present also in proteins that do not bind nucleotides (11, 28). Thus, proteins that display this consensus need not be functionally homologous.

The bacterial permeases, which are multiprotein complexes involved in nutrient uptake, have been extensively analyzed and shown to have homologous protein components with ATP-binding consensus sites (2). The permease systems all have a conserved component which includes a 200-amino-acid (aa) region that displays a high degree of sequence identity and that contains the ATP-binding consensus sites A and B. This extended region of similarity clearly distinguishes these permease components from other nucleotide-binding proteins such as ATP synthase, myosin, and kinases. This extended region is also present in many bacterial proteins that form part of dedicated export systems. This observation serves as the unifying theme presented in this review. Several excellent reviews describing other aspects of the family of proteins containing this conserved region have been written recently (3, 8, 80)

This highly conserved region has been called the ATPbinding cassette (ABC), and proteins which contain this region are called ABC transporters (82) or, alternatively, traffic ATPases (3). For the purposes of this review, we subdivide the superfamily of proteins that contain the ABC into three subfamilies based not on evolutionary differences but, rather, on differences in their source and generalized function as an importer or exporter. Each subfamily is described briefly, but the main emphasis here is placed on the description of the ABC exporters from bacteria.

# **OVÉRVIEW OF ABC TRANSPORTERS**

# Bacterial ABC Importers (Periplasmic Permeases)

The bacterial periplasmic permeases compose a large subfamily of ABC transporters and have been carefully studied (2, 83). These systems have several distinguishing features. They are all multisubunit import systems with a similar structural organization (Fig. 1). They all include a hydrophilic, membrane-associated protein containing an ABC type of ATP-binding domain. In several systems, this conserved component has been shown to associate tightly with two hydrophobic proteins (102). A unique feature of the bacterial ABC importers is that they all have a periplasmic binding protein that interacts with the incoming substrate, binds to it, and presents it to the import complex in the inner membrane. Bacterial ABC importers have the ATP-binding domain and the membrane-spanning domains (MSDs) present on separate polypeptides. In contrast, most but not all of the ATP-binding domains in the bacterial exporter systems are present on the same polypeptide as the MSDs.



FIG. 1. Structural models of various ABC transporters. The prototype systems included are the *E. coli* alpha-hemolysin exporter, the *E. coli* polysialic acid exporter, the *S. typhimurium* histidine importer, the *B. subtilis* subtilin exporter, and the mammalian P-glycoprotein drug exporter. The bacterial exporters are drawn as dimers, consistent with the model of Higgins (80) and others, who propose a minimum of four required "core components." There is no experimental evidence that the bacterial export complexes form dimers. The core components in each complex are shaded.

# **Eukaryotic ABC Transporters**

A second subfamily is composed of the ABC transporters found in eukaryotes. All the eukaryotic ABC transporters have their ATP-binding domain on a single polypeptide with the MSDs. In fact, most of the eukaryotic ABC transporters have a tandem duplication of the structure and do not appear to require other subunits for their function (Fig. 1). Several of these systems are of significant medical importance and have been under intense study since their discovery. These include the multidrug resistance protein, P-glycoprotein, which exports chemotherapeutic drugs from tumor cells when overexpressed (35) and the cystic fibrosis transmembrane regulator found to be defective in patients with cystic fibrosis (145). Other eukaryotic ABC exporters (reviewed in reference 80) include pfMDR, which exports antimalarial drugs from Plasmodium falciparum (44), STE6, which exports a-type mating factor from Saccharomyces cerevisiae (113, 127), and a group of transporters involved in antigen presentation (27, 128, 167, 178). Other, more recently identified eukaryotic ABC transporters include atpgp, an Arabidopsis thaliana P-glycoprotein homolog (30); pmd1, the leptomycin B resistance gene from Schizosaccharomyces pombe (133); hmt1, which encodes a cadmium-specific phytochelatin in S. pombe (135); LEMDR06 and LEMDRF2, two multidrug resistance genes in *Leishmania donovani* (75); SNQ2, a quinoline resistance gene in S. cerevisiae (160); NG-TRA, which is a putative hormone transporter expressed in adrenal glands (6); and ALD, the putative X-linked adrenoleukodystrophy gene in humans (130).

#### **Bacterial ABC Exporters**

The third subfamily of ABC proteins is made up of the bacterial ABC exporters, the largest and fastest-growing group. There are over 40 identified systems (Table 1). A few ABC proteins (FtsE, HepA, MsbA, SurB, Orf1) are included in Table 1 because their structure is consistent with their having a translocation function, even though their role as an exporter has not been demonstrated. We describe the common features of these ABC exporters with particular emphasis on key prototype systems. At the end, we present a survey of all the known and putative bacterial ABC exporters with the hope that it can serve as an easy reference for those who wish to investigate a particular system in more detail.

All ABC transporters in the bacterial export subfamily have the conserved ATP-binding motif. The domain contain-

#### 998 FATH AND KOLTER

Transported substrate	ABC homolog	Trans	locator structural org	ganization		Organism	Refer- ence
		ABC	MSD	Ar-	UMF		
Proteins			TH-D N		Talo	Frank suisbier and	42-
HlyA (alpha-nemolysin)	HIYB Ub-D	HIYB-C	HIYBN		TolC	Escherichia coli Protovo vulcorio	42a
HlyA (alpha-nemolysin)	HIYB Ub-D	HIYBC	HIYBN	HIYD	TOIC	Proteus vuigaris Monogradila monografii	107
Ann A (homolysin)	AnnP	Ann P C		AppD		Morganella morganu	105
AppA (nemolysin)	Арры	Арры—С	Аррь—іч	Арри		niae	17
AshA (hemolysin)	AshB	NA	NA	AshD		Actinobacillus suis	12
LktA (leukotoxin)	LktB	LktB—C	LktB—N			Actinobacillus actinomycetem- comitans	68
LktA (leukotoxin)	LktB	LktBC	LktB—N	LktD		Pasteurella haemolytica	171
CyaA (cyclolysin)	CyaB	CyaB—C	CyaB—N	CyaD	CyaE	Bordetella pertussis	63
PrtA,B,C (proteases A, B, C)	PrtD	PrtD-C	PrtD-N	PrtE	PrtF	Erwinia chrysanthemi	117
AprA (alkaline protease)	AprD	AprD-C	AprD—N	AprE	AprF	Pseudomonas aeruginosa	31, 69
Peptides	<b>a b</b>	a		<u> </u>	<b>T</b> 10		<i>(</i> <b>)</b>
CvaC (colicin V)	CvaB	CvaB_C	CvaB—N	CvaA	TolC	Escherichia coli	62
CylL (hemolysin/bacteriocin)	CylB	CyiB_C	CylB—N	6D		Enterococcus faecalis	60 10 102
SpaS (subtilin)	<b>Spa</b> В	SpaB-C	SpaB-N	SpaD		Baculus subtuis	18, 103
NISA (IIISIN)	INIS I	NIST-C	NISI —N			Lactococcus lactis of 3	30
EpiA (epidermin)	Com	Epil Com A C	Epil Com A N	ComB		Staphylococcus epiaermiais	155
CF (competence factor)		ComA—C	COMA-N	COMB		Streptococcus pneumoniae	92
Len A (lectococcin A)	LonC	readC	LenC N	LanD		Lastososous lastis suban lastis	124
McbA (microcin B17)	McbEF	McbF	McbE	LCIID		Escherichia coli	51
Nonprotein substrates							
Capsular polysaccharide	KpsMT	KpsT	KpsM			Escherichia coli K5	164
Capsular polysaccharide	KpsMT	KpsT	KpsM			Escherichia coli K1	138
B-1,2-glucan	NdvA	NdvA—C	NdvA—N			Rhizobium meliloti	168
β-1,2-glucan	ChvA	ChvA-C	ChvA—N			Agrobacterium tumefaciens	13
Daunorubicin/doxorubicin	DrrAB	DrrA	DrrB			Streptomyces peucetius	67
Capsular polysaccharide	BexABC	BexA	BexB, BexC?			Haemophilus influenzae	110
Capsular polysaccharide	CtrBCD	CtrD	CtrC, CtrD?			Neisseria meningitidis	490
Heme for cytochrome c	CycVWX	CycV	CycW, ORF263			Bradyrhizobium japonicum	143
Heme for cytochrome c	HelABC	HelA	HelB, HelC			Rhodobacter capsulatus	7
Erythromycin resistance	MsrA	MsrA—N,C				Staphylococcus epidermidis	147
Tylosin resistance	TlrC	TlrC-N,C				Streptomyces fradiae	149
Macrolide resistance	SrmB	SrmB-Ń,C				Streptomyces ambofaciens	156
Macrolide resistance	CarA	CarA—N,C				Streptomyces thermotolerans	156
Incomplete systems (substrate or							
transporter not identified)							
? (envelope polysaccharide)	HepA	HepA—C	HepA—N			Anabaena sp. strain PCC7120	86
? (required to restart growth)	SurB	SurB-C	SurB-N			Escherichia coli	162
? (suppresses <i>htrB</i> phenotype)	MsbA	MsbAC	MsbA—N			Escherichia coli	98
? (aerial mycelium formation)	AmtA	AmtAC	AmfA—N			Streptomyces griseus	178a
? (aerial mycelium formation)	AmfB	AmfBC	AmfB—N			Streptomyces griseus	178a
(essential, cell division)	FISE	FtsE	NT 17			Escherichia coli	56
? (lipooligosaccharides?)	Nodlj	Nodi	NodJ			Rhizobium leguminosarum	37
(inpooligosaccharides?)	NOOLI	INODI	NOdJ			Bradyrhizobium japonicum	181
(downstream of NtrA)	UKFI	UKFI				Knizobium meliloti	1
ReteM (ring metallemetal)	N.A.	INA	INA			Knizobium leguminosarum	34
rnom (zinc metalloprotease)	N.A.	INA	NA			Serratia marcescens	118,
Em A (En binding protoic)	NT A	<b>NT 4</b>	N7.4			<b>x</b> 7 · · · · · · · · · · · · · · · · · · ·	172
ColA (collidormin)	N.A.	INA	INA			Neisseria meningitidis	174
L cnG (lactococcin G)	Gall Lon <sup>2</sup>	INA NA	INA NA			Staphylococcus gallidermidis	154
Laciococcin G)	LCII?	INA	INA			Lactococcus tactis	/4

TABLE 1. List of bacterial ABC exporters

<sup>*a*</sup> AF, accessory factor; OMF, outer membrane factor. <sup>*b*</sup> NA, not applicable.

ing this motif can be on the same polypeptide as the MSDs (as in the eukaryotic ABC exporters) or on a polypeptide separate from the hydrophobic domains (as in the bacterial ABC importers). Every bacterial ABC exporter has the components necessary to form the conserved structural organization described in Fig. 1. Notable exceptions are the antibiotic resistance proteins CarA, MsrA, SrmB, and TlrC, which have two ABC cassettes and no identified MSDs.

Accessory factors. Many of the bacterial ABC exporters require additional proteins, besides the protein(s) which include the ABC and the MSD, to form a functional complex. These additional factors have been identified in several gram-negative systems and shown to be needed when the secreted product is destined for immediate release into the extracellular medium. These proteins are referred to here as accessory factors. In several of the prototypes, the accessory factor fractionates mostly to the inner membrane, with minor amounts fractionating to the outer membrane (24, 123), consistent with the hypothesis that the accessory factor is anchored in the inner membrane and spans the periplasm. The accessory factor probably acts to connect the inner and outer membranes and functions to facilitate the export of products through both membranes of the gram-negative cell. The gene encoding the accessory factor is always found linked to the gene encoding the ABC protein.

Many gram-negative bacterial ABC exporters do not have accessory factors, notably those involved in the transport of nonprotein substrates. In these systems, the final destination of the exported product is the periplasm or outer membrane. In a few cases, additional proteins are required to complete translocation of the product into the outer membrane. However, in these systems, outer membrane transport is a distinct process and the proteins involved are different from the accessory factor proteins described above. Several of the gram-positive secretion systems have linked sequences that have some similarity to the accessory factors characterized in gram-negative bacteria (18, 103, 129, 170). However, in the export process of gram-positive bacteria, the secreted product has only one membrane to traverse, so the role of the accessory factor in gram-positive systems is not clear.

A third export-related gene is also observed in several of the gram-negative export systems (63, 117, 185). This gene encodes an outer membrane protein required for secretion in several characterized systems. In Table 1, the third component is referred to as the outer membrane factor. The gene encoding the outer membrane factor can be found either closely linked to the other export genes (63, 117) or physically quite distant (185). In the latter case, the outer membrane factor has been shown to have cellular functions apart from its export function (185). Lastly, genes involved in synthesis of lipopolysaccharide also seem to be required for protein export in several gram-negative systems, suggesting that lipopolysaccharide is also involved in the export process (187).

List of bacterial export systems. Table 1 lists all known bacterial ABC exporters, including several putative exporters. The list is divided into three major groups based on the characteristics of the transported product (and not on phylogeny of the ABC transporter, which is discussed later). These substrate groups are proteins, peptides, and nonprotein substrates. Some other systems have been included even though they have not been completely described.

Table 1 further describes these systems by including the components identified in each case which make up the transporter apparatus. Since there are different structural combinations among the bacterial ABC exporters, the structural organization is broken down into distinct units, the ABC, MSD, accessory factor, and outer membrane factor.

### **PROTOTYPE BACTERIAL ABC EXPORTERS**

# **Protein Transport Systems**

The best-characterized class of bacterial exporters is involved in the extracellular secretion of a family of proteins called the RTX (repeats in toxin) toxins (19) and the related extracellular proteases (183). At present, 10 protein export systems have been identified and characterized from gramnegative bacteria. These RTX protein transport systems share many common features: they all have a large ABC transporter (600 to 750 aa) that contains N-terminal MSDs and a C-terminal ABC, and they all require an accessory factor and an outer membrane factor. The transported product in each of these systems is a large protein ranging in size from 50 kDa (protease A of *Erwinia chrysanthemi*) to 216 kDa (the *Bordetella pertussis* cyclolysin), and each product appears to be transported into the extracellular medium without any detectable periplasmic intermediate. The translocation complex recognizes an export signal located in the C-terminal 60 to 150 residues of the transported protein.

E. coli alpha-hemolysin. (i) Sequence information. The Escherichia coli alpha-hemolysin determinant was the first ABC export system identified in prokaryotes (42a). It has been sequenced and characterized now from several sources including the chromosome of uropathogenic strain J96 (42a), strain LE2001 (96), and plasmid pHly152 (78). It is highly conserved in these strains. Each hemolysin determinant contains an operon encoding four proteins, HlyCABD. The organization of the Hly operon is shown in Fig. 2. hlyA encodes the 1,023-aa (107-kDa) alpha-hemolysin (HlyA), hlyB encodes the 707-aa ABC exporter, hlyD encodes the 477-aa accessory factor, and hlyC encodes a 170-aa protein that has no secretion function but facilitates the activation of HlyA by addition of a fatty acid group (94). The outer membrane factor for alpha-hemolysin is the chromosomally encoded, 495-aa TolC protein (132, 185).

A large amount of genetic and biochemical data has been accumulated for the *E. coli* hemolysin, and these results are the subject of several detailed reviews (8, 9, 19, 65, 87, 88, 91, 190). Here, we will summarize several key points and discuss recent results.

(ii) HlyA secretion signal. The secretion of HlyA has been shown to require neither an N-terminal signal sequence nor the secA gene (53, 65). By using deletions, gene fusions, and point mutants, the secretion signal in HlyA has been localized to the C-terminal 48 to 60 amino acids (77, 100, 101, 108, 169). This C-terminal secretion signal does not resemble a typical N-terminal signal sequence, nor is it removed during HlyA secretion (42). There is surprisingly little primary sequence conservation in the C-terminal secretion signals of HlyA and the other RTX toxins (87, 91, 101, 169). Instead, several secondary structure features—the  $\alpha$ -helical region, the aspartate box, and the amphipathic helix-have been proposed to be important for signal recognition (87, 91, 101, 169). Experiments with protein fusions have shown that the C-terminal signal in HlyA can be used to secrete heterologous proteins out of the cell in a HlyBD-dependent fashion (88, 100, 122). These include proteins that normally cannot be secreted by the Sec-dependent pathway. Over 100 mutations have been generated in the C-terminal region by several research groups. Several conclusions can be drawn from the analysis of these mutants. (i) The HlyA signal region is very tolerant of mutations. In fact, no missense mutations were found that completely abolished HlyA secretion. (ii) Surprisingly, mutations which should alter the proposed secondary structures conserved among the RTX toxins do not greatly affect HlyA secretion. (iii) Several of the putative structural domains in the C-terminal signal which were thought to be critical (the  $\alpha$ -helical region, the aspartate box, and the amphipathic helix) are actually not essential for HlyA secretion. (iv) There appear instead to be



FIG. 2. Organization of selected bacterial ABC exporter operons. The systems included are *E. coli* alpha-hemolysin (*hlyCABD*) (42a), *Erwinia chrysanthemi* protease (*inh*, *prtDEFBCA*) (117), *E. coli* colicin V (*cvaABC*, *cvi*) (62), *B. subtilis* subtilin (*spaDBCS*) (18) (called *spaBTCS* by Klein et al. [103]), and the *E. coli* capsular polysaccharide transporter (*kpsMT*) (138, 164). The *tolC* gene is unlinked to the alpha-hemolysin and the colicin V operons. The symbol for the *spaD* gene means that *spaD* has some sequence similarity to other accessory factors but may not be part of the functional exporter. The gene sizes are roughly to scale. See the text for more details of these systems.

important contact residues scattered throughout the C-terminal secretion signal (101, 169).

(iii) Structures of HlyB and HlyD. The HlyB and HlyD exporters have been studied in some detail. Both proteins fractionate primarily to the inner membrane (96, 123, 157, 173, 188). Topological analysis of HlyD shows that it has one transmembrane domain and that most of the protein is on the periplasmic side of the membrane (157, 188). These results are consistent with the sequence hydrophobicity data and with the model that predicts that HlyD spans the periplasm as part of the secretion complex. The HlyB protein sequence can be divided into three large domains, an N-terminal hydrophilic domain of about 150 aa, a 275-aa central hydrophobic domain, and a 275-aa C-terminal domain that contains the ABC. Hydropathy analysis of HlyB predicts that it would have six transmembrane domains, all localized within the central hydrophobic domain. This would be consistent with the predicted topology of all the other ABC transporters. However, experiments from two research groups provide evidence that, instead, HlyB contains eight transmembrane domains: two in the N-terminal domain and six in the central hydrophobic domain (52, 188). The topological models predicted in these papers differ from models based on hydropathy analysis and from each other. Therefore, at this point it is difficult to conclusively identify the topological domains of HlvB.

HlvB has been difficult to characterize biochemically because of its membrane localization and low yield. Recently, HlyB has been tagged with epitopes from P-glycoprotein to allow further biochemical analysis (96). In addition, antibodies to HlyB have been successfully generated (173). Both of these studies confirmed the inner membrane localization of HlyB in vivo. Mutagenesis of HlyB has been carried out to characterize its function and to identify regions which may interact with HlyA. Linker insertion mutations and point mutations in the MSD of HlyB have been shown to affect HlyA secretion levels (8, 9, 96). Surprisingly, a deletion that removes the MSDs in the N-terminal 467 aa of HlyB and replaces them with part of TetC appears to be able to cause secretion of low levels of HlyA in an HlyBD-dependent fashion (173). This suggests that some secretion signals are in the MSD but that residual secretion activity is retained within the C-terminal ABC region

(iv) Functional complementation. Since the HlyA substrate is able to tolerate many mutations and still remain functional for secretion, the HlyBD exporter must have a relaxed specificity for export of HlyA-like molecules. This hypothesis is confirmed by looking at the ability of HlyBD to functionally complement mutations in exporters from related ABC export systems. HlyBD have been shown to efficiently secrete many related proteins including *Morganella morganii* and *Proteus vulgaris* HlyA (105), AppA (72), LktA (15, 84, 171), CyaA (125, 159), and NodO (152). At a lower level, HlyBD can also secrete ColV (39), AprA (69), PrtSM (118, 172), and PrtB (23). Thus, many proteins that are secreted through a dedicated ABC exporter can be transported by HlyBD, even though some of them have very little sequence similarity with HlyA.

(v) HlyA secretion pathway. The mechanism of HlyA secretion by HlyBD has been carefully studied. In cells carrying the wild-type HlyBD exporters, HlyA can be found in the cytoplasm, the outer membrane, and the culture supernatant. No periplasmic HlyA has ever been detected during secretion (43, 66, 108, 134). In cells that lack either HlyBD or HlyB alone, HlyA localizes to the cytoplasm and the inner membrane; this suggests that some HlyA can target to the membrane independently of HlyBD. HlyA from cells lacking HlyD are also cytoplasmic (65). However, when HlyB is expressed along with two-thirds of HlyD, HlyA targets to the outer membrane, suggesting that this HlyD derivative still retains partial secretion functions (134). The secretion process can be divided into early and late stages on the basis of energy requirements (106). Early steps, possibly HlyA binding to the membrane secretion complex, require the proton motive force. In contrast, late translocation to the outer membrane is independent of the proton motive force, perhaps relying on ATP hydrolysis by HlyB (106). A threestep model for HlyA secretion has been proposed (91, 169): (i) the C terminus of HlyA associates with the inner membrane; (ii) once in the membrane, HlyA interacts with HlyB; and (iii) a secretion complex that includes HlyB, HlyD, and TolC then facilitates secretion directly through both membranes.

*Erwinia chrysanthemi* proteases. The protease secretion system of *E. chrysanthemi* shares most of the functional features with the hemolysin system described above, such as an ABC export system and C-terminal secretion signal. In addition to these features, the extracellular protease systems from *E. chrysanthemi*, *Pseudomonas aeruginosa*, and *Serratia marcescens* share additional commonalities and compose a distinct subgroup of ABC exporters (183). E. chrysanthemi is a phytopathogenic enterobacterium that secretes three related proteases, PrtA, PrtB, and PrtC. The genes required for this function were cloned and expressed in E. coli (Fig. 2) (186). The proteases are distinct but closely related in sequence and size; PrtA is 50 kDa, PrtB is 53 kDa, and PrtC is 55 kDa. Most detailed studies have been carried out on PrtB, which displays slight sequence similarity with HlyA in the RTX repeat region. The deduced protein sequences also reveal that these proteases are processed to remove the N-terminal 16 to 18 residues (25, 55). These processed leaders are not Sec-dependent export signal sequences. Instead, secretion of the proteases requires the three linked genes prtD, prtE, and prtF (117). PrtD is the ABC exporter, PrtE is similar to HlyD, and PrtF has sequence similarity with TolC.

By using protein fusions and deletions of the PrtB protein, a minimum secretion signal region has been localized to the C-terminal 39 residues of PrtB (23). Thus, the proteases have a C-terminal signal similar to the HlyA secretion signal. PrtDEF and HlyBD were shown to have only slight functional conservation. Complementation experiments showed that HlyBD secretes PrtB at only less than 2% and that PrtDEF does not secrete detectable HlyA (23). More recently, CyaA-PrtB and HlyA-PrtB fusions were used to further characterize the C-terminal PrtB secretion signal (119). A large C-terminal region, which contained the RTX repeats, was required for the secretion of larger HlyA-PrtB and CyaA-PrtB fusions. The authors hypothesize that larger regions of PrtB might be necessary to facilitate the unfolding of the fusion polypeptide when it interacts with the PrtDEF secretion proteins (119).

The secretion proteins PrtD, PrtE, and PrtF have also been characterized when expressed in  $E. \ coli$  (24). PrtD is localized in the inner membrane, PrtE is localized to both the inner and outer membranes, and PrtF is localized in the outer membrane. This is consistent with the localization of HlyBD and TolC and provides independent evidence about the structure of the ABC-type secretion complexes. Protease accessibility studies also confirm that the Prt proteins have a membrane organization similar to that found in the Hly system, although they do not address the question of six or eight MSDs (24). The protease exporters are expressed at relatively high levels in  $E. \ coli$  and appear to be more amenable to biochemical analysis than the exporters from hemolysin and colicin V.

# **Peptide Transporters**

The second major group of bacterial transport systems facilitates the secretion of ribosomally encoded peptide antibiotics (for a review, see reference 104) and related small proteins. This group consists of at least seven systems from gram-positive bacteria and two from gram-negative bacteria. Several features distinguish these transport systems from the systems described above that secrete large proteins. The extracellular product in each of these systems is a small protein/peptide (<11 kDa) which lacks any RTX repeat motifs or any HlyA-like C-terminal export domain. In the one peptide secretion system where a secretion signal has been characterized, colicin V, this signal has been localized to the N-terminal region (62). The colicin V secretion signal shares some primary sequence conservation with the N terminus of lactococcin A, but this conservation does not extend to the other peptides in this subfamily. Most of the secreted peptides undergo significant processing and unusual posttranslational modifications that give rise to residues such

as dehydroalanine and lanthionine (45, 95). Peptides containing lanthionine have been termed the lantibiotics (95).

The ABC exporters responsible for peptide secretion are typically on one large polypeptide that contains both the ABC and an MSD, much like HlyB and PrtD. In only one example, the microcin B17 secretion system, the ABC and the MSD are on separate polypeptides (51). Several other features of the microcin B17 secretion system distinguish it from the other systems in this group. No accessory factor has been found, and during microcin B17 secretion, a periplasmic intermediate has been observed (22, 51).

The prototypical colicin V exporter requires both an accessory factor and outer membrane factor (62). In contrast, the systems from gram-positive bacteria, having no outer membrane barrier to cross, may be able to bypass the requirement for the additional components. Although obviously no outer membrane factor component can exist in the gram-positive systems, several of them do have linked genes that encode potential HlyD-like accessory factors (18, 103, 129, 170), and in the lactococcin A system a mutation in the accessory factor is found to abolish lactococcin A activity (170).

E. coli colicin V. Colicin V is an 88-residue peptide antibiotic whose production is encoded in large, low-copynumber virulence plasmids found in E. coli and other members of the Enterobacteriaceae (38, 189). Colicin V is active against a number of gram-negative bacteria and kills sensitive cells by disrupting their membrane potential (193). The genes encoding colicin V production, activity, and immunity were cloned from plasmid pColV-K30 and sequenced (61, 62). Four linked genes were identified: cvaC encodes the 103-aa pro-ColV, cvi encodes a 78-aa immunity protein; and cvaA and cvaB encode 413- and 698-aa secretion proteins, respectively (see Fig. 2). CvaC, the colicin V primary translation product, does not contain a Sec-dependent N-terminal signal sequence. However, the N-terminal 15 residues are removed concomitant with export (40). In addition, mass spectroscopy has shown that, unlike many other peptide antibiotics, colicin V does not undergo posttranslational covalent modifications (40). Secretion of colicin V from E. coli requires the ABC exporter CvaB, the accessory factor CvaA, and the unlinked outer membrane factor TolC (61, 62).

The secretion signal in CvaC was localized to the 39 N-terminal residues by using CvaC-PhoA fusions and point mutations (62). CvaC mutations G14D, G14N, and G38R all decrease secretion levels without affecting intracellular colicin activity. Active colicin V can also be secreted via the HlyBD and the PrtDEF secretion systems at lower efficiencies, despite the lack of similarity between the HlyA/PrtB secretion signals and the CvaC secretion signal (39). The HlyBD system was shown to recognize an N-terminal signal in CvaC, but the CvaC export mutants differentially affect export through HlyBD and CvaAB, suggesting differences in signal specificity between the two systems (39).

The colicin V secretion pathway has been characterized by genetic and biochemical methods. In one experiment, the N-terminal signal sequence from the OmpA protein was fused to the 88-aa mature colicin V peptide and shown to facilitate export of colicin V into the periplasm, showing that colicin V is not incompatible with the Sec-dependent pathway but that it could not pass through the outer membrane unaided (40). During CvaAB-mediated secretion, free colicin V was shown not to accumulate in the periplasm, and mutations in *cvaA*, *cvaB*, or *tolC* also do not result in free periplasmic colicin V. These results suggest that the three transport proteins form a single export complex that directly secretes colicin V into the extracellular medium, consistent with what is seen for the hemolysin secretion system (40). The colicin V system has been reviewed recently (38).

# **Nonprotein Substrates**

Many ABC export systems are known to secrete nonprotein molecules such as lipophilic drugs, antibiotics, and polysaccharides. Among the bacterial ABC exporters there is a surprisingly wide range of substrate structural diversity. In some systems, a single HlyB-like ABC exporter is sufficient to move polysaccharide to the outer membrane of *Rhizobium* and *Agrobacterium* cells (13, 168). In other systems involved in export of capsular polysaccharide, transporters that look very much like inverse periplasmic permeases are found (49, 110, 138, 164). Perhaps most surprisingly, antibiotic efflux systems have been described that lack any identifiable MSD (147, 149, 156). It is yet to be seen whether any unlinked MSDs will be identified for these efflux systems.

None of the secretion systems for nonprotein substrates have accessory factors like HlyD or outer membrane factors like TolC. Consistent with this observation, none of the systems secrete their product directly into the extracellular medium. In the capsular polysaccharide systems, additional proteins are required to complete translocation of the product into the outer membrane. However, in these examples outer membrane transport is a distinct process and the proteins involved are different from the accessory factor proteins described above.

Several members of this exporter subfamily have been reviewed recently and given the name of ABC-2-type transporters (144). It is interesting that these multicomponent systems have so much in common with the periplasmic permeases and yet clearly facilitate export. This observation reinforces the idea that all the ABC systems share a common structural organization and functional history and that there is not a clear structural distinction between exporters and importers (3).

Agrobacterium tumefaciens ChvA and Rhizobium meliloti NdvA. The plant pathogen Agrobacterium tumefaciens causes crown gall tumors when present on wounded dicotyledonous plants. Tn5 mutagenesis identified two chromosomal genes which were required for virulence, *chvA* and *chvB* (29). These mutants fail to produce extracellular  $\beta$ -1,2glucans. The  $\beta$ -1,2-glucans are oligomers 18 to 24 aa long with a molecular mass of 4 to 6 kDa (194, 195), and are involved in the attachment of the bacteria to plant cells. *chvB* mutants are avirulent and affect glucan production but not glucan secretion (141).

β-1,2-Glucans have also been implicated in the attachment of *Rhizobium meliloti* to legumes (54), a process essential for the formation of symbiotic nitrogen-fixing nodules. The *ndvA* and *ndvB* genes were isolated by hybridization to *chvA* and *chvB* and found to be required for nodule development (32). The *ndv* genes are functionally equivalent to their *chv* counterparts; they also function to produce and secrete β-1,2-glucans. *ndvA* mutants do not produce extracellular β-1,2-glucans, although the protein-sugar intermediate is observed in cytoplasmic extracts (168). Sequence analysis shows that *ndvA* encodes a 616-aa ABC export protein with a predicted molecular mass of 67.1 kDa. NdvA probably functions to export the 4- to 6-kDa β-1,2-glucan oligomers to the periplasm, from where they are then localized to the surface of the outer membrane. The *chvA* gene was sequenced soon afterward and found to encode another ABC exporter (13). ChvA is a 588-aa protein with a predicted molecular mass of 64.7 kDa. *chvA* mutations were also shown to accumulate glucan in the cytoplasm and to lack detectable glucan in the periplasm (13). ChvA and NdvA are highly homologous in function and sequence (76% amino acid identity). The discovery of NdvA and ChvA widened the family of ABC exporters to include systems that secrete nonproteinaceous molecules and to include proteins that function to facilitate secretion to the periplasm of gram-negative bacteria.

Capsular polysaccharide exporter-E. coli KpsMT. Many clinical strains of E. coli, Haemophilus influenzae, and Neisseria meningitidis produce capsular (K) polysaccharides that are major virulence determinants (10). To identify genes from E. coli serotypes K5 and K1 that encode the capsular polysaccharide determinant, cosmids from K1 and K5 were cloned into the noncapsular E. coli K12 and shown to be sufficient to allow K12 to produce capsule (10). The capsular polysaccharide determinant was shown to contain three phenotypically distinct genetic regions associated with polysaccharide expression. Region 1 is necessary for the transport of mature, lipid-linked polysaccharide across the outer membrane and its assembly into a capsule. Region 2 is serotype specific and encodes the enzymes for synthesis and polymerization of specific K antigen. Region 3 is involved in the energy-dependent translocation of the polysaccharide across the inner membrane (10).

Two genes have been identified and sequenced from region 3 of E. coli serotypes K1 and K5 and found to encode components of an ABC transporter. In K5, the kpsT gene encodes a 224-aa, 25.5-kDa ABC protein and kpsM encodes a 258-aa, 29.5-kDa hydrophobic protein (164). In K1, kpsT encodes a 219-aa, 24.9-kDa ABC protein and kpsM encodes a 258-aa, 29.6-kDa hydrophobic protein (Fig. 2) (138). Insertion mutations in kpsM or kpsT result in polysaccharide that is cytoplasmic and shorter than surface polymers (112). A mutation in the ATP-binding domain of K1 KpsT (K44E) results in a nonfunctional protein (138). Additional mutants with amino acid substitutions in the ATP-binding domains are also nonfunctional (163), and KpsT has also been shown to bind to 8-azido-ATP, providing biochemical evidence for ATP binding (163). PhoA fusions to KpsM suggest that it is membrane associated and that parts of it are exposed to the periplasm (138), and more extensive analysis of KpsM with fusions to  $\beta$ -lactamase support the model predicted from the hydropathy profile, i.e., that KpsM contains six MSD (163).

Genes in region 3 of serotypes K5, K7, K12, and K92 are able to functionally complement each other in *trans*, suggesting that kpsM and kpsT are functionally equivalent between serotypes (146). Apparently, the KpsMT exporter is capable of transporting a variety of acidic polysaccharides. These *E. coli* kps systems are highly homologous to the capsular polysaccharide genes of *H. influenzae* and *N. meningitidis* described below. The authors suggest that KpsMT acts to transport the growing polysaccharide chain to the periplasm, where it is then secreted to the surface by specific periplasmic proteins, such as KpsD, which are encoded by region 1 (163, 164).

#### **Incomplete Systems**

Several systems have only been partially described. These include systems where the ABC transport protein has been cloned and sequenced but a substrate for the transporter has not been clearly identified. They also include a few systems that contain a secreted protein that is homologous to a known ABC substrate such as hemolysin or epidermin. In these systems, characterization of the operon is not yet complete and putative ABC transporters that facilitate secretion are likely to be identified. Several such systems are described in the survey below.

# SEQUENCE AND PHYLOGENETIC ANALYSIS OF BACTERIAL ABC EXPORTERS

To better understand the evolutionary relationship between the bacterial ABC exporters, we carried out a detailed sequence comparison of the ABC regions from 29 systems. For this analysis, only the ABC domains were compared. Although the MSDs all appear to have similar hydrophobicity profiles, there is much less primary sequence conservation between the MSDs, making detailed sequence comparison more difficult.

The ABC sequence comparison results are shown in Fig. 3, 4, and 5. Figure 3 shows a multiple sequence alignment of the bacterial ABC exporters, including a consensus sequence. The degree of sequence identity varies markedly along the 190 aa in the ABC, suggesting the presence of distinct subdomains within the ABC. The degree of sequence similarity across the ABC domain is shown graphically in Fig. 4. The most highly conserved sequences cover the extended regions around the A and B sites, specifically residues 10 to 34 and 115 to 150. The central subdomain of the ABC has much lower primary sequence conservation, possibly resulting from differences in substrate specificity. Overall, the primary and secondary structures of the ABC domain are conserved between the various bacterial ABC exporters and throughout the entire superfamily of ABCtransport systems.

Analysis of the phylogenetic tree in Fig. 5 shows that the bacterial ABC exporters can be divided into two primary branches. One branch contains all the transport systems where the ABC is present on the same polypeptide as the MSD (group A). The other branch contains all the systems where the ABC and MSD are found on separate polypeptides (group B). We have previously shown that the ABC domains from mouse and human P-glycoprotein belong with the sequences in group A and that the ABC domains from several periplasmic permease systems (MalK, HisP, OppD, OppF, and PstB) as well as from Orf1, NodI, and FtsE belong with the sequences in group B (38). Therefore, the evolutionary relatedness of these domains seems to be more a function of their disposition with the MSD than with their substrate recognition. On the basis of these observations, two hypotheses can be made: the superfamily of ABC transporters originated from a shared common ancestor and the earliest branching of the superfamily resulted in two ABC protein families, those that became fused with an MSD and those that did not.

The transporters in group A appear to function only in export. Proteins in group A include the 17 sequences in Fig. 5 known to have a function in export, as well as the eukaryotic exporter P-glycoprotein, described previously (38). Additional pairwise alignments of the eukaryotic exporters STE6 and the cystic fibrosis transmembrane regulator suggest that they also fall in group A (data not shown). The proteins MsbA and SurB, whose functions are not yet known, also fall within group A, although very near to the base of that branch. From this, it is difficult to predict an export function yet for MsbA or SurB. Within group A are ABC exporters which do and do not contain a tandem duplication of its functional domains. None of the prokaryotic proteins in group A contain this tandem duplication. In contrast, eukaryotic ABC exporters have been found both with and without the tandem duplication (80). This suggests that the tandem duplication did not arise until after the split between prokaryotes and eukaryotes.

The bacterial ABC exporters in group A can be further subdivided. One node within group A contains proteins identified from gram-positive systems (ComA, PedD, CylB, NisT, and SpaB). They appear to more related to each other than to any of the other ABC exporters. The RTX toxin exporters also make up a distinct subfamily in group A, but, interestingly, the protease exporters PrtD and AprD do not fall within this grouping, even though they are quite similar to the RTX exporters.

ABC transporters from group B can facilitate import or export. Examples include the polysaccharide exporters, the putative heme exporters, the microcin B17 exporter, and the daunorubicin exporter, all shown in Fig. 5, as well as the other periplasmic permease importers (38, 83). Interestingly, no eukaryotic ABC transporters have yet been identified that would fall into group B. Several proteins described in this review (Orf1, NodI, and FtsE) fall in group B and do not have any known transport function.

# **INSIGHTS FROM BACTERIAL ABC EXPORTERS**

Several excellent reviews of the ABC transporters include detailed discussion of the bacterial permeases and eukaryotic transporters (3, 8, 80). These reviews also give careful consideration to fundamental aspects of the ABC transporters, including transporter structure, energy coupling during translocation, transporter regulation, and the similarities between channels and transporters. The reader is referred to these for further information. Our goal has not been to repeat these earlier works but to extend the coverage to the third subfamily of ABC transporters, the bacterial exporters. Perhaps we can gain some new insight into the ABC transporters by considering here some of the fundamental issues from the point of view of the bacterial exporters.

# How Do ABC Transporters Recognize Such a Wide Range of Substrates?

An intriguing fact about the ABC transporters is that they can transport such a wide range of structurally unrelated products. The periplasmic permeases import oligopeptides, amino acids, sugars, phosphate, metal ions, and vitamins. The eukaryotic exporters facilitate the secretion of lipophilic drugs, peptides, and pigments (80). The bacterial exporters have an equally wide range of specificity. Transported products for bacterial ABC exporters include protein toxins as large as 216 kDa, small peptide antibiotics, polysaccharides, antibiotics, and possibly even heme molecules.

Results from the bacterial systems suggest that there is specificity involved in these transport systems but that the specificity is rather relaxed, allowing for the recognition of related ABC-protein substrates by various ABC transporters. Mutagenesis of several ABC-protein substrate secretion signals shows that these signals are able to tolerate many mutations and still be secreted. Mutagenesis of the ABC transporters has identified residues in the MSD that affect transport, but, surprisingly, a mutant with deletion of the MSD still retains some function. We suggest that efficient secretion is mediated through the MSD, which retains little sequence conservation among the ABC exporters, but that a

	3 11	16 30	31 45	46 60	61 75	76 90	91 105
	1 1.	, 10	31 33	10 00	<b>75</b>		
1 HIYB	INLSIKQGEVIGIV	RSGSGKSTLTKLIQR	FYIPENGQVLIDG	HDLALADPNWLRRQV	GAADOUATTUKSTI	DNISLANPGM	SVERVIIAAALAGAH
2 HlyB-	Pv INLNIKQGEIIGIV	RSGSGKSTLTKLIOR	FYIPENGQVLIDG	HDLALADPNWLRRQV	GVVLQDNVLLNRSII	DNIALADPGM	PVEKVIHAAKLAGAH
3 T.k+B-	Aa INLDISOGEVIGIV	RSGSGKSTLTKLTOR	FYIPEOCOVLIDG	HDLALADPNWLRROV	GVVLODNVLLNRSIR	ENIALTNPGM	PMEKVIAAAKLAGAH
4 Th+D		DECECKERT THE LOP	EVID-ENCOVI IDC	UDI AT ADDNINT DOOT	CUALODNALINDELD	ENTAL-SDPCM	DMEDUTYAAKT.ACAH
4 LKCB-	PH VNLEIRQGEVIGIV	- RSGSGRSTLIKLLQR	FITEENGOVETEG	ADDADADE MADRINGT	Georgene musicality	ENTAL SUPERA	
5 AppB	VNLSIQQGEVIGIV	RSGSGKSTLTKLIQV	FYIPENGQVLIDG	HDLALADPNWLRRQV	GVVLQDNVLLGRSIR	DNIALADPGM	PMEKIVHAAKLAGAH
6 CvaB	VSLRTAPGEVVGVV	RSGSGKSTLTRLTOR	MEVADRGRVLIDG	HDIGIVDSASLEROL	GVVLOESTLENRSVR	DNIALTRPGA	SMHEVVAAARLAGAH
	VODENCE		12 VII DIGICULIDO	TO I CONTRACT DIST	ATTEODACIINETD	ENIDICDETA	TOARIARAAAAAAA
/ ChVA	VSFTAKAGETVALV	PIGAGKITLINLLQR	VIDPDSGQILIDG	TDISTVIKNSLENSI	ATVFQDAGLLNRSTR	ENIRLGREIA	IDAEVVEAAAAAAA
8 NdvA	VSFKAKAGQTIAIV	F PTGAGKTTLVNLLQR	VHEPKHGQILIDG	VDIATVTRKSLRRSI	ATVFQDAGLMNRSIG	ENIRLGREDA	SLDEVMAAAEAAAAS
9 Hot A	TTLTTERCKTTALV	ASGAGKTTLADLTPR	FYDPTEGOTLVDG	LDVOYFE INSLERKM	AVVSODTFIFNTSIR	DNIAYGTSGA	SEAEIREVARLANAL
10 Maha	THE KIDS OF THE V	Decercreated	EVDI DECETIMOC	UDI DEVITI ACI DNOU	AT UCONBULL ENDING	NULTRY ADTROY	COPOTEENADMAVAM
IU MSDA	INLKIPAGKIVALV	RSGSGRSTIASLIIR	FIDIDEGEILADG	HULKEIIIASLRIVQV	ALVSQNVHLENDIVA	NNIAIARIEQI	SREQIEEAANATAT
11 CvaB	LSLSVAPGESVAIT	G ASGAGKTTLMKVLCG	LFEPDSGRVLING	IDIRQIGINNYHRMI	ACVMODDRLFSGSIR	ENICGFAEEM	DEEWMVECARASHIH
12 AprD	LTLAIPAGSVVGVI	PSGSGKSSLARVVLG	IWPTLHGSVRLDG	AEIROYERETLGPRI	GYLPODIELFAGTVA	ENIARFGEV	QADKVVEAARLAGVH
12 D-+D	TUPPET ON CRUTT VIT	ACCCRECT ADT TWO	ACCR TOCKUPI DC	ADINOVANTECOTT	CVT DODUOT FKCST A	ENTADFCDA	DEEKINAAAKT.ACVH
15 PICD	THE SUQAGE ILLY ILL	ASGSGRSSLARULIVG	AQ3FIQGKVKLDG	ADDIQ VDIQ11 GE 11	GILLE QD VQLLE NOSLLA		
14 ComA	INLTVPQGSKVAFV	ISGSGKTTLAKMMVN	FYDPSQGEISLGG	VNLNQIDKKALRQYI	NYLPQQPYVFNGTIL	ENLLLGAKEGT	TOEDILRAVELAEIR
15 LcnC	IELSIKENEKLTIV	MSGSGKSTLVKLLVN	FFQPTSGTITLGG	IDLQQFDKHQLRRLI	NYLPQQPYIFTGSIL	DNLLLGANENA	SQEEILKAVELAEIR
16 PedD	VSLTIPHHOKITIV	MSGSGKTTLAKLLVG	FFEPOEOHGETOTNH	HNTSDISRTILROYI	NYVPOEPFIFSGSVL	ENLLLGSRPGV	TOOMIDOACSFAEIK
10 1000							
17 CylB	ISFDIRKGDKVAIV	RSGSGKSTLLKLLAG	LLQPSNGEILYEG	YPLSNNSNNRRNIFY	VNQNAHIFNETIEKN	ISLEFKPNSSIN	EKKRLKGSMSKSKMD
18 SurB	ISLOVNAGEHIAILA	RTGCGKSTLLQQLTR	AWDPQQGEILLND	SPIASLNEAALROTI	SVVPQRVHLFSATLR	DNLLLASP	GSSDEALSEILRRVG
19 NieT	INLSFEKGELTATV	KNGSGKSTLVKTISG	LYOPTMGTTOYDK	MRSSIMPEEFYOKNT	SVIFODEVKYELTIR	ENIGLISTISSOWEDE	KIIKVLDNLGLDFLK
20 0-20	TNEET UKCEDUA TW	DICCCRCTETATIOC	IVEN OCCUTIINC	TNITKELDMOCYMMOT	ANT FOREMEVEMENT K	ENTCECOIDET HOTEN	KNUEVI DIVRADELK
20 SpaB	INVSLARGERVALV	PNGSGKSTF IKILLIG	PIEA	INIKELDADSIMAQI	AADPODIMIEMIEK	ENIGEGUIDKINGIN	NAME VID I VRADE DR
21 BexA	INFELQKGEKIGILA	RNGAGKSTLIRLMSG	VEPPTSGTIERSM	SISWPLAFSGAFQGS	LTGM	DNLRFICRLYDVDPD	YVTRFTKEFS
22 CtrD	INFSLOKGEKVGTU	PNCACKSTT.VPLTSC	VEDDTSCETKDTM	STSWPLAFSCAFOCS	T.TCM	DNLRFTCRTYNVDTD	YVKAFTEEFS
			VELL ISOSTIUM	VICHDUCI DOCTOCO	LICH	ENTWERIADI VAKODE	INEDIDEREP C
23 Kpsr-	KI LNIIPPKGINIALIO	QNGAGKSTLLRIIGG	IDRPDSGNIITER	KISWPVGLAGGFQGS	DIGR	ENVREVARLIARRDE	LNERVDF VEEFS
24 KpsT-	K5 LNIEIPSGKSVAFIC	RNGAGKSTLLRMIGG	IDRPDSGKIITNK	TISWPVGLAGGFQGS	LTGR	ENVKFVARLYAKQEE	LKEKIEFVEEFA
25 DrrA	LDI MURACI MUCTI	DICACKSTTIDMLAT	LLPP-DCCTARVEC	UDWTGFDDTWDDDTC	VTCOVASVDECT.TCT	ENLUMMORTOCYSWA	PAR
			MILLE DOGIARVIG	IDVISEDIVIARIS	VIGQIAD VDEGEIGI		
26 Nod1	LSFTIAAGECFGLLO	F PNGAGKSTITRMILG	MTSPSVGK1TVLG	AQEPGQVRLARAKIG	IVSQFDNLDLEFTVR	ENLLVIGLIFRMSTR	EIETVIPSLLEFA
27 CycV	LDFEAVSGEAVAVV	RNGSGKTSLLRLIAG	LLIPAGGTIA	LDGGDAELTLPEQCH	YLGHRDALKPALSVA	ENLSFWADFLGGERL	DAHESLATV
28 HelA	VSFSLAAGHALVLR	PNGIGKTTLLRTLAG	LOPPLAGRVS	MPPEGIA	YAAHADGLKATLSVR	ENLOFWAAIHATDTV	ETALARM
Lo nem	101 011101111101111						
29 McbF	LSLKIEOGELIGLIG	ENPAGKTTLFNLIRG	GVSNYEGTLKENE	SGGELVSLPOVINLS	GILRNEEVLDLICCF	NKLTKKOAW	TDVNHKWNDNFF1
	DODIGEOGODEODD		CION IDOIDIGUE				
30 CONSE	NSUS INLSIPAGEV?AIV	RSGSGKSTLTKLIOG	FYIPDSGQILIDG	HDLALVD?NSLRRQI	GVV?QDNVLFN?SIR	ENIALA?EGA	SMEKVV?AAELAGAH
30 CONSE	NSUS INLSIPAGEV?AIV	RSGSGKSTLTKLIQG	FYIPDSGQILIDG	HDLALVD?NSLRRQI	GVV?QDNVLFN?SIR	ENIALA?EGA	SMEKVV?AAELAGAH
30 CONSE	NSUS INLSIPAGEV?AIV	<u>; RSGSGKST</u> LTKLIQG A site	FYIPDSGQILIDG	HDLALVD?NSLRRQI	GVV?QDNVLFN?SIR	ENIALA?EGA	Smekvv?aaelagah
30 CONSE	NSUS INLSIPAGEV?AIV	<u>RSGSGKST</u> LTKLIQG A site	FYIPDSGQILIDG	HDLALVD?NSLRRQI	GVV?QDNVLFN?SIR	ENIALA?EGA	Smekvv?aaelagah
30 CONSE	NSUS INLSIPAGEV?AIV	RSGSGKSTLTKLIQG A site 121 135	FYIPDSGQILIDG	HDLALVD?NSLRRQI	GVV?QDNVLFN?SIR	ENIALA?EGA 181 195	SMEKVV?AAELAGAH
30 CONSE	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG	RSGSGKSTLTKLIQG A site 121 135 EOGAGLSGGORORIA	FYIPDSGQILIDG	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488-	SMEKVV?AAELAGAH -667) 180
30 CONSE	106 120 DFISELREGYNTIV	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA FOCACLSCCOPOPIA	FYIPDSGQILIDG	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488-	SMEKVV?AAELAGAH -667) 180 -667) 180
30 CONSE 1 HlyB 2 HlyB-H	106 120 DFISELREGYNTIVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488-	SMEKVV?AAELAGAH -667) 180 -667) 180
30 CONSE 1 HlyB 2 HlyB-H 3 LktB-A	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Aa DFISELREGYNTVVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPRILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESENI FDEATSALDYESENI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QNRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- LIIAHRLSTV (488-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180
30 CONSE 1 HlyB 2 HlyB-H 3 LktB-H 4 LktB-H	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Aa DFISELREGYNTVVG Ph DFISELREGYTTIVG	RSGSCKSTLTKLIQG A site 121 135 EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA	136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPRILI IARALVNNPKILI	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESENI FDEATSALDYESENI FDEATSALDYESENI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QRTV	ENIALA?EGA 181 195 111AHRLSTV (488- 111AHRLSTV (488- L11AHRLSTV (489- L11AHRLSTV (489-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180
30 CONSE 1 HlyB 2 HlyB-H 3 LktB-H 4 LktB-H	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Aa DFISELREGYNTVG Ph DFISELREGYTTIVG	RSCSCKSTLTKLIQG A site 121 135 EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESENI FDEATSALDYESEHI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QNRTV IMQNMQKIC-QGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180
30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-F	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Aa DFISELREGYNTIVG Ph DFISELREGYTTIVG Ph DFISELREGYTTIVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA	136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QNRTV IMQNMQKIC-QGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- LIIAHRLSTV (488- ILIAHRLSTV (489- ILIAHRLSTV (489-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180
30 CONSE 1 HlyB 2 HlyB-H 3 LktB-H 4 LktB-H 5 AppB	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA	136     150       136     150       1ARALVNNPKILI     1ARALVNNPKILI       1ARALVNNPKILI     1ARALVNNPKILI       1ARALVNNPKILI     1ARALVNNPKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489- ILIAHRLSTV (489- IIIAHRLSTV (488-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180
<pre>30 CONSE 1 HlyB 2 HlyB-H 3 LktB-H 4 LktB-H 5 AppB 6 CyaB</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG EFICQLPEGYDTMLG	RSCSCKSTLTKLIQG A site 121 135 EQCAGLSCGORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIHRPRVLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (488- ILIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSAV (493-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -672) 180
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-F</li> <li>AppB</li> <li>CyaB</li> <li>ChyA</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG AA DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFICQLPEGYDTMLG DFIDSRINGYLTQVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQPNRIA ENGVGLSGGQPQRIG ERGNRLSGGERQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIHRPRVLI IARALIKNAPILV	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- LIIAHRLSTV (488- ILIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (354-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -672) 180 -533) 180
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIDSRINGYLTQVG	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA ENGVGLSGCGQQRIA ERGNRLSGCERQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIHPRVLI IARALIKNAPILV IARAILKNAPILV	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QNRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (354- FIIAHRLSTV (354- FIIAHRLSTV (354-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -672) 180 -533) 180 -533) 180
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTVG DFIEDRLNGYDTVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGORQRIA EQGAGLSGGORQRIA EQGAGLSGGORQRIA EQGAGLSGGORQRIA ENGVGLSGGORQRIG ERGNRLSGGERQRIA ERGNRLSGGERQRIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIHRPRVLI IARAILKNAPILV IARAILKNAPILV	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYETEAR LDEATSALDVETEAR	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QRTV IMQNMQKIC-QGRTV IQRNMRDIC-KGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- FIIAHRLSTV (385-	SMEKVV?AAELAGAH -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -533) 180 -564) 180
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG ADFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFICQLPEGYDTMLG DFIDSRINGYLTVG DFIEDRLNGYDTVVG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGCORORIA EQGAGLSGCORORIA EQGAGLSGCORORIA EQGAGLSGCORORIA EQGAGLSGCORORIA ENGNUSGCGCRORIA ERGNRLSGCERORIA	136       150         136       150         IARALVNNPKILI       IARALVNNPKILI         IARALVNNPKILI       IARALVNNPKILI         IARALVNNPKILI       IARALVNNPKILI         IARALVNNPKILI       IARALVNNPKILI         IARALVNNPKILI       IARALIKNAPILI         IARALIKNAPILI       IARALIKNAPILI         IARALIKNAPILV       IARAILKNAPILV	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-GGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (355- FIIAHRLSTV (355- FIIAH	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -533) 180 -564) 180 -564) 180
<pre>30 CONSE 1 HlyB 2 HlyB-F 3 LktB-F 4 LktB-F 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIDERLNGYDTVG QFIEEMPEGFDTKLG	RSCSCKSTLTKLIQG A site 121 135 EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA EQCAGLSCCORORIA ENGVGLSCCORORIA ERGNRLSCCERORIA DRGVRLSCCORORIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALINNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDPEILI	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QNRTV IMHNMHKIC-QRTV IMRNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (385- IIAHRLSTI (383- IAIAHRLSTI (383-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -672) 180 -572) 180 -533) 180 -564) 180
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIEDRLNGYDTVVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA ENCVGLSGGQRQRIA ERGNRLSGGERQRIA DRGVRLSGGQRQRIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDPEILI IARALLRDPEILI	HDLALVD?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSVSERL LDEATSRLDTESERA	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMQNMQKIC-GGRTV IQRNMRDIC-KGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (362-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -572) 180 -533) 180 -564) 180 -562) 180 -542) 181
<ul> <li>30 CONSE</li> <li>1 HlyB</li> <li>2 HlyB-I</li> <li>3 LktB-I</li> <li>4 LktB-I</li> <li>5 AppB</li> <li>6 CyaB</li> <li>7 ChvA</li> <li>8 NdvA</li> <li>9 HetA</li> <li>10 MsbA</li> <li>11 CvaB</li> </ul>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIDSRINGYLTQVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG DFINKMDNGETLG	RSCSCKSTLTKLIQC A site 121 135 EQCAGLSCGQRQRIA EQCAGLSCGQRQRIA EQCAGLSCGQRQRIA EQCAGLSCGQRQRIA EQCAGLSCGQRQRIA ENCYCLSCGQRQRIA ERGNRLSCGERQRIA ENCYLLSCGQRQRIA ELGEGLSCGCQRQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDPEILI IARALLRDPEILI IARALRDPEILF IARALRDPEILF	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSSERL LDEATSALDSESEHF	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQDAIDALR-KNRTV IQDAIDALQ-KNRTS IQDAIDALQ-KNRTS	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (488- IIIAHRLSTV (385- IIAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRLSTI (512-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -672) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AppP</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFICQLPEGYDTMLG DFIDSRINGYLTQVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG EFU BLDGCYNTUG	RSCSCKSTLTKLIQG A site 121 135 EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA ENGVGLSCGORORIA ERGNRLSCGERORIA ENGVLLSCGORORIA ELGEGLSCGORORIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSSEEHA MDEATSALDSESEHA	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMQKIC-QGRTV VKAAVDALR-KORTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR VIVAIKNMNITR	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRETTI (512- VIIAHRETTI (512- VIIAHR	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 190
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG ELVLRLPQGYDTVLG	RSCSGKSTLTKLIQG A site 121 135 EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA EQCAGLSCGORORIA ENGRUSCGCERORIA ERGNRLSCGERORIA ENGVLLSCGORORIA ELGEGLSCGORORIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPRILV IARALLRDPEILI IARALLRDPEILI IARALLRDSPILF LARALYGAPTLVV	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSVSERL LDEATSALDSSEHF LDEPNSNLDDSGEQA	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-KGRTV IQRNMRDIC-GGRTV IQRNMRDIC-DGRTV VKDAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRETTL (512- LLITHRAGVL (352-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -573) 180 -573) 180 -564) 180 -562) 180 -562) 180 -542) 181 -689) 178 -531) 180
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIG DFIS	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA ENGVGLSGCGCRQRIA ERGNRLSGGERQRVA DRGVRLSGCQRQRIA ELGEGLSGCQRQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDSPILI IARALLRDSPILI IARALYGAPTLVV	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSVSERL LDEATSALDSEEHA MDEATSALDSEEHA	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIAHRLSTI (383- IVIAHRLSTI (362- VIIAHRETTL (512- LLITHRAGVL (352-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -562) 180 -542) 181 -689) 178 -531) 180
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG EFICQLPEGYDTMLG DFIDSRINGYLTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG ELVLRLPQGYDTVIG ELILSLPNGYDTELG	RSCSCKSTLTKLIQG A site 121 135 EQCAGLSGCORQRIA EQCAGLSGCORQRIA EQCAGLSGCORQRIA EQCAGLSGCORQRIA EQCAGLSGCORQRIA ENGVGLSGCORQRIA ERGNRLSGCERQRIA DRGVRLSGCORQRIA ELGEGLSGCORQRIA ELGEGLSGCORQRIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPRILV IARALLRDPEILI IARALLRDPEILI IARALLRDPEILF IARALYGPFCLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSVSERL LDEATSALDSSEHF LDEPNSNLDDSGEQA LDEPNASLDSEGDQA	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (383- IIIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRETTL (512- LLITHRAGVL (352- VLITHRPALT (350-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFIEDRLNGYDTVG DFIEDRLNGYDTVG DFIEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG ELVLRLPQGYDTVG ELILSLPNGYDTELG EDIERMPLNYQTELI	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA ERGNRLSGGERQRIA ERGNRLSGGERQRIA ENGVLLSGGQRQRIA ELGEGLSGGQRQRIA DGGGGLSGGQRQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALINNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDPEILI IARALLRDSPILF IARALYKKPGILF LARALYGAPTLVV LARAMYGDPCLLI LARALTDAPVLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVERL LDEATSALDSVERL LDEPNSNLDDSGEQA LDEPNSSLDSEGDQA LDEATSSLD ILTERR	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQLIC-KGRTV IQRNMRDIC-DGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTL	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (384- FIIAHRLSTV (385- IAIAHRLSTI (382- UVIAHRLSTI (362- VLITHRASTI (512- UIIAHRETTL (512- IIITHRAGVL (350- IFIAHRLTI (350- IFIAHRLTI (510-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -672) 180 -533) 180 -564) 180 -564) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179
1 HlyB 2 HlyB-H 3 LktB-A 4 LktB-H 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 Lorc	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG ELVLRLPQGYDTVLG ELILSLPNGYDTELG EDIERMFLNYQTELG EDIERMFLNYQTELG	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQGAGLSGCQRQRIA     ENGVLLSGCGRQRIA     ENGVLLSGCQRQRIA     ELGEGLSGCQRQRIA     DGGGGLSGCQRQRIA     DGGGGLSGCQRQRIA     SDASLSCCOKORIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPEILI IARALLRDPEILI IARALLRDPEILI IARALYRPGLVV LARALYGPCLLI LARALTDAPVLI LARALTDAPVLI LARALTDAPVLI LARALTDAPVLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSEEHF LDEPNSNLDDSGEQA LDEPNASLDSEGDQA LDEATSSILDMITEKK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QRTV IMRNMHQIC-GGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVMNLIALDKTI	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRETTL (512- LLITHRAGVL (352- VLITHRPALT (350- IFIAHRLTA (503- IFIAHRLSA (503- IFI	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179
1 HlyB 2 HlyB-I 3 LktB-J 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LenC	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG Ph DFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFIDSRINGYLTVG DFIEDRLNGYDTVG OFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG ELVLRLPGYDTLG ELILSLPNGYDTELG EDIERMPLNGYDTELG EDIECMPLNGYTELS NDIECMGLGYOTELS	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRVA     DRGVRLSGGCRQRVA     DRGVRLSGGQRQRIA     ELGEGLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALHRPRVLI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPELLI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSERL LDEATSALDSERL LDEPNSNLDDSGEQA LDEPNSSLDJLTEKR LDEATSNLDMITEKK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMHNMHKIC-QRTV IMHNMHKIC-QRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-KGRTV IQRNMRDIC-KGRTV IQRNMRDIC-GRTV IQRNMRDIC-KGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTL ILKNLLPLDKTL	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489- ILIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTI (383- LVIAHRLSTI (363- IVIAHRLSTI (362- VLITHRPALT (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLST) (501- IFIAHRLSTA (501- IFIAH	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -573) 180 -564) 180 -564) 180 -562) 180 -562) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -679) 179 -679) 100
1 HlyB 2 HlyB-I 3 LktB-J 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG         PV       DFISELREGYNTIVG         PM       DFISELREGYNTIVG         EFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         GFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTVG         QFIEEMPEGFDTKLG         DFINKMONGLDTVG         QFIEEMPEGFDTKLG         DVIMNMEMSYETLIG         ELVIRLPQGYDTLG         ELILSLPNGYDTELG         EDIERMPINYQTELIA         ADIEQMQIGYQTELS         TDIENLPQGYHTRLS	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA EQCAGLSGCQCQRIA ENGVRLSGCGCQCQRIA ENGVLLSGCQCQRIA ELGEGLSGCQCQRIA ELGEGLSGCQCQRIA DGGGGLSGCQCQRIA SDASGLSGCQCQRIA ESGFNLSGCQCQRIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDSPILI IARALYGAPTLVV LARAMYGDPCLLI LARALLTDAPVLI IARALLSPAKILI IARALLSPAKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSESEHF LDEPNSNLDDSGEQA LDEPNASLDSESEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF LDEPNASLDSEEHF	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KORTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS LLAAIQALKARGCTV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (362- VLIAHRLSTI (362- VLIAHRLSTI (352- VLITHRAGVL (352- VLITHRAGVL (352- IFIAHRLTA (503- IFIAHRLTA (505- IFIAHRLTA (505-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -529) 180 -691) 179 -691) 179 -686) 182
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>CyaB</li> <li>ThvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>ComA</li> <li>LcnC</li> <li>FedD</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFIDSRINGYLTQVG DFIDSRINGYLTQVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMONGLDTVIG ELVLRLPQGYDTVG ELILSLPNGYDTELG EDIERMPLNYQTELT ADIEQMQLGYQTELS TDIENLPQGYHTRLS	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCORORIA EQCAGLSGCORORIA EQCAGLSGCORORIA EQCAGLSGCORORIA EQCAGLSGCORORIA EQGAGLSGCORORIA ENGVGLSGCORORIA ENGVRLSGCERORIA ENGVLSGCORORIA ELGEGLSGCORORIA ELGEGLSGCORORIA SDASSLSGCORORIA ESGFNLSGCORORIA	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPILV IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPELIF LARALYGAPTLVV LARAMYGDPCLLI LARALLSPAQCFI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSEEHF LDEPNSNLDDSGEQA LDEPNSSLDSEGDQA LDEATSSLDIITERK FDESTSNLDMITEKK FDESTSNLDMITEKK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-GGRTV IQNNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI ILKNLLPLDKTI IVSKLLFMKDKTI	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTI (383- LVIAHRLSTI (383- LVIAHRLSTI (350- IIIAHRETTL (512- LLITHRAGVL (350- IFIAHRLTI (503- IFIAHRLSTA (501- IFIAHRLSTA (501- IFVAHRLNIA (505-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -553) 180 -564) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>PrtD</li> <li>ComA</li> <li>LcnC</li> <li>FedD</li> <li>Cv1B</li> </ol>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFIEDRLNGYDTVG DFIEDRLNGYDTVG DFIEDRLNGYDTVG ELVLRLPQGYDTVG ELILSLPNGYDTELG EDIERMPLNYQTELT ADIEQMQLGYQTELS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA ERGNRLSGGERQRIA ERGNRLSGGGRQRIA ERGVRLSGGQRQRIA ELGEGLSGGQRQRIA DGGGGLSGGQRQRIA SDASSLSGGQRQRIA SDASSLSGGQRQRIA ESGFNLSGGQRQRIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPELII IARALLSPACLII IARALSPACLII IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSERA MDEATSALDSERA LDEPNSNLDDSGEQA LDEPNSSLDSEGDQA LDEATSSLDILTERR LDEATSNLDMITEKK FDESTSNLDTITEHK LDEPTSAMDNISEFE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI VFSNLLDEKRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (383- LVIAHRLSTI (365- IIIAHRLSTI (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLSVA (501- IFIAHRLSVA (501- IFVAHRLNIA (505- ITVAHRLSTV (491-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -573) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180
1 HlyB 2 HlyB-I 3 LktB-Z 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISELREGYNTIG DFISE	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQGAGLSGCQRQRIA     ENGNRLSGGERQRVA     DRGVRLSGGERQRVA     DRGVRLSGGCRQRIA     ELGEGLSGCQRQRIA     ELGEGLSGCQRQRIA     DGGGLSGCQRQRIA     SDASLSGCQRQRIA     SDASLSGCQRQRIA     SDASLSGCQRQRIA     SDASSLSGCQRQRIA     ESGFNLSGCQRQRIA     ENGSNFSGCQRQKIA     ENGSNFSGCQRQKIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPILV IARALLRDPEILI IARALYRPGILI IARALYGPTLVV LARAMYGDPCLLI IARALLSPAQCFI IARALSPAQCFI LARAFYSNVNTLL IARALHDAPVLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSVSERL LDEPNSNLDDSGEQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSLDILTEKK FDESTSNLDTITEHK LDEPTSAMDNISEFE LDEPTSAMDNISEFE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKDAIDALR-KORTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI VFSNLLDEKRTV	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VIIAHRLSTI (352- VLITHRAGVL (352- VLITHRAGVL (352- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTV (491- IFVAHRLNIA (505- ITVAHRLSTV (491- IVVHRLEGI (359- ISTVAHRLSTV (491- INVTHRLEGI (359- ISTVAHRLSTV (491- INTHRLEGI (359- ISTVAHRLSTV (491- INTHRLEGI (359- ISTVAHRLSTV (491- INTHRLEGI (359- ISTVAHRLSTV (491- INTHRLEGI (359- ISTVAHRLSTV (350- ISTVAHRLSTV (350- ISTV	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>ComA</li> <li>PrtD</li> <li>ComA</li> <li>LcnC</li> <li>PedD</li> <li>CylB</li> <li>SurB</li> <li>SurB</li> <li>Ndvam</li> </ol>	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFIDSRINGYLTQVG DFIDSRINGYLTQVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG ELILSLPNGYDTELG EDIERMPLNYQTELT ADIEQNGLGYQTELS TDIENLPQGYHTRLS EVILGIPQYEKTIVS LEKLLEDAGINSMLG	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRVA     DRGVRLSGGCRQRVA     DRGVRLSGGCRQRIA     ELGEGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     ESGFNLSGGQRQKIA     ESGFNLSGQQQX	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPILV IARALLRDPEILI IARALLRDPEILI IARALLRDPELILI IARALLSPACLILI IARALLSPACCFI IARALLSPACLVY	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSEEHF LDEPNSNLDDSGEQA LDEPNSSLDJLTEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHCC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV IQRNMRDIC-DGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IEDVEUMCENTY IEDVEUMCENTY	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTI (363- IVIAHRLSTI (362- VLIAHRLSTI (362- VLITHRPALT (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (505- ITVAHRISTV (491- IMTHHRLRGU (359- IMTHHRLRGU (359- IMTHR	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -573) 180 -553) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -555) 195
1 HlyB 2 HlyB-I 3 LktB-J 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB 19 NisT	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTVG DF	RSGSGKSTLTKLIQG         A site         121       135         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIG         ERGNRLSGGERQRIA         ERGNRLSGGERQRIA         ENGVLLSGGQRQRIA         ELGEGLSGGQRQRIA         SDGAGLSGGQRQRIA         DGGGGLSGGQRQRIA         SDASLSGQQRQRIA         ESGFNLSGGQRQKIA         EGGRQLSGGQRQKIA         EGGRQLSGGQRQKIA         EGGRQLSGGQRQKIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARALLRDPEILI IARALYRAPGILF IARALYGAPTLVV LARAMYGDPCLLI LARALSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSESEHF LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPTSSILDILTERR LDEPTSAMDNISEFE LDEPTSGLDATTESQ LDEPTSGLDATTESQ LDEPSAALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IVDNLIALDKTL IVSKLLFMKDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFVAL-SENNIS	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (362- VLIAHRLSTI (362- VLIAHRLSTI (352- VLITHRPALT (505- IFIAHRLSVA (501- IFIAHRLSVA (501- IFIAHRLSVA (501- IFIAHRLSTV (491- IMVTHRLRGL (359- IFIAHRLSTV (491- IMVTHRLGL (359- IFISSINAA (371-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -572) 180 -533) 180 -564) 180 -522) 180 -542) 181 -6689 178 -531) 180 -529) 180 -529) 180 -691 179 -691 179 -691 179 -691 179 -691 179 -691 180 -536) 182 -570 180 -536) 178 -535) 185
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>CyaB</li> <li>ThvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PretD</li> <li>PrtD</li> <li>ComA</li> <li>PrtD</li> <li>ComA</li> <li>LcnC</li> <li>PedD</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DF ISELREGYNTIVG         PV       DF ISELREGYNTIVG         PV       DF ISELREGYNTIVG         PV       DF ISELREGYNTIVG         EF ISELREGYNTIVG       EF ISELREGYNTIVG         DF ISELREGYNTIVG       DF ISELREGYNTIVG         QF IEERPEGFDTKLG       DF INSMONGLDTVIG         QF IEEMPEGFDTKLG       DF INKMONGLDTVIG         QF IEEMPEGFDTKLG       ELIV.RLPQGYDTVG         QF IEEMPEGFDTKLG       ELIV.RLPQGYDTVG         QF IEEMPEGFDTKLG       ELILSLPNGYDTUG         ELILSLPNGYDTUG       ELILSLPNGYDTUG         ELILSLPNGYDTELG       ED IERMPLNYQTELT         AD IEQMQLGYQTELS       TD IENLPQGYHTRLS         EVLLGIPQYEKTIVS       LEKLLEDAGINSWLG         TNNQYVLDTGLGRWF       SHSSYQFDTQLGRWF	RSGSGKSTLTKLIQG           A site           121         135           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           EQGAGLSGCQORQRIA           ENGVILSGCQCQRQRIA           DRGVRLSGGQRQRIA           ENGVILSGCQRQRIA           ENGVILSGCQRQRIA           EGGGGLSGGQRQRIA           SDASSLSGCQRQRIA           ESGFNLSGCQRQRIA           ESGFNLSGCQRQRIA           ESGFNLSGCQRQRIA           ENGSNFSGCQRQKIA           EGGRQLSGCGQNQKIA           DEGRQLSGCQWQKIA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPVLI IARALIRDPEILI IARALIRDPEILI IARALIRDPEILI IARALIRDPEILI IARALIRDPEILI IARALIRDPEILI IARALISPACILI IARALISPACILI IARALISPACILI IARALISPACILI IARALISPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR IDEATSALDVETEAR IDEATSALDSEEHF LDEPNSNLDDSGEQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPTSALDMITEKK FDESTSNLDTITEHK LDEPTSANDNISEFE LDEPTSALDPVAEKE LDEPSSALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMQKIC-QGRTV VKDAIDALR-KORTT VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFFSL-SKDKIG	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTI (383- LVIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRLSTI (352- VIITHRAGVL (352- VLITHRPALT (350- IFIAHRLSTV (501- IFVAHRLSTV (501- IFVAHRL	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -672) 180 -533) 180 -542) 180 -542) 181 -531) 180 -529) 180 -531) 180 -529) 180 -686) 182 -670) 180 -536) 178 -536) 178 -536) 178 -536) 178 -536) 185 -542) 185
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>ComA</li> <li>Dist</li> <li>PedD</li> <li>CylB</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG         PV       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFIEDRLNGYDTVG         DFIEDRLNGYDTVG         DFIEDRLNGYDTVG         ELILSLPNGYDTVG         ELILSLPNGYDTVG         ELILSLPNGYDTELG         DIENRPLNYQTELT         ADIEQMQLGYQTELS         TDIENLPQGYHTRLS         EVLLGIPQYEKTIVS         LEKLLEDAGINSWG         TNNQYVLDTQLGNWF         SHSSYQFDTQLGLWF	RSGSGKSTLTKLIQG         A site         121       135         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         EQGAGLSGGQRQRIA         ERGNRLSGGERQRIA         ERGVLLSGGQRQRIA         ELGEGLSGGQCQRIA         ELGEGLSGGQRQRIA         SDASSLSGGQRQRIA         SDASSLSGGQRQRIA         ENGSNFSGGQRQRIA         ENGSNFSGGQRQKIA         EGGRQLSGGQWQKIA         QEGHQLSGGQWQKIA         DGGGQLSGGQWQKIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALHRPRVLI IARALHRPRVLI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPELII IARALLROPELII IARALLROPELII IARALLROPELII IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI	HDLALVD ?NSLRRQI 151, 165 FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDVETEAR MDEATSALDSVSERL LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSSLDJLTERR LDEATSSLDILTERR LDEATSSLDILTERK FDESTSNLDTITEHK LDEPTSAMDNISEFE LDEPTSGLDATTESQ LDEPSAALDPVAEKE LDEPSSALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-GGRTV IQRNMRDIC-GGRTV IQRNMRDIC-KGRTV IQRSIEKLS-VGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT VKDAIDALR-KORTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFVAL-SENNIS TFDTFFSL-SKDKIG	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- ILIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (354- FIIAHRLSTV (354- FIIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VLITHRPALT (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLSTV (491- INVTHRLRGL (359- IFISHSLNAA (371- IFISHSLNAA (371- IFISHRLVAA (358-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -555) 185 -542) 185
<pre>30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB 19 NisT 20 SpaB 21 BexA</pre>	NSUS INLSIPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTVG DFI	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA EQCAGLSGCQRQRIA ENGVGLSGCGCRQRIA ERGNRLSGGERQRVA DRGVRLSGCGRQRIA ELGEGLSGCQRQRIA ELGEGLSGCQRQRIA SDASGLSGCQRQRIA SDASGLSGCQRQRIA ESGFNLSGCQRQRIA ESGFNLSGCQRQRIA ESGFNLSGCQRQRIA ESGFNLSGCQRQRIA ESGFNLSGCQRQRIA ESGFNLSGCQRQRIA EGGRQLSGCGPQCKIA DEGRQLSGCGPQQKIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPVLI IARALLRDPEILI IARALLRDSPILI IARALYGAPTLVV LARAMYGDPCLII LARALSPAKILI IARALSPAKILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEATSALDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPTASLDILTEKK FDESTSNLDTITEHK LDEPTSAMDNISEFE LDEPTSALDVISEFE LDEPTSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KORTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS NVAAIKM—-NITR ILAAIQAIKARGGTV IMQAIVALQKRGATV IVDNLIALDKTL IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFVAL-SENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRST	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VIIAHRLSTI (352- VLITHRAGVL (352- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (505- IFIAHRLSTA (505- IFIAHRLSTA (505- IFIAHRLSTV (491- IMVTHRLRGL (359- IFISHSLNAA (371- IFISHRLVAA (358- ILVSHSPSAM (24-1)	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -681) 179 -679) 180 -536) 182 -670) 180 -536) 178 -535) 185 -542) 185 -542) 185 -542) 185
<ol> <li>CONSE</li> <li>H HyB</li> <li>H HyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>FrtD</li> <li>ComA</li> <li>IS</li> <li>PrdD</li> <li>CyaB</li> <li>NisT</li> <li>SpaB</li> <li>ExA</li> </ol>	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFIDSRINGYLTQVG DFIDSRINGYLTQVG DFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG ELILSLPNGYDTVIG ELILSLPNGYDTVIG ELILSLPNGYDTELI ADIEQMQLGYQTELIS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS LEKLLEDAGINSWLG TNNQYVLDTQLGWWF SHSSYQFDTQLGUWF	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRVA     DRGVRLSGGCRQRVA     DRGVRLSGGCRQRIA     ELGEGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     ESGFNLSGCQRQRIA     ESGFNLSGCQRQRIA     ENGSNFSGGQRQKIA     EGGRQLSGGQRQKIA     QEGHQLSGGQWQKIA     OGGGLSGGQWQKIA     OGGQLSGGQWQKIA     OGGQLSGGQWQKIA     OGGQLSGGQWQKIA     OGGQLSGGQWQKIA     OGGQLSGGWQKIA     OGGQLSGGWQXIA     OGGULSGCWQXIA     OGGQLSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGGWQXIA     OGGULSGCWQXIA     OGGULSGCWQXIA     OGGULSGCWQXIA     OGGULSGCWQXIA     OGGULSGCWQXIA     OGGULSGULX	FYIPDSGQILIDG 136 150 1ARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALHRPRVLI IARALLRDPEILI IARALLRDPEILI IARALLRDPEILI IARALLRDPELI IARALLSPACLI IARALLSPACCI IARALLSVEFDCYL	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSEEHF LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK LDEPTEGLDATTESQ LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV IQRNMRDIC-DGRTV IQRNMRDIC-BRTV IQRNMRDIC-KGRTV IQRNMRDIC-GRTV IQRNMRDIC-KGRTV IQRNMRDIC-GRTV IQRNMRDIC-BRTV IQRNMRDIC-BRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IFDYFVAL-SENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRSI	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTI (363- IVIAHRLSTI (362- VLITHRPALT (350- IFIAHRLSTI (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLSVA (501- IFIAHRLSVA (501- IFIAHRLSVA (501- IFIAHRLSVA (503- IFIAHRLSVA (503- IFIAH	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -562) 180 -564) 180 -562) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -535) 185 -542) 180 -542) 180 -542) 181 -542) 182 -542) 185 -542) 185 -5420
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>MsbA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>TcvaB</li> <li>AprD</li> <li>PrtD</li> <li>PrtD</li> <li>PrtD</li> <li>PrtD</li> <li>FedD</li> <li>CylB</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>PrtD</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG         PV       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         QFIEENEGYNTIVG         QFIEENEGYNTVG         DFINKMDNGLDVIG         DFINKMDNGLDVIG         ELULSLPNGYDTVG         EVLLGIPQYGYTELS         TDIENLPQGYHTRLS         EVLLGIPQYEKTIVS         LEKLLEDAGINSNIG         TNNQYVLDTQLGNWF         SHSSYQFDTQLGLWF         ELGQYLYEPVKK         ELGQYLYEPVKK	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQQRIA EQCAGLSGCQQQQRIA EQCAGLSGCQQQQRIA EQCAGLSGCQQQQRIA ERGNRLSGGEQQQRIA ERGNRLSGGEQQQRIA ERGNRLSGCQQQQRIA ELGEGLSGCQQQQRIA DGGGGLSGCQQQQRIA DGGGGLSGCQQQRIA SDASSLSGCQQQRIA ESGFNLSGCQQQRIA ESGFNLSGCQQQQRIA EGGQLSGCQQQQKIA EGGQLSGCQQQQKIA DGGGQLSGCQQQXIA EGGQLSGCQQQXIA DGGQLSGCQQQXIA DGGQLSGCQQQXIA DGGQLSGCQQQXIA COCHOLSGCQQQXIA DGGQLSGCQQQXIA COCHOLSGCQQQXIA COCHOLSGCQQQXIA COCHOLSGCQQQXIA COCHOLSGCQQQXIA COCHOLSGCQQQXIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARAILKNAPILV IARALLRDSPILI IARALYRAPGILF IARALYGAPTLVV LARAMYGDPCLLI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKIVI I	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSESEHF LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPSANDNISEFE LDEPTSGLDATTESQ LDEPSAALDPVAEKE LDEPSSALDP IAEKE IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IVDNLIALDKTL IVSKLLFMKDKTI IVSKLLPMKDKTI IVSKLLPMKDKTI IVSKLLPMKSENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRSI	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IIAIAHRLSTV (385- IAIAHRLSTI (362- VLITHRPALT (350- IFIAHRLSTI (352- VLITHRPALT (350- IFIAHRLSTA (505- IFIAHRLSTA (505- IFVAHRLSTV (491- IFVAHRLSTV (491- IMVTHRLRGL (359- IFISHSLNAA (371- IFISHRLVAA (358- ILVSHSPSAM (24-1 ILVSHSPSAM (24-1	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -572) 180 -533) 180 -533) 180 -542) 181 -6689 178 -531) 180 -529) 180 -529) 180 -529) 180 -529) 180 -531) 180 -552) 181 -6691 179 -679 179 -686) 182 -670) 180 -536) 178 -555) 185 -542)
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>AppB</li> <li>CyaB</li> <li>7 ChvA</li> <li>NdvA</li> <li>9 HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PHEA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>ConA</li> <li>FrdD</li> <li>FedD</li> <li>CylB</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>KpsT-F</li> </ol>	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG OFIEDRLNGYDTVG OFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMONGLDTVIG DFINKMONGLDTVIG ELILSLPNGYDTVG ELILSLPNGYDTVG ELILSLPNGYDTUG ELILSLPNGYDTELG EDIERMFLNGYTELS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS LEKLLEDAGLNSWLG TNNQYVLDTQLGKWF SHSSYQFDTQLGLWF ELGQYLYEPVKK ELGQYLYEPVKK	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQGAGLSGCQRQRIA     ENGVGLSGCQRQRIA     ENGVLSGCGRQRIA     DRGVRLSGCGRQRIA     ELGEGLSGCQRQRIA     DGGGLSGCQRQRIA     DGGGLSGCQRQRIA     SDASSLSGCQRQRIA     SDASSLSGCQRQRIA     ESGFNLSGCQRQRIA     ESGFNLSGCQRQRIA     EGGRQLSGCQRQKIA     CEGRQLSGCQRQKIA     DGGRQLSGCQWQKIA    YSSGMKARLA    YSSGMRSRLA	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPVLI IARALIRDPEILI IARALIRDPEILI IARALIRDSPILI IARALIRDSPILI IARALYGPCLLI LARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACIVI FALSLSVEFDCYL FALSLAVEFDCYL FGLSMAFKFDYYL	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSEEHI LDEPNSNLDDSGEQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPSANLDMITEKK FDESTSNLDTITEHK LDEPTSANDNISEFE LDEPTSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEVIAVGDSRFAEK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKDAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTL ILKNLLFLDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IFDYFYAL-SENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRSI CKYELFEKR-KDRSI	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- IFIAHRLSTA (352- IFIAHRLSTA (351- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLSTV (491- IMVTHRLRGL (359- IFISHSLNAA (371- IFISHSLNAA (371- IFISHSLNAA (371- IIVSHSPSAM (24-1) ILVSHSPSAM (24-1) ILVSHSPSAM (24-1) ILVSHSPSAM (24-1)	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -543) 180 -552) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -686) 182 -670) 180 -536) 178 -555) 185 -542) 185 -542 -542) 185 -542) 185 -542) 185 -542) 185
30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB 19 NisT 20 SpaB 21 BexA 22 CtrD 23 KpsT-F 24 KpsT-F	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG       EFISELREGYNTIVG         EFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTVG         DFIEDRLNGYDTVG       DFIEDRLNGYDTVG         DFIEDRLNGYDTVG       ELISLPNGYDTVG         ELILSLPNGYDTVG       ELISLPNGYDTELG         EULLGIPQYEKTIVS       LEKLLEDAGLNSWLG         TNNQYVLDTQLGNWF       SHSSYQFDTQLGLWF         ELGGYLYEPVKR       ELGGYLYEPVKR         K1       ELGKYFDMPIKT	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRIA     ERGNRLSGGERQRIA     DRGVRLSGGQRQRIA     DRGVLSGGQRQRIA     ELGEGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     EGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     CONNFSGGQRQRIA     DGGRQLSGGQRQRIA     CONNFSGGQRQRIA     DGGRQLSGGQRQRIA     CONNFSGGQRQRIA     CONNFSGGRQRQRIA     CONNFSGGRQRVANA     CONNFSGGRQRVANAA     CONNFSGGRQRVANAAA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGGRAPRIA     CONNFSGRAPRIA      CONNFSGRAPRIA     CONNF	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALNNPKILI IARALIRPRVLI IARALIRNAPILV IARAILROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IA	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSVSERL LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK LDEPTSAMDNISEFE LDEPTSGLDATTESQ LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEPSSALDPVAEKE	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHCC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-GGRTV IQRNMRDIC-GGRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IVKDAIDALR-KDRTT IVXALFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSSNLLDEKRTV ILELLAEM-MREKTV IFDJFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CSDIFDKIR-EKSHL	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTV (385- IIAIAHRLSTI (383- LVIAHRLSTI (383- LVIAHRLSTI (362- VLITHRPALT (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLSVA (501- IFVAHRLNIA (505- ITVAHRLSTV (491- IVVHRLRGL (359- IFISHSLNAA (371- IFISHSLNAA (371- IFISHSLNAA (358- ILVSHSSFAM (24-1) ILVSHSSFAM (24-1) IMVSHSENAL (24-1) IMVSHSENAL (24-1)	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -564) 180 -564) 180 -562) 180 -562) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555] 185 -542] 185 -545 -545 -545 -545 -54
<ol> <li>CONSE</li> <li>HlyB-1</li> <li>LktB-2</li> <li>LktB-2</li> <li>LktB-3</li> <li>LktB-4</li> <li>LktB-4</li> <li>LktB-7</li> <li>LktB-7</li> <li>AppB</li> <li>CyaB</li> <li>CyaB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>MsbA</li> <li>MsbA</li> <li>MsbA</li> <li>SappB</li> <li>PrtD</li> <li>FrtD</li> <li>FrtD</li> <li>FrtD</li> <li>FrtD</li> <li>FrtD</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>KpsT-F</li> <li>KpsT-F</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG       DFISELREGYNTIVG         EY       DFISELREGYNTIVG         EFISELREGYNTIVG       DFISELREGYNTIVG         EFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       DFISELREGYNTIVG         QFIEEMPEGFDTKLG       DFINKMDNGLDTVIG         DFINKMDNGLDTVIG       DFINKMDNGLDTVIG         ELVLRLPQGYDTLG       ELVLRLPQGYDTVLG         ELILSLPNGYDTELT       ADIEQMQLGYQTELS         TDIENLPQGYHTRLS       EVILGIPQYEKTIVS         LEKLLEDAGLNSWIG       TNNQYULDTQLGNWF         SHSSYQFDTQLGLWF       ELGQYLYEPVKR         ELGQYLYEPVKR       ELGQYLYEPVKR         KS       ELGKYFDMP IKT	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQCQQRIA     EQCAGLSGCQCQQRIA     EQCAGLSGCQCQQRIA     EQCAGLSGCQCQQRIA     EQCAGLSGCQCQQRIA     EQGAGLSGCQCQQRIA     ENGVILSGCQCQQRIA     ENGVILSGCQCQQRIA     ELGEGLSGCQCQRIA     DCGCGLSGCQCQRIA     DCGCGLSGCQCQRIA     ESGFNLSGCQCQRIA     ESGFNLSGCQCQQRIA     ESGFNLSGCQCQQRIA     DCGCGLSGCQCQQRIA     DCGCGLSGCQCQRIA     ESGFNLSGCQCQQRIA     DCGCGLSGCQCQQRIA     DCGCGLSGCQCQQRIA     DCGCGLSGCQQQRIA     DCGCGLSGCQQQRIA     DCGCGLSGCQQQQXIA     CCCCCCQCQCCCCCCCCCCCCCCCCCCCCCCCCC	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPPVLI IARALIRPPVLI IARALLRDPEILI IARALLRDSPILI IARALYGAPTLVV LARAMYGDPCLII IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKIVI FALSLSVEFDCYL FGLSMAFKFDYVI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSEEHI LDEATSALDSEEHA LDEATSALSEEHA LDEATSALDSEEHA LDEATSALDSEEHA LDEATSALDSEEHA LDEATSAL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMUNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAIDALR-KNRT IQAIDALQ-KNRTS VNVAIKM—-NITR ILAAIQAIKARGATV IVDNLIALDKTL IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFVAL-SENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTV (385- IAIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- IIIAHRLSTI (352- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (503- IFIAHRLSTA (505- ITVAHRLSTV (491- INVTHRISTV (491- INVTHRIRGL (359- IFISHSINAA (371- IFISHRLVAA (358- ILVSHSPSAM (24-1 ILVSHSPSAM (24-1 INVSHSERAL (24-1	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -564) 180 -562) 180 -542) 181 -689) 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -536) 178 -536) 178 -536) 178 -536) 185 -542) 185 -84) 161 -84) 161 -86) 163 -86) 163 -66)
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>FrtD</li> <li>ComA</li> <li>IC Carbon</li> <li>CylB</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>EexA</li> <li>CtrD</li> <li>KpsT-F</li> <li>KpsT-F</li> <li>Span</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG         PV       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFIEDRINGYLTQVG         DFIEDRINGYLTQVG         DFIEDRINGYLTQVG         DFIEDRINGYDTVLG         ELILSLPNGYDTVLG         ELILSLPNGYDTELG         EDIERMPLNYQTELT         ADIEQMQLGYQTELS         TDIENLPQGYHTRLS         EVLLGIPQYEKTIVS         LEKLLEDAGLNSWIG         TNNQYUDTQLGWF         SHSSYQFDTQLGWF         ELGCYLYEPVKR         ELGCYLYEPVKR         ELGCYLYEPVKR         CLCENDENLYT	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRIA     DRGVRLSGGGRQRIA     DRGVRLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGLSGGQRQRIA     DGGGLSGGQRQRIA     DGGGLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     CGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     CGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     CGGRQLSGGQRQRIA     DGGRQLSGGQRQRIA     CGGRQLSGGQRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGQRQRIA     CGGRQLSGGQRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGQRQRIA     CGGRQLSGGRQRIA     CGGRQLSGGRGRGRGRGRCA     CGGRQLSGGRQRIA     CGGRQLSGQRGRAA     CGGRQLSGQRG	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRNAPILV IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPELII IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKILI IARALLSPAKIVI FALSLSVEFDCYL FGLSMAFKFDYVI FGLSMAFKFDYVI	HDLALVD ?NSLRRQI 151, 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDSERL LDEATSALDSERL LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK LDEPTEGLDATTESQ LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPIAEKE IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQNMQKIC-QGRTV IQNMRDIC-GGRTV IQNMRDIC-GGRTV IQNMRDIC-GGRTV IQNMRDIC-GGRTV IQNATOBLQ-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATOBLQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IFDTFFSL-SKDKIG CKYELFEKR-KDRSI CKYELFEKR-KDRSI CAQLFKERH-KESSF	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTV (385- IIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRESTI (362- VIIAHRESTI (362- IIFIAHRLSTI (350- IFIAHRLSTV (491- IFIAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (358- IIVAHRLSTV (358- IIVAHRLSTV (358- IIVSHSPSAM (24-1 IIVSHSPSAM (24-1 IMVSHSENAL (24-1 IMVSHSLNSL (24-1	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -513) 180 -523) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -536) 178 -536) 178 -536) 178 -536) 180 -536) 180 -537) 180 -538) 180 -538) 180 -539) 180 -531) 180 -531) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 181 -668) 180 -542) 180 -542) 181 -668) 180 -542) 185 -542)
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>KapsB</li> <li>CyaB</li> <li>PrtD</li> <li>P</li></ol>	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTVG DFISELREGYNG DFISELRE	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQCQRIA     EQCAGLSGCQCQRIA     EQCAGLSGCQCQRIA     EQCAGLSGCQCQRIA     EQCAGLSGCQCQRIA     EQCAGLSGCQCQRIA     EQGAGLSGCQCQRIA     ERGNRLSGGERQRIA     ERGNRLSGCGCQQRIA     DRGVRLSGCQCQRIA     DRGVRLSGCQCQRIA     DGGGGLSGCQCQRIA     DGGGGLSGCQCQRIA     ESGFNLSGCQCQRIA     ESGFNLSGCQCQRIA     EGGRQLSGCCQQQKIA     DEGRQLSGCCQQQKIA     DEGRQLSGCQCQXIA    YSSGMKARLA    YSSGMRRLD	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARAILKNAPILV IARAILKNAPILV IARALLRDSPILI IARALYGAPTLVV LARAMYGDPCLLI LARALSPAKILI IA	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSESEHF LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPNASLDSEGDQA LDEPTSAMDNISEFE LDEPTSANDNISEFE LDEPTSALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPSAALDPVAEKE LDEPTSALDPVAEKE LDEPTGLDATTESQ LDEPTAGLAFKKK VDEVTAVGDARFKEK	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHQIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IVQAIVALQKRGATV IVDNLIALDKTL ILKNLLPLDKTL IVSKLLFMKDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDJFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF VWDIVRALVDAGTT-	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (362- VLITHRPALT (350- IFIAHRLSTI (352- VLITHRPALT (350- IFIAHRLSTA (505- IFIAHRLSTA (505- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IFVAHRLSTV (491- IMVTHRLRGL (359- IFISHSINAA (371- IFISHRLVAA (358- ILVSHSPSAM (24-1 ILVSHSPSAM (24-1 IMVSHSERAL (24-1 IMVSHSENSL (24-1) VLLTTQYLDE (27-2	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -533) 180 -542) 181 -6689 178 -531) 180 -529) 180 -529) 180 -529) 180 -529) 180 -531) 180 -529) 180 -552) 185 -531) 180 -555) 185 -542) 180 -555) 185 -542) 180 -555) 185 -542) 180 -536) 178 -555) 185 -542) 180 -542) 180 -542) 180 -542) 180 -542) 180 -542) 181 -6689 182 -679) 179 -686) 182 -555) 185 -542) 185 -542) 180 -542) 180 -542) 180 -542) 180 -542) 180 -542) 180 -542) 181 -6699 178 -555) 185 -542) 180 -555) 185 -542) 185 -542] 185 -542] 185 -542] 185 -542] 185 -542] 185 -542] 155 -542] 155 -545 -545 -545 -545 -545 -545 -545 -545 -545 -545 -545 -545 -545 -545
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>FrtD</li> <li>ConA</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>KpsT-F</li> <li>DrrA</li> <li>NodI</li> </ol>	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG OFIEDRLNGYDTVG QFIEEMPEGFDTKLG DFINKMDNGLDTVIG ELILSLPNGYDTVIG ELILSLPNGYDTVG ELILSLPNGYDTUG ELILSLPNGYDTELG EDIERMPLNYQTELS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS LEKLLEDAGLNSWLG TNNQYVLDTQLGNWF SHSSYQFDTQLGLWF ELGQYLYEPVKK ELGQYLYEPVKK ELGQYLYEPVKR CLGDARDRLLKT RLESKANTRVAD	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQCAGLSGCQRQRIA     EQGAGLSGCQRQRIA     ENGVGLSGCQRQRIA     ENGVLSGCGRQRIA     DRGVRLSGGERQRIA     DRGVRLSGCGRQRIA     ELGEGLSGCQRQRIA     DGGGLSGCQRQRIA     DDGGGLSGCQRQRIA     SDASSLSGCQRQRIA     SDASSLSGCQRQRIA     ESGFNLSGCQRQRIA     EGGRQLSGCQRQRIA     DEGRQLSGCQRQKIA     CEGRQLSGCQRQKIA     CEGRQLSGCQRQXXA     CEGRXRRT	FYIPDSGQILIDG 136 150 IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPPVLI IARALIRPPVLI IARALIRDPEILI IARALIRDPEILI IARALIRDSPILI IARALIRDSPILI IARALYRPGLVU LARAMYGDPCLLI IARALSPAQCFI IARALSPAQLVI FALSLSVEFDCYL FALSLSVEFDCYL FALSLAVEFDLLF IARALNDPQLLI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSEGUQA LDEPNASLDSEGUQA LDEPNASLDSEGUQA LDEPNASLDSEGUQA LDEPTSANDNISEFE LDEPTSANDNISEFE LDEPTSANDNISEFE LDEPSSALDPVAEKE LDEPSSALDPVAEKE LDEPTSALDPVAEKE LDEPTSALDATTESQ LDEPTSALDATESQ LDEPSALDPVAEKE LDEPTSALDPVAEKE LDEPTSALDPVAEKE LDEPTGLDRSRNQ LDEPTGLDPRSRNQ LDEPTGLDPRSRNQ	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV VKDAIDALR-KNRTT VKDAIDALR-KNRTT VKDAIDALR-KNRTT VKDAIDALR-KNRTT IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI ILKNLLPLDKTI IVSKLLFMKDKTI IVSKLLFMK-DKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFFSL-SKDKIG CKYELFEKR-KDRSI CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF VWDIVRALVDAGTT IWERLRSLLARGKT-	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (354- FIIAHRLSTV (354- FIIAHRLSTV (385- IAIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- VIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIIHRAGVL (352- IIIIIHRAGVL (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIHRAGVL (352- IIIIHRAGVL (352- IIIIAHRLSTI (352- IIIIAHRLSTI (352- IIIIIAHRLSTI (352- IIIIIHRAGVL (352- IIIIIHRAGVL (352- IIIIHRAGVL (352- IIIIIHRAGVL (352- IIIIIHRAGVL (352- IIIIIHRAGVL (352- IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -533) 180 -533) 180 -542) 181 -531) 180 -529) 180 -529) 180 -531) 180 -529) 180 -686) 182 -670) 180 -536) 179 -686) 182 -555) 185 -542) 185 -542
30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB 19 NisT 20 SpaB 21 BexA 22 CtrD 23 KpsT-I 24 KpsT-I 25 DrrA 26 NodI 27 CvcV	NSUS INLS IPAGEV?AIV 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFIEDRLNGYDTVG DFIEDRLNGYDTVG OFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFINKMDNGLDTVIG DFILENPGYDTVLG ELULSLPNGYDTELT ADIEQMQLGYQTELS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS LEKLLEDAGLNSWLG TNNQYVLDTQLGNWF SHSSYQFDTQLGLWF ELGQYLYEPVKR	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRIA     ERGNRLSGGERQRIA     ERGNRLSGGERQRIA     ERGVRLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     EGGRQLSGGPGRLA     QEGHQLSGGQWQKIA     DGGRQLSGGQWQKIA     DGGRQLSGGQWQKIA     DGGRQLSGGQWQKIA     DGGRQLSGGQWQKIA     DGGRQLSGGQWQKIA     CONNYSGMKARLA     CONNYSGMKARLA     CONNYSGMKARLA     CONNYSGMKRLT     CONNYSGMK	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRNPKILI IARALIRNPFILV IARAILROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPFILI IARALLROPFLUV LARAMYGDPCLLI IARALSPAQCFI LARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCVI FALSLSVEFDCVI FALSLSVEFDCVI FGLSMAFKFDYVI IARALSVYPDLLF LAGALINDPQLLIF LAGALINDPQLUF	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEHK LDEPTSALDVIAVGDRFACK IDEVIAVGDSRFACK IDEVIAVGDSRFACK IDEVIAVGDRFACK VDEVTAVGDARFKEK LDEPTTGLDPRSRNQ LDEPTTGLDPRSRNQ LDEPTTGLDPARHL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHCC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KDRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VIVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDYFVAL-SENNIS TFDTFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF VWDIVRALVDAGTT- IWERLSLLARGKT- FGGIMRDHLARGCT.	ENIALAYEGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (354- FIIAHRLSTV (385- IIAIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRLSTI (352- VLITHRPALT (350- IFIAHRLSTI (350- IFIAHRLTIA (503- IFIAHRLTIA (505- ITVAHRISTV (491- IFVAHRLNIA (358- IIVSHSEAM (24-1) ILVSHSESAM (24-1) INVSHSENAL (24-1) IMVSHSENAL (24-1) IMVSHSENAL (24-1) IMVSHSENAL (24-1) IMVSHSENAL (24-1) ILTTVILDE (21-2) ILTTVILDE (21-2) ILTTVILDE (21-2) IAATHMALGI (20-1)	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -572) 180 -533) 180 -564) 180 -522) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -555) 185 -542) 185 -542) 185 -542) 180 -689 178 -531) 180 -536) 178 -555) 185 -542) 185 -542 185 -543 185 -543 185 -543 185 -544 161 -86 163 -86 163 -80 163 -91 175 -91 169 -91 175 -91 175 -91 175 -91 175 -91 169 -91 175 -91 175 -91 175 -91 169 -91 175 -91 175 -91 175 -91 175 -91 175 -91 175 -91 169 -91 175 -91
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-G</li> <li>CyaB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>LcnC</li> <li>PedD</li> <li>PrtD</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>KpsT-F</li> <li>KpsT-F</li> <li>NodI</li> <li>CycV</li> <li>Parb</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DF ISELREGYNTIVG         PV       DF ISELREGYNTIVG         EF ISELREGYNTIVG         EF ISELREGYNTIVG         DF ISELREGYNTIVG         DF ISELREGYNTIVG         EF ISELREGYNTIVG         DF ISELREGYNTVG         QF IEEMPEGFDTKLG         DF INKMONGLDTVIG         ELVLRLPQGYDTVLG         ELVLRLPQGYDTVLG         ELISLPMGYDELT         AD IEQMQLGYDTELS         TD IENLPQGYHTKLS         EVLLG IPQYEKTIVS         LEKLLEDAGLNSWIG         TNNQYVLDTOLGNWF         SHSSYQFDTQLGLWF         ELGQYLYEPVKR	RSGSGKSTLTKLIQG     A site     121 135     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     EQCAGLSGCQQQQIA     ERGNRLSGGEQQQQIA     DRGVRLSGGCQQQQIA     ELGEGLSGCQQQQIA     ELGEGLSGCQQQQIA     DGGGLSGCQQQQIA     DGGGLSGCQQQQIA     EGGQLSGCQQQQIA     ESGFNLSGCQQQQIA     DGGQLSGCQQQQIA     DGGQLSGCQQQQIA     CGGQLSGCQQQQIA     DGGQLSGCQQQQIA     DGGQLSGCQQQQIA     DGGQLSGCQQQQIA     CGGQLSGCQQQQIA     CGGQQAGAAAAA     CCGAAAAAAAAAAAAAAAAAAAAAAAAA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPPVLI IARALIRPPVLI IARALLRDPEILI IARALLRDPEILI IARALLRDSPILI IARALYGAPTLVV LARAMYGDPCLII IARALLSPAQCFI LARALSPAQCFI LARALSPAQCFI LARALSPAQCFI LARALSPAC-ILI IARALLAPPE-SU IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-ILI IARALSPAC-YI FALSLSVEFDCYL FGLSMAFKFDYII IARALSVVTPDLLF LAGALINDPQLUI LARLSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ IARALSVVTPDLUZ	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSUSERL LDEATSALDSUSERL LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK LDEPTSALDNISEFE LDEPTSALDVETEAR LDEPTSALDVETEAK IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK LDEPTTGLDPRSRNQ LDEPTTGLDPRARHL LDEPTTGLDPRARHL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHQIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAIDALR-KNRT IQAIDALR-KNRT IQAIDALQ-KRGTV IQAIDALQ-NRTS VNVAIKNMNITR ILAAIQAIKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI VFSNLLDEKRTV ILELLAEM-MREKTV IFDJFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF VWDIVRALVDAGTT- IWERLRSLLARGKT- FGGIMEDHLARGCI	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (362- VIIAHRLSTI (352- VIIAHRLSTI (352- IIIAHRLSTI (352- IFIAHRLSTA (503- IFIAHRLSTA (359- IFIAHRLSTA (359- IFIAHRLSTA (359- IFIAHRLSTA (359- IFIAHRLSTA (359- IIVSHSSAM (24-1 IMVSHSENAL (24-1 IMVSHSE	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -672) 180 -572) 180 -533) 180 -564) 180 -562) 180 -562) 180 -542) 181 -689) 178 -531) 180 -531) 180 -531) 180 -542) 181 -669) 179 -679) 179 -686) 182 -570) 180 -536) 178 -536) 178 -555) 185 -542) 185 -542
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>KyB-J</li> <li>Roya</li> <li>NdvA</li> <li>HetA</li> <li>NdvA</li> <li>HetA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>CvaB</li> <li>AprD</li> <li>CvaB</li> <li>AprD</li> <li>CvaB</li> <li>AprD</li> <li>CvaB</li> <li>PhetA</li> <li>CvaB</li> <li>PhetA</li> <li>CvaB</li> <li>PhetA</li> <li>CvaB</li> <li>PhetA</li> <li>SpaB</li> <li>BexA</li> <li>CtrD</li> <li>KpsT-F</li> <li>DrrA</li> <li>KodI</li> <li>CycV</li> <li>HelA</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         EFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTIVG         DFISELREGYNTVG         DFIEDRLNGYDTVG         DFIEDRLNGYDTVG         DFIEDRENGYDTVG         ELILSLPNGYDTFGG         EUVLRIPQGYDTVG         ELILSLPNGYDTELG         EDIERMPLNYQTELT         ADIEQMQLGYQTELS         TDIENLPQGYHTRLS         EVLLGIPQYEKTIVS         LEKLLEDAGLNSWIG         TNNQYUDTQLGWWF         SHSSYQFDTQLGLWF         ELGCYLYEPVKK	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRIA     ERGNRLSGGERQRIA     DRGVRLSGGQRQRIA     DRGVRLSGGQRQRIA     DRGVRLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     ESGFNLSGGQRQRIA     CSGNFSGGQRQRIA     CSGNFSGGRQRIA     CSGNFSGGRQRIA     CSGNFSGGRRRIA     CSSGNFSRIA     CSSGMFRRIA     CSSGMFRARIA     CSSGMFRRIA     CSSGMFRRIA     CSSGMFRRIA     CSSGMFRRIA     CSSGMFRRIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRARIA     CSSGMFRRIA     CSSGMFRARIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALNNPKILI IARALIRPRVLI IARALIRNPILV IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPEILI IARALLROPELII IARALLROPELII IARALLSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKVI IARALSVEFDCYL FALSLSVEFDCYL FGLSMAFKFDYVI IAASIVVTPDLLF IARALSVEFDVWI	HDLALVD ?NSLRRQI 151, 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDVETEAR LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEHK LDEPTSLDATTESQ LDEPTSAMDNISEFE LDEPTSGLDATTESQ LDEPSAALDPVAEKE LDEPTSGLDATESQ LDEPTSGLDATESQ LDEPTGLDPRSRNQ LDEPTTGLDPRSRNQ LDEPTTGLDPARHL LDEPTVSLDAASVAL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHCC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-GGRTV IQRNMRDIC-GGRTV IQRNMRDIC-KGRTV IQRATOBLQ-KNRTT VKDAIDALR-KDRTT VKDAIDALR-KDRTT IQES IEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IVDNLIALDKTI IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IFDTFFSL-SKDKIG CKYELFEKR-KDRSI CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CSDIFDKIR-EKSHL CAQLFFERH-KESSF VWDIVRALVDAGTT FGGLMRDHLARGGLI FAEAVRAHLAAGGAA	ENIALA?EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (354- FIIAHRLSTV (354- FIIAHRLSTI (362- VIAHRLSTI (362- VIAHRLSTI (362- VIAHRLSTI (352- VLITHRPALT (350- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLTIA (503- IFIAHRLSVA (501- IFIAHRLSVA (501- IFIAHRLSVA (501- IFVAHRLNIA (505- ITVAHRLSVA (501- IFVAHRLNIA (505- ITVAHRLSVA (501- IFIAHRLSVA (501- IIVSHSSAM (24-1) ILVSHSSAM (24-1) VLLTTQVLDE (27-2) IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLGI (201- IAATHMLSVA (501- IAATHMLGI (201- IAATHMLSVA (501- IAATHMLSVA (501-	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -533) 180 -564) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -535) 185 -542) 185 -542) 185 -542) 185 -542) 185 -543 -555) 185 -542) 185 -542) 185 -542) 185 -543 -555) 185 -542) 180 -555) 185 -542) 185 -542) 180 -555) 185 -542) 185 -542
<ol> <li>CONSE</li> <li>HlyB</li> <li>HlyB-I</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>LktB-J</li> <li>KovA</li> <li>PhetA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>MsbA</li> <li>MsbA</li> <li>PrtD</li> <li>P</li></ol>	NSUS       INLS IPAGEV?AIV9         106       120         DFISELREGYNTIVG       DFISELREGYNTIVG         PV       DFISELREGYNTIVG         EFISELREGYNTIVG       EFISELREGYNTIVG         EFISELREGYNTIVG       DFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         DFISELREGYNTIVG       EFISELREGYNTIVG         QFIEEMPEGFDTKLG       DFINKMONGLDTVIG         DFINKMONGLDTVIG       ELVLRLPQGYDTULG         ELVLRLPQGYDTLIG       EUILSLPNGYDTELG         EDIERMPINYQTELI       ADIEQMQLGYQTELS         TDIENLPQGYHTRLS       EVILGIPQYEKTIVS         ELGLYLEPQGYLTELS       TONQYVLDTQLGNWF         SHSSYQFDTQLGLWF       ELGQYLYEPVKR	RSGSGKSTLTKLIQG A site 121 135 EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA EQCAGLSGCQQQRIA ENGVLSGCQQQRIA ERGNRLSGGERQRIA ENGVLSGCQQQRIA ELGEGLSGCQQQRIA DGGGLSGCQQQRIA SDASGLSGCQQQRIA SDASGLSGCQQQRIA ESGFNLSGCQQQRIA ESGFNLSGCQQQRIA EGGQLSGCQQQRIA EGGQLSGCQQQKIA EGGQLSGCQQQKIA DEGQLSGCQQQKIA DEGQLSGCQQQKIA DEGQLSGCQQQKIA DEGRQLSGCQQQKIA	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPRVLI IARALIRPPEILI IARALIRDSPILI IARALIRDSPILI IARALYRAPTLVV IARALYRAPTLVV LARAMYGDPCLLI LARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKILI IARALSPAKIVI IARALSPAKIVI IARALSPAKIVI IARALSPAKIVI IARALSPAKIVI IARALSPAKIVI IARALSPAKVLI IARALSPAKVLI IARALSPAKVVI IARALSPAKVVI FGLSMAFKFDVVI IAASIVVTPDLLF LAGALINDPQLUI LARLLVTGRPVWI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI IDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDSESEHF LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPNASLDSEGQA LDEPTSALDNITTEKK FDESTSNLDTITEKK FDESTSNLDTITEKK IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK IDEVIAVGDARFKEK LDEPTTGLDPRSRNQ LDEPTTGLDPRSRNQ LDEPTTGLDPRARHL LDEPTTGLDPASVAL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHQIC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-DGRTV VKAAVDALR-KNRTT VKDAIDALR-KNRTT IQESIEKLS-VGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI ILELLAEM-MREKTV IFDJFFSL-SKDKIG CKYELFEKR-KDRSI CSDIFDKIR-EKSHL CAQLFKERH-KESSF VWDIVRALVDAGTT- IWERIRSLLARGKT- FGGLMRDHLARGGLI FAEAVRAHLAAGGAA	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIAHRLSTV (385- IIAHRLSTV (385- IAIAHRLSTI (362- VLITHRAGVL (352- VLITHRAGVL (352- VLITHRAGVL (352- IFIAHRLTIA (503- IFIAHRLTA (505- IFVAHRLSTV (491- IFVAHRLSTV (491- IMVTHRLRGL (359- IFISHSINAA (371- IFISHRLVAA (358- ILVSHSPSAM (24-1) ILVSHSPSAM (24-1) IMVSHSERAL (24-1) IMVSHSERAL (24-1) IMVSHSERAL (24-1) IMVSHSERAL (24-1) IMVSHSERAL (20-1) IAATHMALGI (20-1) IMATHIDLGL (22-1)	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -667) 180 -572) 180 -533) 180 -564) 180 -522) 180 -542) 181 -669) 178 -531) 180 -529) 180 -529) 180 -529) 180 -542) 181 -669) 178 -531) 180 -529) 180 -531) 180 -529) 180 -555) 185 -542) 180 -555) 185 -542) 185 -542) 180 -555) 185 -542) 155 -542) 155 -542 -542) 155
<ol> <li>CONSE</li> <li>HlyB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>LktB-I</li> <li>AppB</li> <li>CyaB</li> <li>ChvA</li> <li>NdvA</li> <li>HetA</li> <li>MsbA</li> <li>CvaB</li> <li>AprD</li> <li>AprD</li> <li>PrtD</li> <li>CvaB</li> <li>PrtD</li> <li>ConA</li> <li>PrtD</li> <li>CvaB</li> <li>PrtD</li> <li>CvaB</li> <li>PrtD</li> <li>CvaB</li> <li>PrtD</li> <li>Solar</li> <li>PrtD</li> <li>SurB</li> <li>NisT</li> <li>SpaB</li> <li>BexA</li> <li>KpsT-F</li> <li>DrrA</li> <li>NodI</li> <li>CycV</li> <li>HelA</li> <li>McbF</li> </ol>	NSUS       INLS IPAGEV?AIV9         106       120         DF ISELREGYNTIVG         PV       DF ISELREGYNTIVG         EF ISELREGYNTIVG         EF ISELREGYNTIVG         DF ISELREGYNTVG         DF IEEMPEGFDTKLG         DF INSKMONGLDTVIG         DF ISELNGYDTVIG         ELISLPNGYDTVIG         ELILSLPNGYDTELG         ED IERMPLNYQTELI         AD IEQMQLGYQTEIS         TD IENLPQGYHTRLS         EVLLG IPQYEKTIVS         LEKLEDAGINSWIG         ELGYLYEPVKK	RSGSGKSTLTKLIQG     A site     121 135     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     EQGAGLSGGQRQRIA     ERGNRLSGGERQRVA     DRGVRLSGGGRQRIA     ELGEGLSGGQRQRIA     DGGGGLSGGQRQRIA     DGGGGLSGGQRQRIA     SDASSLSGGQRQRIA     SDASSLSGGQRQRIA     ENGSNFSGGQRQKIA     EGGRNLSGGEPGRLA     QEGHQLSGGQWQKIA     DEGRQLSGGQWQKIA    YSSGMKARLA    YSSGMRRLG    LSGGMKRRLT    LSGQRKRLI    VSYGEKRWLI	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALIRPPVLI IARALIRPPVLI IARALIRDPEILI IARALIRDPEILI IARALIRDSPILI IARALYRPGILF IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACILI IARALSPACCYL FALSLSVEFDCYL FALSLAVEFDYII IAASIVVTPDLIF IARALSVTRPIWL IARLSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL IARALSVTRPIWL ISLMVTLCKNARLFL	HDLALVD ?NSLRRQI 151, 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR MDEATSALDVETEAR MDEATSALDVETEAR LDEPTSALDVETEAR LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK FDESTSNLDTITEKK IDEPTSAMDNISEFE LDEPTEGLDATTESQ LDEPSAALDPVAEKE LDEPTSALDPIAEKK IDEVIAVGDSRFAEK IDEVIAVGDSRFAEK LDEPTTGLDPRSRNQ LDEPTTGLDPRSRNQ LDEPTTGLDPRARHL LDEPTVGIDIQYRMM	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHQIC-KGRTV IQRNMRDIC-DGRTV IQRNMRDIC-DGRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKSKNIG IFDTFFSL-SKDKIG CKYELFEKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KGRSI CKYELFEK	ENIALA7EGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (493- FIIAHRLSTV (385- IIIAHRLSTI (363- IVIAHRLSTI (362- VIAHRLSTI (362- VIAHRETTL (512- ILITHRAGVL (352- VLITHRPALT (350- IFIAHRLSTA (501- IFIAHRLSTA (501- IFIAHRLSTA (501- IFIAHRLSTA (503- IFIAHRLSTA (503- IIVSHSSA (24-1) MVSHSSERA (24-1) MVSHSSE	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -668) 180 -668) 180 -668) 180 -562) 180 -564) 180 -562) 180 -562) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -686) 182 -670) 180 -536) 178 -535) 185 -542) 157 -589) 169 -59 -59 -59 -59 -59 -59 -59 -5
30 CONSE 1 HlyB 2 HlyB-I 3 LktB-I 4 LktB-I 5 AppB 6 CyaB 7 ChvA 8 NdvA 9 HetA 10 MsbA 11 CvaB 12 AprD 13 PrtD 14 ComA 15 LcnC 16 PedD 17 CylB 18 SurB 19 NisT 20 SpaB 21 BexA 22 CtrD 23 KpsT-I 24 KpsT-I 25 DrrA 26 NodI 27 CycV 28 HelA 29 McbF 30 CONSEN	NSUS INLS IPAGEV?ATV9 106 120 DFISELREGYNTIVG PV DFISELREGYNTIVG EFISELREGYNTIVG EFISELREGYNTIVG DFISELREGYNTIVG DFISELREGYNTIVG DFIEDRLNGYDTVG OFIEEMPEGFDTKLG DFINKMDNGLDTVIG DFILDRLNGYDTVG CFIEMPEGFDTKLG DFINKMDNGLDTVIG DFILDRLNGYDTVG ELULSLPNGYDTVLG ELULSLPNGYDTELT ADIEQMQLGYQTELS TDIENLPQGYHTRLS EVLLGIPQYEKTIVS LEKLLEDAGLNSWLG TNNQYVLDTOLGNWF SHSSYQFDTQLGLWF ELGGYLYEPVKR RLGKYFDMPIKT RLSKANTRVAD GLDAATHLPAAF NLNALEHRAAAS RYDKIRRKRTYT	RSGSGKSTLTKLIQG A site 121 135 EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA EQGAGLSGGQRQRIA ERGNRLSGGERQRIA ERGNRLSGGERQRIA ERGNRLSGGERQRIA ERGVLSGGQRQRIA ELGEGLSGGQRQRIA DGGGGLSGGQRQRIA DGGGGLSGGQRQRIA SDASSLSGGQRQRIA ESGFNLSGGQRQRIA ESGFNLSGGQRQRIA EGGRQLSGGQRQRIA EGGRQLSGGQRQRIA DGGGQLSGGQRQRIA ESGFNLSGGQRQRIA DGGGGLSGGQRQRIA ESGFNLSGGQRQRIA DGGGGLSGGQRQRIA ESGFNLSGGQRQRIA DGGGQLSGGQRQRIA COMPACTION SDASSLSGGRQRIA ESGFNLSGGQRQRIA DGGRQLSGGQRQRIA DGGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA DCGRQLSGGQRQRIA COMPACTION DCGRQLSGQRQRIA DCGRQLSGQRQRIA DCGRQLSGQRQRIA DCGRQLSGQRQRIA DCGRQLSGQRQRIA DCGRQRIA DCGGQRGQRQRIA DCGGQRGQRQRIA DCGGQRGQRQRIA DCGGQRGQRQRIA DCGGGGGGQRQRIA DCGGGGGGQRQRIA DCGGGGGGGQRQRIA DCGGGGGGGQRQRIA DCGGGGGGGQRQRIA DCGGGGGGGGGGGRQRIA DCGGGGGGGGGGGQRQRIA DCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	FYIPDSGQILIDG FYIPDSGQILIDG IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALVNNPKILI IARALKNAPILV IARAILKNAPILV IARAILKNAPILV IARALLRDSPILI IARALLRDSPILI IARALLRDSPILI IARALLRDSPILI IARALLRDSPILI IARALLRDSPUV LARAMYGDPCLLI IARALSPAQCFI LARALSPAQCFI LARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCFI IARALSPAQCYI FALSLSVEFDCYL FALSLSVEFDCYL FGLSMAFKFDYYI IARAILVTGRPVW ISLMVTLCKNARLFL IARALLNNPPILI	HDLALVD ?NSLRRQI 151 165 FDEATSALDYESEHV FDEATSALDYESEHV FDEATSALDYESEHI FDEATSALDYESEHI FDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDYESEHI LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEATSALDVETEAR LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDDSGEQA LDEPNSNLDTITEHK LDEPTSALDVIJEKK FDESTSNLDTITEHK LDEPTSALDVAEKE LDEPTSALDPVAEKE LDEPTSALDPVAEKE LDEPTGLDARFACK VDEVTAVGDARFACK LDEPTTGLDPRSNQ LDEPTTGLDPARHL LDEPTTGLDPASVAL LDEPTVGLDAASVAL	GVV?QDNVLFN?SIR 166 180 IMRNMHKIC-KGRTV IMRNMHKIC-QGRTV IMRNMHKIC-QGRTV IMRNMHCC-QGRTV IMQNMQKIC-QGRTV IMQNMQKIC-QGRTV IQRNMRDIC-GGRTV IQRNMRDIC-GRTV IQRNMRDIC-GRTV IQRATDELQ-KGRTV IQAATDELQ-KNRTS VNVAIKNMNITR LLAAIQALKARGCTV IMQAIVALQKRGATV IVDNLIALDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFMKDKTI IVSKLLFKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KDRSI CKYELFEKR-KDRSI CGUFKERH-KESSF VWD IVRALVAGGTI- IWELINKITADGKT IMDNLRALKGRTV	ENIALAYEGA 181 195 IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (488- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (489- IIIAHRLSTV (385- IIIAHRLSTV (385- IIIAHRLSTV (385- IIIAHRLSTI (383- LVIAHRLSTI (362- VIIAHRETTI (512- LLITHRAGVL (352- VLITHRPALT (350- IFIAHRLTIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (505- IFVAHRLNIA (358- ILVSHSEAM (24-1 ILVSHSESAM (24-1 IMVSHSENSL (24-1 IMVSH	SMEKVV?AAELAGAH -667) 180 -667) 180 -667) 180 -667) 180 -668) 180 -667) 180 -572) 180 -533) 180 -564) 180 -522) 180 -542) 181 -689 178 -531) 180 -529) 180 -681) 179 -679) 179 -689 182 -670) 180 -536) 182 -555) 185 -542) 185 -542) 185 -542) 185 -543 180 -536) 178 -555) 185 -542) 185 -542) 185 -542) 185 -542) 185 -542) 185 -555) 185 -542) 185 -542) 185 -542) 180 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -536) 178 -542) 180 -536) 178 -536) 175 -542) 1

FIG. 3. Sequence alignment and consensus sequence of the 190-aa domain containing the ABC. The proteins listed are described in detail in Table 1 and in the text. The first 20 sequences fall into evolutionary group A, the group of ABC exporters that contain the ABC and MSDs on one polypeptide. The last nine sequences fall into evolutionary group B, which have the ABC and MSDs on separate polypeptides (see Fig. 5). The alignment was generated by using the pattern construction algorithm PIMA (pattern-induced multiple alignment) written by Smith and Smith at the Molecular Biology Computer Research Resource, Dana-Farber Cancer Institute (165, 166). All pairwise comparisons between sequences in the set were performed, and the resulting scores were clustered into families by using MASE (41) to move gaps from conserved secondary-structure domains to predicted turn regions. The numbers in parentheses next to the alignments show the amino acid boundaries for the particular protein sequence. The number to the right of the parentheses on each line is the total size of the domain being compared from each system. The consensus was generated from MASE and simply shows the amino acid that occurs in each column with the highest frequency.

low level of secretion can occur through interaction with the ABC domain. Functional complementation experiments are consistent with this hypothesis.

# Why Are Dedicated Exporters Required?

Proteins can be transported through an ABC transporter even though they contain any number of posttranslational modifications including fatty acylation and unusual residues such as lanthionine. Some proteins which normally use the Sec-dependent pathway can also be secreted through ABC systems (100, 122). However, more interestingly, many molecules that cannot be transported through the Sec pathway can be transported via an ABC transporter. Given the limits of the Sec pathway, the ABC pathway is a useful alternative.

In the gram-negative bacterial systems, the presence of additional factors such as the accessory factor and the outer membrane factor allow for the secretion of proteins directly



FIG. 4. Sequence similarity between 29 of the bacterial ABC export systems calculated across the 190-aa ABC domain. The sequence similarity score was determined by comparing the amino acid in each position with the consensus amino acid sequence. The degree of sequence similarity was calculated by using the similarity score rules devised by Smith and Smith (165): identical matches were given a score of 5, conservative changes were scored 1 to 3 depending on the degree of conservation, and mismatches were scored as zero. The 1-aa window is shown by the dotted line, and a 7-aa window is shown by the thick solid line. Sequences corresponding to several of the highly conserved regions are also shown above the graph.

through both membranes. This allows proteins which could not normally pass through the outer membrane to move through it. The ABC export system is very efficient in this respect and is a very simple system when compared with the general secretion pathway for pullulanase, which requires over 15 proteins to facilitate protein transport through the outer membrane (139).

One possible reason why the dedicated exporters are required is that the proteins that utilize them cannot be transported via the Sec pathway. In at least one case, this hypothesis is not correct. We have shown that the OmpA signal sequence, when fused to the mature colicin V peptide, can permit colicin V to translocate across the inner membrane by using the Sec pathway (40). Interestingly, this colicin V remains trapped in the periplasm. Thus, the dedicated colicin V exporter has a marked advantage over the Sec pathway because it facilitates export through both membranes. It would be interesting to see whether other ABC protein/peptide substrates can be secreted by the Sec pathway, especially those from gram-positive systems such as the lantibiotics, which do not have an outer membrane to cross but which do have extensive posttranslational modification.

# What Does the Complex Look Like?

A basic ABC transporter structure has been proposed which includes four "core components," basically consisting of a dimer of ABCs and MSDs (Fig. 1) (80). This configuration is likely to be conserved among the bacterial ABC exporters, although the macrolide efflux systems currently are an exception since they lack MSDs. Most of the bacterial ABC exporters have their MSD and ABC in the same polypeptide. In these cases, the four core components would consist of a homodimer of each bacterial ABC exporter. There are few experimental data on the multimeric state of the bacterial ABC exporters, but the generation of antibodies and epitope tags to HlyB should allow such analysis to begin. (See Addendum in Proof.)

The presence of additional accessory factors is unique to bacterial ABC exporters. Accessory factors are present when secretion occurs through both membranes. In the absence of an accessory factor, some ABC transporters are still able to facilitate export into the periplasm. This raises an interesting question: is the accessory factor an integral part of the export complex? If the accessory factor is a peripheral component, the four core domains of the ABC exporter should be sufficient to facilitate translocation across the inner membrane (as in the NdvA system) even in the absence of an accessory factor and outer membrane factor. Thus, deletions of the accessory factor from other systems should result in periplasmic accumulation. Instead, deletion of the accessory factor in several systems results in cyto-



FIG. 5. Unrooted phylogenetic tree showing the evolutionary relatedness of the 190-aa domain containing the ABC. The computer program PAUP (Phylogenetic Analysis Using Parsimony) written by D. L. Swofford, Illinois Survey of Natural History, was used to generate the phylogenetic tree. The PAUP program uses DNA as well as protein sequence information when generating an evolutionary tree. For this analysis, DNA from the second position of each codon was used for the analysis, which is appropriate when comparing a diverse family of sequences. The total length of the tree is 607 units, and the consistency index for the tree is 0.427. The horizontal branch lengths indicate relative evolutionary distance.

plasmic accumulation of substrate (40, 65). These results suggest that there are several functionally distinct classes of bacterial ABC exporters. One class evolved in the presence of accessory factors and outer membrane factors and became dependent on all three components to form a functional export complex through both membranes of a gramnegative cell. The second class evolved to export molecules through a single membrane and can function independently of additional factors. To date, though, no physical interactions have been demonstrated between the ABC and the accessory factors, and the exact location of the accessory factors in the export complex has not been determined.

#### How Does the Export Complex Interact with Substrate?

The most detailed analyses of the interactions between substrate and exporter have been carried out on the bacterial permeases and eukaryotic exporters (3, 80). Specific interactions between bacterial ABC exporters and their substrates have not been well characterized, but a few ideas have been proposed. The secretion signal is located at the C terminus of several transported ABC protein substrates; therefore in these systems, secretion must be posttranslational. The large size of the proteins makes it likely that they pass through the ABC exporter in an unfolded state, but the chaperones GroEL, GroES, and SecB are not involved in HlyA or PrtB secretion (53, 184). There is, as of yet, no experimental evidence of unfolding, but if it does occur, it could be mediated either by host-encoded chaperones or by the secretion proteins themselves.

Newly synthesized HlyA can be found in the inner membrane even in the absence of the HlyBD exporters. This raises the possibility that interaction of HlyA with the exporters occurs within the plane of the membrane. Export of glucans and capsular polysaccharides also appears to occur after these oligomers are synthesized at inner membrane biosynthetic complexes. This type of membrane interaction is similar to the "flippase" model suggested for the P-glycoprotein, which recognizes and transports lipophilic drugs. It was suggested that P-glycoprotein interacts with drugs that are in the plane of the inner membrane and functions to flip them from the inner face of the lipid bilayer to the outer face (81). Localization of the transported substrate to the inner membrane could restrict the movement of these molecules and increase the likelihood of substratetransporter interaction. Under these conditions, substrate specificity could be more relaxed, because only few types of molecules would be properly localized to interact with the export complex in the proper orientation.

# CONCLUSIONS

The bacterial exporter subfamily of ABC transporters is a large and diverse family of proteins sharing a conserved ATP-binding domain and a common export function. In the past few years, the number of identified bacterial systems has greatly increased. Many proteins secreted by bacterial ABC exporters are important virulence factors or have important industrial applications. Other putative exporters have been found to be essential for bacterial growth and development and should provide insight into our understanding of basic cellular physiology. Each bacterial ABC exporter that is discovered demonstrates the diversity and the importance of this superfamily. Much has been learned about these exporters, and our knowledge about these systems should significantly contribute to our overall understanding of the superfamily of ABC transporters.

# THE SURVEY

# Proteus vulgaris and Morganella morganii Hemolysins

The hemolysins from *Proteus vulgaris* and *Morganella* morganii have the same genetic organization and functional

characteristics as the *E. coli* hemolysin system (105). Complementation experiments show that export proteins HlyB and HlyD are completely exchangeable between these two systems and the *E. coli* hemolysin system. The *M. morganii* hemolysin was also shown to be biochemically and immunologically related to the *E. coli* hemolysin (33).

The hlyB gene has been sequenced only from Proteus vulgaris (107). P. vulgaris HlyB is 707 aa and is highly homologous to the E. coli HlyB. In initial functional studies, several mutations were constructed in the region adjacent to the B site of P. vulgaris HlyB. Two conservative changes, G605A and K625R, had little effect on extracellular hemolysin levels, but the mutations G608R and K625I greatly reduced hemolysin activity. Biochemical analysis will be necessary to determine the effect of these mutations on P. vulgaris HlyB function.

#### Actinobacillus Hemolysins and Leukotoxins

At least three Actinobacillus species produce extracellular hemolysins or leukotoxins in the RTX family. Actinobacillus pleuropneumoniae and A. suis each secrete a 104-kDa RTX hemolysin (AppA and AshA, respectively) thought to be virulence factors in porcine pleuropneumonia, and A. actinomycetemcomitans secretes a closely related 125-kDa leukotoxin implicated in juvenile periodontitis. The A. pleuropneumoniae hemolysin was first studied immunologically, and antibodies raised against A. pleuropneumoniae hemolysin serotype 1 strain CM-5 were shown to cross-react with supernatants from other A. pleuropneumoniae serotypes, Pasteurella haemolytica, A. suis, and alpha-hemolysin-producing E. coli (26).

The E. coli HlyC, HlyB, and HlyD proteins were shown to be capable of secreting active A. pleuropneumoniae hemolysin from E. coli (72). More recently, the genes encoding one of the complete A. pleuropneumoniae hemolysin operons were cloned into E. coli and sequenced. One operon contains the appCA genes with a downstream appBpseudogene (16). An unlinked operon contains the appBD genes found with an appA pseudogene upstream. AppBD, HlyBD, and P. haemolytica LktBD all function in E. coli to secrete the AppA hemolysin (17). Analysis of operons from several different A. pleuropneumoniae serotypes suggest that there are, in fact, several distinct hemolysin determinants in each of the serotypes including AppA (17), HlyIIA (47, 72), and ApxIIIA (94a). The operons that contain these determinants appear to have undergone rearrangement by deletion to yield noticeable genetic variation (46, 94a). In A. suis, the hemolysin determinant has also been sequenced and found to be in a split configuration, where the ashCA and ashBD genes are unlinked and flanked by pseudogenes (12).

The A. actinomycetemcomitans leukotoxin production genes (*lktCA*) have sequence similarity to *hlyCA*. A. actinomycetemcomitans LktC is 168 aa. The 1,055-aa A. actinomycetemcomitans LktA protein toxin appears to remain associated with the outer membrane, in contrast to the extracellular secretion of hemolysin (109, 115). This membrane localization could be due to a C-terminal hydrophobic domain unique to the A. actinomycetemcomitans LktA. The gene encoding the A. actinomycetemcomitans leukotoxin ABC transporter is *lktB* (68). A 200-bp region of the gene encoding LktD has been sequenced and predicts a partial protein containing 72% identities with HlyD (109). LktB and LktD have not yet been shown to be required for secretion, and no other secretion functions have been identified.

# Pasteurella haemolytica Leukotoxin

Pasteurella haemolytica serotype A1 is a virulence factor involved in a bovine respiratory pneumonia known as shipping fever. *P. haemolytica* secretes a ruminant-specific leukotoxin that migrates as a 102-kDa protein on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The genes encoding the leukotoxin system have been cloned and sequenced (84, 171). The operon is identical in form with the *E. coli* operon. LktC is 167 aa, LktA is 983 aa, LktB is 708 aa, and LktD is 478 aa. The Hly genes are highly homologous to the Lkt genes, with amino acid sequence similarities ranging from 50 to 90%. LktBD and HlyBD are also functionally conserved—HlyBD and LktBD are interchangeable and each set functions to secrete either toxin (15, 84, 171).

# Bordetella pertussis Cyclolysin

Bordetella pertussis is one of the few prokaryotic organisms to secrete an adenylate cyclase, called cyclolysin. The B. pertussis cyclolysin is a key virulence factor and has multiple activities, including adenylate cyclase, invasin, and hemolysin. These properties are included on the single large polypeptide, CyaA (63). CyaA is produced as a 216-kDa protein which is secreted into the extracellular medium (79). Several groups have reported lower-molecular-mass species of cyclolysin (114, 126), but these appear to be proteolytic by-products. Genetic and functional analyses demonstrate that the N-terminal 450 aa have the adenylate cyclase and invasiveness activities while the C-terminal 1,250 aa have the hemolytic activity (63).

A total of five linked genes have been found to be required for cyclolysin production, activation, and secretion. cyaAencodes the 1,706-aa cyclolysin, and cyaC encodes the 185-aa protein required to activate the hemolytic moiety of CyaA (5). Three genes encode the secretion functions: cyaBencodes the 712-aa ABC transporter, cyaD encodes the 440-aa accessory factor, and cyaE encodes the 474-aa outer membrane factor (63). CyaE was the first outer membrane component shown to be associated with an ABC transport complex. All three genes have been shown to be required for secretion of the cyclolysin (63).

The HlyBD secretion system is capable of secreting CyaA when cloned into *E. coli* (125, 159). To facilitate analysis of cyclolysin, the complete secretion system has been reconstructed in *E. coli* (158). Surprisingly, the *cyaBDE* genes are synthesized in *E. coli* but are unable to promote CyaA secretion (158). In *E. coli*, CyaA secretion is facilitated by the presence of the HlyBD/TolC system, which recognizes a secretion signal at the C-terminal 217 aa of CyaA. When larger fragments of CyaA are fused to the C-terminal secretion signal, the efficiency of CyaA secretion decreases (159). Several groups studying heterologous secretion have observed that the size of the secreted product affects the efficiency of heterologous secretion (39, 119, 159).

# Pseudomonas aeruginosa Alkaline Protease

*P. aeruginosa* is an opportunistic pathogen whose virulence is related to the secretion of several proteins, including an alkaline protease (90). The genes in the alkaline protease operon could function to secrete active protease from *E. coli*, but only when put under the control of the *tac* promoter (70). No processing of protease was observed, and protease secretion was found to be SecA independent (71).

# 1008 FATH AND KOLTER

The alkaline protease operon includes five genes. aprA encodes the 479-aa alkaline protease, and aprI encodes a 130-aa protease inhibitor. Secretion requires three other genes: aprD encodes the 593-aa ABC exporter, aprE encodes the 432-aa accessory factor, and aprF encodes the 481-aa outer membrane factor (31). Mutations in the AprDEF region prevent protease secretion in *E. coli*. The alkaline protease could be secreted from *E. coli* by PrtDEF with high efficiency and by HlyBD/TolC at low efficiency, consistent with their degree of sequence similarity (31, 69). This system has been recently reviewed (176).

# Serratia marcescens Zinc Metalloprotease

S. marcescens produces an abundant extracellular metalloprotease. The gene encoding the zinc metalloprotease (prtSM) was cloned, sequenced, and found to encode a protein with 470 aa and a predicted molecular mass of 50.6 kDa. PrtSM lacks a signal peptide but possesses a 17-aa N-terminal propeptide which is cleaved off during processing. This protease was found to have significant sequence similarity to the Erwinia chrysanthemi and P. aeruginosa proteases (131). When expressed alone in E. coli, the protease is not secreted. However, significant amounts of extracellular protease were observed when the PrtSM protein was produced in the presence of either Erwinia chrysanthemi PrtDEF (118) or E. coli HlyBD/TolC (172). Additional experiments localized the secretion signal recognized by PrtDEF to the C terminus of PrtSM (118). To date, no PrtSM secretion genes have been identified from S. marcescens, but it is likely that an ABC-type exporter will be found associated with PrtSM secretion.

### Rhizobium leguminosarum NodO

The nodulation region of the R. leguminosarum genome contains a gene called nodO, which encodes a 284-aa secreted calcium-binding protein (34). NodO does not have a signal sequence and is not processed at its N terminus (25a). The sequence is slightly similar to HlyA, CyaA, and LktA in the RTX glycine repeat region (25a, 34). NodO secretion from E. coli could be facilitated by either HlyBD/TolC or PrtDEF. All three secretion genes must be present for NodO secretion through either heterologous system, providing strong evidence that an ABC exporter is required for NodO secretion (152). There are two genes linked to nodO in the R. leguminosarum chromosome, nodI and nodJ, which encode a multicomponent ABC transport system (see below). However, surprisingly, nodI mutants still produce secreted NodO (34). The ndvA gene, also found in R. leguminosarum, is an unlinked ABC transporter involved in glucan export (see above). Mutants with mutations in ndvA were also able to secrete NodO (152). Thus, another, as yet unidentified, R. leguminosarum gene(s) must be involved in the secretion of NodO.

# Neisseria meningitidis FrpA and FrpC

When starved for iron, *N. meningitidis* produces a variety of outer membrane proteins (174). One of these proteins, FrpA, has been cloned and sequenced and found be a 1,115-aa protein with sequence similarity to HlyA and other RTX toxins (174). A monoclonal antibody to FrpA crossreacts with HlyA and CyaA. A second outer membrane protein, FrpC, has been identified by using cross-reacting antibody that is also related to the RTX toxins. No transporters have been identified for FrpA or FrpC, but their sequence and outer membrane localization suggest that an ABC transporter may play a role in their secretion (174).

# Enterococcus faecalis Bacteriocin/Hemolysin

Sixty percent of clinical Enterococcus strains have a cytolytic toxin which lyses human, rabbit, and horse erythrocytes. This toxin is also bactericidal against a broad range of gram-positive bacteria (93). Toxin activity requires two components. The L component is an 11-kDa precursor of the lytic toxin which is activated extracellularly by the 27-kDa A (activator) component (64). The genes encoding these components are localized to the large transmissible plasmid, pAD1. Mutagenesis of pAD1 identified a 7.0-kb region required for bacteriocin/hemolysin production (93). A fragment from the L region was sequenced and found to encode a 714-aa, HlyB-like protein called CylB. Mutations in cylB prevent secretion of component L but, interestingly, do not affect secretion of component A. CylB is expressed in E. coli and can complement a cylB mutation in trans, but it fails to complement an hlyB mutation (60). The nucleotide sequences of components A and L, and other associated factors have not yet been determined. The sequence adjacent to cylB is likely to encode additional functions necessary for hemolysin/bacteriocin maturation or secretion. CylB was the first described ABC exporter identified in gram-positive bacteria.

# Bacillus subtilis Subtilin

Subtilin is one of a number of lantibiotics produced and secreted by gram-positive bacteria. Mature subtilin is a 32-aa peptide antibiotic with many unusual amino acids including lanthionine and dehydroalanine (95). The subtilin operon was cloned from *B. subtilis* ATCC 6633 and sequenced by two groups, who report somewhat different results (18, 103). *spaS* encodes the 56-aa subtilin primary product, which is posttranslationally modified and processed to the mature 32-aa form. Adjacent to *spaS* is *spaC*, which encodes a 441-aa product that is probably involved in subtilin processing.

Two genes also linked to spaS probably have secretionrelated functions: spaB (called spaT by Klein et al. [103]) and spaD (called spaB by Klein et al. [103]). SpaB is an ABC exporter predicted to have 599 or 614 aa depending on which ATG is used. Both groups have disrupted SpaB by cassette mutagenesis, with conflicting results. Chung et al. found that spaB insertions prevent external subtilin activity, whereas Klein et al. observed unusual clumping and slow-growth phenotypes but still observed external subtilin activity. Klein et al. suggest that their mutant is unable to secrete subtilin, leading to an internal accumulation which may be responsible for the observed effects. Neither experiment demonstrated a specific secretion function for SpaB.

The two published sequences for SpaD have significant inconsistencies. Both groups identify an identical 177-aa open reading frame (ORF), but the remaining details differ. The SpaD sequence of Chung et al. predicts a 177-aa protein with some sequence similarity to HlyD; the sequence of Klein et al. identifies the 177-aa ORF as part of a 390-aa protein that is homologous to the C terminus of EpiB, a 990-aa epidermin biosynthetic protein. Klein et al. show that mutants with mutations in SpaD (called SpaB in their paper) do not secrete subtilin, but no specific secretion function is demonstrated. At this point, it is still unknown whether SpaD function is related to secretion or biosynthesis.

#### Lactococcus lactis Nisin

Nisin is a commercially important lantibiotic secreted by certain strains of Lactococcus lactis and active against a wide range of gram-positive bacteria. It is widely used as a food preservative. The nisin peptide (NisA) is ribosomally synthesized as a 57-aa propeptide (97) and processed and modified into a 32-aa mature form (95). The three-dimensional structure of nisin has been determined by nuclear magnetic resonance spectroscopy (180a). The nisin operon has been sequenced, and three adjacent genes have been found: nisB, nisT, and nisC (36). nisT encodes a 600-aa ABC export protein. Expression of nisT in E. coli is toxic, and no secretion function has yet been identified for nisT. nisB encodes a 993-aa membrane-bound protein, and nisC encodes a 418-aa protein. Both have homologues in the subtilin and epidermin systems and probably encode biosynthetic enzymes. No other export-related genes were found in the nisin operon.

# Staphylococcus epidermidis Epidermin

Epidermin is another lantibiotic active against many grampositive bacteria (95). Epidermin is synthesized as a 52-aa propeptide and processed to the tetracyclic 21-aa mature form (155). The epidermin genes are located on a 54-kb plasmid, and at least six genes are involved in epidermin synthesis (155). *epiA* encodes the 52-aa propeptide. Downstream of *epiA* are *epiB*, *epiC*, *epiD*, *epiP*, and *epiQ*, which all appear to have a role in epidermin biosynthesis (4, 153).

Two intriguing open reading frames are found in the opposite orientation upstream of epiA. One, called epiT, could encode a maximum of 148 aa and has a good Shine-Dalgarno sequence but no identifiable start codon. With a -1frame shift, a further reading frame of 275 aa follows and continues past the end of the cloned fragment. The second ORF, called epiT', has the consensus A site, but the sequence ends before the beginning of any identifiable B site. It is possible that these genes compose a two-part ABC exporter or that they form a single ABC exporter through a shift in reading frame. Surprisingly, the 13.5-kb region cloned from this plasmid is sufficient to yield extracellular epidermin when expressed in the nonproducing organism Staphylococcus carnosus (153), even though epiT' does not contain a complete ABC domain. Much work is necessary to elucidate the mechanism of epidermin secretion, but it appears that an ABC exporter may be involved.

# Staphylococcus gallidermidis Gallidermin

Gallidermin is a lantibiotic produced by S. gallidermidis and is nearly identical in structure to epidermin (154). The gdmA gene encodes a 52-aa propeptide which is processed to a 21-aa mature form. Gallidermin differs from epidermin by only one amino acid: the N-terminal Ile in epidermin is a Leu in gallidermin (154). A gallidermin operon structure has been reported which matches that of epidermin (103). No DNA sequences have been published, but there is a clear possibility that gallidermin secretion is also mediated by an ABC exporter.

# Streptococcus pneumoniae Competence Factor

Competence factor is a 10-kDa secreted protein required to facilitate genetic transformation in *Streptococcus pneumoniae* (175). The *com* locus was cloned (14), sequenced (92), and found to contain several genes required for transformation. One of these genes, *comA*, encodes a 717-aa ABC export protein with a predicted molecular mass of 80.3 kDa (92). Mutations in *comA* and the adjacent *comB* gene have a severe competence deficiency (14, 129). *comB* encodes a 49-kDa protein that has no significant similarities to known proteins (129). ComA and ComB are both hypothesized to be part of a dedicated export complex (129). If this is indeed the case, it would be the first example of a required accessory factor in a gram-positive secretion system. The gene(s) encoding the competence factor has not yet been identified and appears to be unlinked to the *comAB* locus.

# Pediococcus acidilactici Pediocin PA-1

Pediocin PA-1 is a small, heat-stable bacteriocin produced from plasmid pSRQ11 in *Pediococcus acidilactici* PAC1.0 and is active against many gram-positive bacteria (76). The protein has been purified and sequenced and found to be a 44-aa, 4,629-Da peptide with sequence similarity to bacteriocins from other gram-positive bacteria (120). The pediocin production operon was cloned and sequenced (124). Four ORFs were found: *pedA*, *pedB*, *pedC*, and *pedD*. PedA is the 62-aa precursor of pediocin A; PedB and PedC are 112 and 174 aa, respectively, and have no known function. PedD is a 724-aa ABC exporter. The *pedD* gene was mutated and shown to be required for pediocin production.

# Lactococcus lactis Lactococcins A and G

The lactococcin A system was cloned and characterized from several strains of Lactococcus lactis (89, 170, 179, 180). Lactococcin A is synthesized as a 75-aa precursor with a 21-aa N-terminal extension (89). The complete lactococcin A operon contains four genes: lcnA encodes the 75-aa precursor, and *lciA* encodes a 98-aa immunity protein (89), whereas lcnC and lcnD encode a putative export system (170). LcnC is a 716-aa ABC transport protein with a predicted molecular mass of 79.9 kDa. LcnD is a 474-aa protein (52.5 kDa) with slight sequence similarity to the HlyD and PrtE accessory factors. Both LcnC and LcnD have been inactivated by transposon mutagenesis and found to be essential for lactococcin A activity (170). Another newly discovered bacteriocin, lactococcin G from L. lactis, also requires additional genes for its synthesis. These genes have been cloned and sequenced. One encodes a putative immunity protein and the other encodes an ABC exporter (74).

## Escherichia coli Microcin B17

Microcin B17 is a small, highly modified 43-aa peptide antibiotic that inhibits DNA replication (21). The proteins McbE and McbF act to transport microcin B17 into the periplasm, where it associates with the outer membrane and passes through into the extracellular medium. McbEF also play a role in conferring immunity to the producing cell by secreting microcin B17 out away from its cytoplasmic target, DNA gyrase (51). The *mcbEF* genes were cloned and sequenced from the microcin-producing plasmid pMccB17. McbE is a 241-aa hydrophobic protein with six potential transmembrane domains and a predicted molecular mass of 27.9 kDa. McbF is a 28.9-kDa, 247-aa protein that contains the consensus A and B sites common to ABC transporters. Insertion mutations in *mcbE* and *mcbF* result in a complete loss of extracellular microcin B17, but active microcin can be detected in cytoplasmic extracts (51). McbEF is the first multicomponent ABC transporter to be identified that is clearly involved in export (51).

# Streptomyces peucetius DrrAB

The gram-positive bacterium S. peucetius produces the antitumor agents daunorubicin and doxorubicin. Several genes involved in the biosynthesis of daunorubicin and doxorubicin have been identified by being cloned into the nonproducing strain S. lividans (136). Genes that are involved in daunorubicin and doxorubicin resistance have been sequenced and found to include a multicomponent ABC transport system (67). The genes, drrAB, encode two proteins: DrrA, an ABC protein with 330 aa and a predicted molecular mass of 35.7 kDa; and DrrB, a very hydrophobic protein with 283 aa and a predicted molecular mass of 30.6 kDa. When the drrAB operon is cloned into S. lividans, the genes confer Dnr and Dxr resistance to the heterologous host. drrAB are cotranscribed and are expressed only during antibiotic production (67). The authors suggest that DrrA may energize transport of antibiotics by ATP hydrolysis through a membrane pore made by DrrB. The discovery of bacterial ABC exporters that facilitate export of Dnr and Dxr is particularly exciting because these are two of the antibiotics that are exported by the mammalian P-glycoprotein encoded by the mdr gene (35).

# Haemophilus influenzae BexAB

A polymer of ribose and ribitol-5-phosphate makes up the capsule of type b strains of *H. influenzae* (20). The locus involved in capsule formation is organized as a directly repeated duplication separated by a 1.3-kb "bridge" region (85). Recombination that removes this bridge results in loss of capsule, suggesting that an essential promoter or gene is present in the region (110). The bridge region was sequenced and found to contain the gene *bexA*, which encodes a 24.7-kDa, 217-aa ABC transporter (110). A frameshift mutation in *bexA* results in accumulation of internal polysaccharide (110).

The region around *bexA* was then sequenced and three additional ORFs were identified: *bexB*, *bexC*, and *bexD* (111). BexB is 30.2-kDa, 265-aa hydrophobic protein homologous to KpsM. Although the organization of the *bex* locus is distinct from that of the *kps* locus, they each contain genes sufficient for production and export of capsular polysaccharide. The BexAB and KpsTM exporters seem to be structurally and functionally similar.

# Neisseria meningitidis CtrCD

The capsular polysaccharides produced by *N. meningitidis* are also major virulence factors. The group B *N. meningitidis* capsule is composed of  $\alpha$ -2,8-linked polysialic acid and is related to the *E. coli* and *H. influenzae* capsules described above. A 24-kb plasmid sufficient to produce capsule in *E. coli* K-12 was characterized, and five phenotypically distinct regions were identified (50). Mutations in region A abolish all polysaccharide synthesis; those in region B produce mutants that contain only cytoplasmic polysaccharide; those in region C produce mutants that contain polysaccharide in the cytoplasm and periplasm; and those in regions D and F affect levels of polysaccharide (48).

Region C, which seems to be required for secretion of polysaccharide from the periplasm to the cell surface, was sequenced (49), and four genes, ctrABCD, were found. CtrC is a 264-aa, 30.1-kDa inner membrane protein homologous to BexB, and CtrD is a 215-aa, 24.6-kDa cytoplasmic ABC transporter homologous to BexA. Together, CtrCD could form an ABC transporter. These results are somewhat unexpected because the ctrABCD cluster is located in region C, which is implicated in outer membrane transport, while the homologous bex locus from H. influenzae is implicated in inner membrane transport. In a recent review, the authors propose that CtrBCD form an inner membrane ABC transport complex that associates with the outer membrane protein CtrA. This complex would recognize membraneassociated polysaccharide and facilitate its transfer to the outer membrane (48). In this scenario, CtrB and CtrC are heterodimeric subunits of the transport complex. However, the hydropathy analysis of CtrB is not consistent with that of CtrC, and CtrB may not have the transport role as proposed.

# Bradyrhizobium japonicum CycVW

Bacterial *c*-type cytochromes consist of an apoprotein covalently attached to a 600-Da protoheme IX cofactor. Cytochrome c is found free in the periplasm. Pleiotropic Tn5 insertion mutations were identified in Bradyrhizobium *japonicum* that gave a negative reaction in the cytochrome c oxidase colony assay (143). These mutants were cloned and sequenced, and several ORFs were found. Three of these mutants no longer made c-type cytochromes and were designated cycVWX. CycV is a 200-aa, 21.1-kDa ABC transporter, and CycW is a 222-aa, 22.7-kDa protein predicted to be in the inner membrane. Mutations in the small (61-aa) protein CycX also abolish cytochrome activity. An additional ORF, ORF263, was predicted to encode a 263-aa inner membrane protein, but mutations in ORF263 did not abolish cytochrome activity (143). The authors speculate that CycVW may form a complex that transports out an essential component of cytochrome c. Other experiments suggest that the apoprotein is exported by a signal-sequencedependent method and the heme group is attached to the apoprotein in the periplasm (137). That suggests that CycVW may export the cytochrome c-heme lyase or possibly the heme molecule itself.

#### **Rhodobacter capsulatus HelABC**

A region homologous to the *B. japonicum cyc* region was cloned and sequenced from the photosynthetic bacterium *Rhodobacter capsulatus* and subjected to insertion mutagenesis (7). Four genes were characterized and found to be required for cytochrome c biogenesis. HelA (the CycV homolog) is a hydrophilic 214-aa ABC transporter; HelB is a hydrophobic 218-aa protein homologous to CycW; HelC is a hydrophobic 242-aa protein homologous to ORF263; and ORF52 is a small (52-aa) protein homologous to CycX.

The observation that HelC is required for cytochrome biosynthesis in *R. capsulatus* contrasts with the *B. japonicum* result which suggests that ORF263 is not an essential gene. The small ORF52 and CycX proteins have no counterparts in other bacterial ABC systems. Their role in the translocation process and in cytochrome c production is not yet clear. Cytochrome c-PhoA fusions show that the apocytochrome is secreted to the periplasm in Hel<sup>-</sup> mutants. The

authors suggest that the *hel* genes are not involved in import of iron or in export of holo- or apocytochrome c. Instead, they favor the model that the Hel complex exports the heme group to the periplasm where it is then attached to the apocytochrome.

# Anabaena Sp. Strain PCC 7120 HepA

The cyanobacterium Anabaena sp. produces nitrogenfixing heterocysts when starved for nitrogen under aerobic conditions (73). A specialized external polysaccharide layer is formed early in heterocyst development. A mutant that reduces the cohesiveness of the polysaccharide layer was identified and sequenced (86). The mutated gene was originally called *hetA* (86) but was recently renamed *hepA* to reflect the fact that it affects heterocyst envelope polysaccharide (36a). *hepA* encodes a protein with a maximum size of 607 aa, but some evidence suggests that a shorter (532-aa) protein might be produced instead. The HepA protein is similar to other members of the ABC exporter family and is hypothesized to be involved in export of the heterocyst envelope polysaccharide.

#### Escherichia coli SurB

The surB gene was identified by looking for genes required for E. coli survival during the stationary phase (161, 177). Mutations in one gene, surB, resulted in cells unable to exit from the stationary phase and resume aerobic growth at high temperature (162). The surB gene was cloned, mapped to 19.5 min on the E. coli chromosome, and sequenced. The sequence indicates that surB encodes a putative 573-aa ABC transporter with both an ABC and potential MSDs. The gene directly upstream of surB (called orfU) was also sequenced and found to contain a 243-aa ABC. The genetic organization of orfU and surB is particularly interesting. orfU ends with the sequence TAA, and the ATG of surB starts immediately afterward in the same frame. There is an intriguing possibility that OrfU and SurB form as a single polypeptide at some frequency as a result of readthrough. Insertion mutagenesis has demonstrated that surB is not essential for normal E. coli growth (162). surB or orfU may be allelic with cydC, which is involved in heme-d production or secretion. Consistent with this, a surB mutant does lack a functional cytochrome d oxidase (162). Therefore one surB function may be to aid in assembly of heme-d into the cytochrome d oxidase complex.

# Escherichia coli MsbA

The htrB gene is required for growth of E. coli at temperatures above 33°C. Genes were identified that suppress a temperature-sensitive htrB mutant when present on multicopy plasmids (98, 99). One of the suppressor genes is msbA, which maps to 20.5 min on the E. coli chromosome and encodes a 582-aa protein with a predicted molecular mass of 64.5 kDa (98). The MsbA protein shares significant sequence similarities to the bacterial ABC exporters. Insertion mutations in msbA were found to be lethal, demonstrating that msbA is an essential gene in E. coli. This is the first example of an essential ABC exporter. The msbA gene is cotranscribed with another essential gene called orf E, which could encode a 328-aa, 35.6-kDa protein. The authors suggest that MsbA and OrfE could function together to increase the export of polyamines, which restore growth of htrB mutants at high temperatures, or that they could export a toxic molecule that accumulates at nonpermissive temperatures in the absence of HtrB (98).

# Escherichia coli FtsE

An operon involved in cell division was found at 76 min on the E. coli chromosome and contains three genes, including one encoding the putative ABC transporter, ftsE (57). ftsE is the second gene in a three-gene operon and encodes a 222-aa, 24.4-kDa protein. The genes in this operon are, in order, ftsY-ftsE-ftsX (57). FtsY is a 497-aa protein which runs as a 92-kDa protein on SDS-PAGE and has sequence similarity to the SR $\alpha$  protein of eukaryotes involved in protein secretion (59). FtsX is a hydrophobic 352-aa, 38.5kDa protein. All three proteins are associated with the inner membrane (58). Conditional-lethal mutations were isolated in FtsE and found to cluster in the ABC region, between the A site and B site and near to important cystic fibrosis transmembrane regulator mutations (56). No specific transport function has been identified yet, but these studies suggest that FtsYEX may form a complex that transports septation-specific proteins to the periplasm (56).

Several genes that encode putative ABC transporters have now been identified on the chromosome of *E. coli*, including *msbA*, ORF, *surB*, and *ftsE*. It is important to remember that these genes do exist in the *E. coli* chromosome and that they may be able to complement other plasmid-borne ABC exporters at low levels. So far, though, none of these genes appear to complement mutations in ABC exporters such as HlyB, PrtD, or CvaB when these systems are expressed in *E. coli*, because mutations in these ABC transporters are not complemented when expressed in cells containing wild-type MsbA, SurB, OrfU, and FtsE.

# Streptomyces griseus AmfA and AmfB

Genes involved in aerial mycelium formation in *Strepto-myces griseus* were identified by looking for genes which can suppress an A-factor mutation when expressed at high copy number (178a). Among the genes found were two that encode putative ABC translocators, AmfA and AmfB. The authors hypothesize that AmfA and AmfB are involved in the transport of proteins and peptides that serve as intercellular signals during aerial mycelium formation.

# Rhizobium leguminosarum and Bradyrhizobium japonicum NodLJ

NodIJ were identified on the symbiotic plasmid pRL1J1 of R. leguminosarum and shown to be involved in nodulation (37). nodI encodes a 311-aa ABC protein with a predicted molecular mass of 34.3 kDa. nodJ encodes a hydrophobic 259-aa protein with a predicted molecular mass of 27.7 kDa. There is no direct evidence that NodIJ has any transport role, and it is known that NodIJ does not function to secrete the hemolysin-like NodO protein (34). It was postulated that NodIJ may form a permease that imports a compound exuded from plant roots and taken up by Rhizobium cells during nodulation (37). Recently the NodIJ locus from B. japonicum was also cloned and sequenced and found to be highly homologous to R. leguminosarum NodIJ and the polysaccharide exporters (181). The authors of this report hypothesize that NodIJ may actually be involved in export of lipopolysaccharides which are produced by other nod genes (181), but the role of NodIJ as an importer or exporter has not yet been determined.

# Rhizobium meliloti ORF1

A gene has been identified upstream of the transcriptional regulator *ntrA* in *R. meliloti* that is a member of the ABC transporter family (1). The gene, called ORF1, encodes a 230-aa protein and is part of a larger operon that contains additional unidentified upstream genes. Genes similar to ORF1 have been found upstream of *ntrA* in *Salmonella typhimurium* and *Klebsiella pneumoniae*. No function has been identified for these genes, although there is evidence that ORF1 is an essential gene in *R. meliloti* and is unrelated to *ntrA* function. No transport function has been identified for ORF1 yet, but the sequence suggests that it may encode an ABC transporter.

# Staphylococcus epidermidis MsrA

A 1.9-kb DNA fragment containing novel erythromycin resistance genes from *S. epidermidis* was introduced into a sensitive *S. aureus* strain and shown to be functional (147). The fragment contained a gene, *msrA*, which encodes a 488-aa protein. MsrA contains two ATP-binding domains separated by a long Q-linker. Q-linkers make up a family of interdomain sequences found in many prokaryotic multidomain proteins but not yet seen in ABC transport proteins (192). MsrA alone, when subcloned into *S. aureus*, is sufficient to promote efflux of [<sup>14</sup>C]erythromycin. This function is not unlike the immunity function observed by the microcin B17 exporters McbEF (51). If any hydrophobic proteins are required by MsrA to facilitate erythromycin resistance, they must also be present in *S. aureus*.

# Streptomyces TIrC, SrmB, and CarA

The tylosin-producing strain *S. fradiae* contains at least three genes specifying tylosin resistance. One of these genes, *tlrC*, was cloned and sequenced and found to encode a 548-aa, 59.1-kDa protein that has the ABC-Q-ABC structure seen in MsrA (149). Additional macrolide resistance genes have been found in related *Streptomyces* strains: *srmB* from *S. ambofaciens* and *carA* from *S. thermotolerans* (156). The proteins TlrC, SrmB, and CarA have 66 to 76% sequence identity with each other. These macrolide resistance proteins share sequences common to other ABC transporters but have an unusual domain configuration and lack any identifiable hydrophobic components. Additional work is necessary to determine the precise functional mechanism in these systems and their precise place in the family of bacterial ABC exporters.

# **ACKNOWLEDGMENTS**

We are grateful to many researchers in the field for communicating unpublished results with us. We thank Temple Smith and Thomas Grath, Molecular Biology Computer Research Resource, for their valuable assistance with the phylogenetic analysis; Zico Kolter for help in preparing Fig. 4; and Kira and Emma Fath for inspiration.

Our research has been supported by Cystic Fibrosis Foundation grant Z138 and Public Health Service grant AI25944 from the National Institutes of Health.

# ADDENDUM IN PROOF

Since the submission of this paper, at least five new bacterial ABC export systems have been identified or partially described. These include a third RTX toxin from A.

actinomycetemcomitans (Y.-F. Chang et al., DNA Cell Biol. 12:351-362, 1993), a leukotoxin system from a *P. haemolyt*ica-like organism (Y.-F. Chang et al., Infect. Immun. 61: 2089-2095, 1993), an ABC exporter in *Aeromonas salmonicida* associated with a secreted surface array protein (S. Chu and T. J. Trust, J. Bacteriol. 175:3105-3114, 1993), an oleandomycin resistance determinant from *Streptomyces* antibioticus (A. M. Rodriguez et al., Mol. Microbiol. 8:571-582, 1993), and FrpC, a second RTX toxin gene from *N. meningitidis* (S. A. Thompson et al., Mol. Microbiol. 9:85-96, 1993).

In addition, significant progress has been made in analyzing the ABC domain from HlyB. The C terminus of HlyB has been overproduced and shown to form a dimer possessing both ATP binding and ATPase activity, with activities similar to those of the importer MalK and the P glycoprotein (V. Koronakis et al., Mol. Microbiol. 8:1163–1175, 1993). These results provide direct evidence in support of several of the hypotheses discussed in this review.

#### REFERENCES

- 1. Albright, L. M., C. W. Ronson, B. T. Nixon, and F. M. Ausubel. 1989. Identification of a gene linked to *Rhizobium* meliloti ntrA whose product is homologous to a family to ATP-binding proteins. J. Bacteriol. 171:1932–1941.
- Ames, G. F.-L. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. Annu. Rev. Biochem. 55:377-425.
- 3. Ames, G. F.-L., C. S. Mimura, and V. Shyamala. 1990. Bacterial periplasmic permeases belong to a family of transport proteins operating from *E. coli* to human: traffic ATPases. FEMS Microbiol. Rev. 75:429-446.
- Augustin, J., R. Rosenstein, B. Wieland, U. Schneider, N. Schnell, G. Engelke, K.-D. Entian, and F. Götz. 1992. Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of *Staphylococcus epidermidis*. Eur. J. Biochem. 204:1149-1154.
- Barry, E. M., A. A. Weiss, I. E. Ehrmann, M. C. Gray, E. L. Hewlett, and M. S. Goodwin. 1991. Bordetella pertussis adenylate cyclase toxin and hemolytic activities require a second gene, cyaC, for activation. J. Bacteriol. 173:720–726.
- Becker, K. F., H. Allmeier, and V. Hollt. 1992. New mechanisms of hormone secretion—MDR-like gene products as extrusion pumps for hormones. Horm. Metab. Res. 24:210–213.
- 7. Beckman, D. L., D. R. Trawick, and R. G. Kranz. 1992. Bacterial cytochromes c biogenesis. Genes Dev. 6:268–283.
- Blight, M. A., and I. B. Holland. 1990. Structure and function of haemolysin B, P-glycoprotein and other members of a novel family of membrane translocators. Mol. Microbiol. 4:873–880.
- Blight, M. A., B. Kenny, C. Chervaux, A. L. Pimenta, C. Cecchi, and I. B. Holland. 1992. The structure and function of the *Escherichia coli* haemolysin A (HlyA) toxin and the membrane localised transport protein, HlyB, p. 207-216. *In* E. Quagliariello and F. Palmieri (ed.), Molecular mechanisms of transport. Elsevier Science Publishing, Inc., New York.
   Boulnois, G. J., and I. S. Roberts. 1990. Genetics of capsular
- Boulnois, G. J., and I. S. Roberts. 1990. Genetics of capsular polysaccharide production in bacteria. Curr. Top. Microbiol. Immunol. 150:1–20.
- 11. Bradley, M. K., T. F. Smith, R. H. Lathrop, D. M. Livingston, and T. A. Webster. 1987. Consensus topography in the ATP binding site of the simian virus 40 and polyomavirus large tumor antigens. Proc. Natl. Acad. Sci. USA 84:4026-4030.
- Burrows, L. L., and R. Y. C. Lo. 1992. Molecular characterization of an RTX toxin determinant from *Actinobacillus suis*. Infect. Immun. 60:2166-2173.
- Cangelosi, G. A., G. Martinetti, J. A. Leigh, C. C. Lee, C. Theines, and E. W. Nester. 1989. Role of Agrobacterium tumefaciens ChvA protein in export of β-1,2-glucan. J. Bacteriol. 171:1609-1615.
- Chandler, M. S., and D. A. Morrison. 1987. Competence for genetic transformation in *Streptococcus pneumoniae*: molecu-

lar cloning of *com*, a competence control locus. J. Bacteriol. **169:**2005–2011.

- Chang, Y.-F., R. Young, T. L. Moulds, and D. K. Struck. 1989. Secretion of the *Pasteurella* leukotoxin by *Escherichia coli*. FEMS Microbiol. Lett. 60:169–174.
- Chang, Y. F., R. Young, and D. K. Struck. 1989. Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophilus) pleuropneumoniae. DNA 8:635-647.
- Chang, Y. F., R. Young, and D. K. Struck. 1991. The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked appCA and appBD loci flanked by pseudogenes. J. Bacteriol. 173:5151-5158.
- Chung, Y. J., M. T. Steen, and J. N. Hansen. 1992. The subtilin gene of *Bacillus subtilis* ATCC 6633 is encoded in an operon that contains a homolog of the hemolysin B transport protein. J. Bacteriol. 174:1417-1422.
- 19. Coote, J. G. 1992. Structural and functional relationships among the RTX toxin determinants of Gram-negative bacteria. FEMS Microbiol. Rev. 88:137-162.
- Crisel, R. M., R. S. Baker, and D. E. Dorman. 1975. Capsular polysaccharide of *Haemophilus influenzae* type b. J. Biol. Chem. 250:4926–4930.
- Davagnino, J., M. Herrero, D. Furlong, F. Moreno, and R. Kolter. 1986. The DNA replication inhibitor microcin B17 is a forty-three amino acid protein containing sixty percent glycine. Proteins 1:230–238.
- 22. Davagino, J., and R. Kolter. Unpublished results.
- 23. Delepelaire, P., and C. Wandersman. 1990. Protein secretion in Gram-negative bacteria. J. Biol. Chem. 265:17118-17125.
- Delepelaire, P., and C. Wandersman. 1991. Characterization, localization and transmembrane organization of the three proteins PrtD, PrtE and PrtF necessary for protease secretion by the Gram-negative bacterium *Erwinia chrysanthemi*. Mol. Microbiol. 5:2427-2434.
- Delepelaire, P., and S. Wandersman. 1989. Protease secretion by Erwinia chrysanthemi. J. Biol. Chem. 264:9083–9089.
- 25a.de Maagd, R. A., A. H. M. Wijfjes, H. P. Spaink, J. E. Ruiz-Sainz, C. A. Wijffelman, R. J. H. Okker, and B. J. J. Lugtenberg. 1989. nodO, a new nod gene of the Rhizobium leguminosarum biovar viciae sym plasmid pRLIJI, encodes a secreted protein. J. Bacteriol. 171:6764-6770.
- Devenish, J., S. Rosendal, R. Johnson, and S. Hubler. 1989. Immunoserological comparison of 104-kilodalton proteins associated with hemolysis and cytolysis in Actinobacillus pleuropneumoniae, Actinobacillus suis, Pasteurella haemolytica, and Escherichia coli. Infect. Immun. 57:3210–3213.
- Deverson, E. V., I. R. Gow, W. J. Coadwell, J. J. Monaco, G. W. Butcher, and J. C. Howard. 1990. MHC class II region encoding proteins related to the multidrug resistance family of transmembrane transporters. Nature (London) 348:738– 741.
- Doolittle, R. F., M. S. Johnson, I. Husain, B. V. Houten, D. C. Thomas, and A. Sancar. 1986. Domain evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. Nature (London) 323:451-453.
- Douglas, C. J., R. J. Staneloni, R. A. Rubin, and E. W. Nester. 1985. Identification and genetic analysis of an *Agrobacterium tumefaciens* chromosomal virulence region. J. Bacteriol. 161: 850–860.
- Dudler, R., and C. Hertig. 1992. Structure of an mdr-like gene from *Arabidopsis thaliana*: evolutionary implications. J. Biol. Chem. 267:5882-5888.
- Duong, F., A. Lazdunski, B. Cami, and M. Murgier. 1992. Nucleotide sequence of a cluster of genes involved in synthesis and secretion of alkaline protease in *Pseudomonas aeruginosa*: relationships to other secretion pathways. Gene 121:47–54.
- 32. Dylan, T., L. Ielpi, S. Stanfield, L. Kashyap, C. Douglas, M. Yanofsky, E. Nester, D. R. Helinski, and G. Ditta. 1986. *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. USA 83:4403-4407.
- 33. Eberspacher, B., F. Hugo, M. Pohl, and S. Bhakdi. 1990. Functional similarity between the haemolysins of *Escherichia*

coli and Morganella morganii. J. Med. Microbiol. 33:165-170.

- 34. Economou, A., W. D. O. Hamilton, A. W. B. Johnston, and J. A. Downie. 1990. The *Rhizobium* nodulation gene nodO encodes a Ca<sup>2+</sup>-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. EMBO J. 9:349-354.
- Endicott, J. A., and V. Ling. 1989. The biochemistry of P-glycoprotein-mediated multidrug resistance. Annu. Rev. Biochem. 58:137-171.
- Engelke, G., Z. Gutowski-Eckel, M. Hammelmann, and K.-D. Entian. 1992. Biosynthesis of lantibiotic nisin: genomic organization and membrane localization of the NisB protein. Appl. Environ. Microbiol. 58:3730–3743.
- 36a.Ernst, A., T. Black, Y. Cai, J.-M. Panoff, D. N. Tiwari, and C. P. Wolk. 1992. Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. J. Bacteriol. 174:6025– 6032.
- 37. Evans, I. J., and J. Downie. 1986. The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins; nucleotide sequence analysis of the *nodI* and *nodJ* genes. Gene 43:95–101.
- 38. Fath, M. J., R. Skvirsky, L. Gilson, H. K. Mahanty, and R. Kolter. 1992. The secretion of colicin V, p. 331–348. In R. James, F. Lazdunski, and F. Pattus (ed.), Bacteriocins, microcins, and lantibiotics. Springer-Verlag KG, Heidelberg, Germany.
- Fath, M. J., R. Skvirsky, and R. Kolter. 1991. Functional complementation between bacterial MDR-like export systems: colicin V, α-hemolysin, and *Erwinia* protease. J. Bacteriol. 173:7549-7556.
- 40. Fath, M. J., L. H. Zhang, J. Rush, and R. Kolter. Unpublished data.
- Faulkner, D. V., and J. Jurka. 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. 13:321-322.
- 42. Felmlee, T., S. Pellett, E.-Y. Lee, and R. A. Welch. 1985. *Escherichia coli* hemolysin is released extracellularly without cleavage of a signal peptide. J. Bacteriol. 163:88–93.
- 42a.Felmlee, T., S. Pellett, and R. A. Welch. 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. J. Bacteriol. 163:94–105.
- Felmlee, T., and R. A. Welch. 1988. Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion. Proc. Natl. Acad. Sci. USA 85:5269– 5273.
- 44. Foote, S. J., J. K. Thompson, A. F. Cowman, and D. J. Kemp. 1989. Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. Cell 57:921– 930.
- 45. Freund, S., and G. Jung. 1992. Lantibiotics: An overview and conformation studies on gallidermin and Pep5, p. 75–92. *In* R. James, C. Lazdunski, and F. Pattus (ed.), Bacteriocins, microcins and lantibiotics. Springer-Verlag KG, Heidelberg, Germany.
- Frey, J., M. Beck, U. Stucki, and J. Nicolet. 1993. Analysis of hemolysin operons in *Actinobacillus pleuropneumonia*. Gene 123:51–58.
- 47. Frey, J., H. van den Bosch, R. Segers, and J. Nicolet. 1992. Identification of a second hemolysin (HlyII) on *Actinobacillus pleuropneumoniae* serotype 1 and expression of the gene in *Escherichia coli*. Infect. Immun. 60:1671–1676.
- Frosch, M., and U. Edwards. 1993. Molecular mechanisms of capsule expression in *Neisseria meningitidis* serogroup B, p. 49–57. *In* J. Roth, U. Rutishauser, and F. A. Troy (ed.), Polysialic acid. Birkhauser Verlag, Basel.
- Frosch, M., U. Edwards, K. Bousset, B. Kraube, and C. Weisgerber. 1991. Evidence for a common molecular origin of the capsule gene loci in Gram-negative bacteria expressing group II capsular polysaccharides. Mol. Microbiol. 5:1251– 1263.
- 50. Frosch, M., C. Weisgerber, and T. F. Meyer. 1989. Molecular characterization and expression in *Escherichia coli* of the gene complex encoding the polysaccharide capsule of *Neisseria*

meningitidis group B. Proc. Natl. Acad. Sci. USA 86:1669-1673.

- Garrido, M. C., M. Herrero, R. Kolter, and F. Moreno. 1988. The export of the DNA replication inhibitor microcin B17 provides immunity for the host cell. EMBO J. 7:1853-1862.
- Gentschev, I., and W. Goebel. 1992. Topological and functional studies on HlyB of *Escherichia coli*. Mol. Gen. Genet. 232:40– 48.
- 53. Gentschev, I., J. Hess, and W. Goebel. 1990. Change in the cellular localisation of alkaline phosphatase by alteration of its carboxy-terminal sequence. Mol. Gen. Genet. 222:211–216.
- 54. Geremia, R. A., S. Cavaignac, A. Zorreguieta, N. Toro, J. Olivares, and R. A. Ugalde. 1987. A *Rhizobium meliloti* mutant that forms ineffective pseudonodules in alfalfa produces exopolysaccharide but fails to form beta-(1,2) glucan. J. Bacteriol. 169:880-884.
- 55. Ghigo, J.-M., and C. Wandersman. 1992. Cloning, nucleotide sequence and characterization of the gene encoding the *Erwinia chrysanthemi* B374 PrtA metalloprotease: a 3rd metalloprotease secreted via a C-terminal secretion signal. Mol. Gen. Genet. 236:135-144.
- 56. Gibbs, T. W., D. R. Gill, and G. P. C. Salmond. 1992. Localised mutagenesis of the *ftsYEX* operon: conditionally lethal missense substitutions in the FtsE cell division protein of *Escherichia coli* are similar to those found in the cystic fibrosis transmembrane conductance regulator protein (CFTR) of human patients. Mol. Gen. Genet. 234:121-128.
- Gill, D. R., G. F. Hatfull, and G. P. C. Salmond. 1986. A new cell division operon in *Escherichia coli*. Mol. Gen. Genet. 205:134-145.
- Gill, D. R., and G. P. C. Salmond. 1987. The Escherichia coli cell division proteins FtsY, FtsE and FtsX are inner membrane associated. Mol. Gen. Genet. 210:504–508.
- Gill, D. R., and G. P. C. Salmond. 1990. The identification of the *Escherichia coli ftsY* gene product: an unusual protein. Mol. Microbiol. 4:575-583.
- Gilmore, M. S., R. A. Segarra, and M. C. Booth. 1990. An HlyB-type function is required for expression of the *Entero*coccus faecalis hemolysin/bacteriocin. Infect. Immun. 58: 3914–3923.
- Gilson, L., H. K. Mahanty, and R. Kolter. 1987. Four plasmid genes are required for colicin V synthesis, export, and immunity. J. Bacteriol. 169:2466-2470.
- Gilson, L., H. K. Mahanty, and R. Kolter. 1990. Genetic analysis of an MDR-like export system: the secretion of colicin V. EMBO J. 9:3875–3884.
- Glaser, P., H. Sakamoto, J. Bellalou, A. Ullmann, and A. Danchin. 1988. Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of *Bordetella pertussis*. EMBO J. 7:3997–4004.
- Granato, P. A., and R. W. Jackson. 1969. Bicomponent nature of lysin from *Streptococcus zymogenes*. J. Bacteriol. 100:865– 868.
- 65. Gray, L., K. Baker, B. Kenny, N. Mackman, R. Haigh, and I. B. Holland. 1989. A novel C-terminal signal sequence targets *Escherichia coli* haemolysin directly to the medium. J. Cell Sci. Suppl. 11:45–57.
- 66. Gray, L., N. Mackman, J.-M. Nicaud, and I. B. Holland. 1986. The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*. Mol. Gen. Genet. 205:127–133.
- 67. Guilfoile, P. G., and C. R. Hutchinson. 1991. A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. Proc. Natl. Acad. Sci. USA 88:8553-8557.
- Guthmiller, J. M., D. Kolodrubetz, M. P. Cagle, and E. Kraig. 1990. Sequence of the *lktB* gene from *Actinobacillus actino*mycetemcomitans. Nucleic Acids Res. 18:5291.
- 69. Guzzo, J., F. Duong, C. Wandersman, M. Murgier, and A. Lazdunski. 1991. The secretion genes of *Pseudomonas aeruginosa* alkaline protease are functionally related to those of *Erwinia chrysanthemi* proteases and *Escherichia coli* α-haemolysin. Mol. Microbiol. 5:447-453.

- 70. Guzzo, J., M. Murgier, A. Filloux, and A. Lazdunski. 1990. Cloning of the *Pseudomonas aeruginosa* alkaline protease gene and secretion of the protease into the medium by *Esche*-
- richia coli. J. Bacteriol. 172:942–948.
  71. Guzzo, J., J.-M. Pages, F. Duong, A. Lazdunski, and M. Murgier. 1991. *Pseudomonas aeruginosa* alkaline protease: evidence for secretion genes and study of secretion mechanism. J. Bacteriol. 173:5290–5297.
- Gygi, D., J. Nicolet, J. Frey, M. Cross, V. Koronakis, and C. Hughes. 1990. Isolation of the *Actinobacillus pleuropneumoniae* haemolysin gene and the activation and secretion of the prohaemolysin by HlyC, HlyB, and HlyD proteins of *Escherichia coli*. Mol. Microbiol. 4:123–128.
- 73. Haselkorn, R., and W. J. Buikema. 1992. Nitrogen fixation in cyanobacteria, p. 165–189. *In* G. Stacey, H. W. Evans, and R. H. Burris (ed.), Biological nitrogen fixation. Chapman & Hall, New York.
- 74. Håvarstein, L. S., and I. Nes. Personal communication.
- 75. Henderson, D. M., C. D. Sifri, M. Rodgers, D. F. Wirth, N. Hendrickson, and B. Ullman. 1992. Multidrug resistance in *Leishmania donovani* is conferred by amplification of a gene homologous to the mammalian *mdr1* gene. Mol. Cell. Biol. 12:2855–2865.
- Henderson, J. T., A. L. Chopko, and P. D. v. Wassenaar. 1992. Purification and primary structure of pediocin PA-1 produced by *Pediococcus acidilactici* PAC-1.0. Arch. Biochem. Biophys. 295:5-12.
- Hess, J., I. Gentschev, W. Goebel, and T. Jarchau. 1990. Analysis of the haemolysin secretion system by PhoA-HlyA fusion proteins. Mol. Gen. Genet. 224:201-208.
- Hess, J., W. Wels, M. Vogel, and W. Goebel. 1986. Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence. FEMS Microbiol. Lett. 34:1-11.
- Hewlett, E. L., V. M. Gordon, J. D. McCaffery, W. M. Sutherland, and M. C. Gray. 1989. Adenylate cyclase toxin from Bordetella pertussis. J. Biol. Chem. 264:19379-19384.
- Higgins, C. F. 1992. ABC transporters: from microorganisms to man. Annu. Rev. Cell Biol. 8:67–113.
- 81. Higgins, C. F., and M. M. Gottesman. 1992. Is the multidrug transporter a flippase? Trends Biochem. Sci. 17:18-21.
- 82. Higgins, C. F., I. D. Hiles, G. P. C. Salmond, D. R. Gill, J. A. D. I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell, and M. A. Hermodson. 1986. A family of related ATPbinding subunits coupled to many distinct biological processes in bacteria. Nature (London) 323:448-450.
- Higgins, C. F., S. C. Hyde, M. M. Mimmack, U. Gileadi, D. R. Gill, and M. P. Gallagher. 1990. Binding protein-dependent transport systems. J. Bioenerg. Biomembr. 22:571-592.
- Highlander, S. K., M. Chidambaram, M. J. Engler, and G. M. Weinstock. 1989. DNA sequence of the *Pasteurella haemolytica* leukotoxin gene cluster. DNA 8:15-28.
- Hoiseth, S. K., C. J. Connelly, and E. R. Moxon. 1985. Genetics of spontaneous high-frequency loss of b capsule in *Haemophilus influenzae*. Infect. Immun. 49:389–395.
- Holland, D., and C. P. Wolk. 1990. Identification and characterization of *hetA*, a gene that acts early in the process of morphological differentiation of heterocysts. J. Bacteriol. 172: 3131-3137.
- Holland, I. B., M. A. Blight, and B. Kenny. 1990. The mechanism of secretion of hemolysin and other polypeptides from Gram-negative bacteria. J. Bioenerg. Biomembr. 22:473–491.
- Holland, I. B., B. Kenny, and M. Blight. 1990. Haemolysin secretion from *E. coli*. Biochimie 72:131-141.
- Holo, H., Ø. Nilssen, and I. F. Nes. 1991. Lactococcin A, a new bacteriocin from *Lactococcus lactis* subsp. *cremoris*: isolation and characterization of the protein and its gene. J. Bacteriol. 173:3879-3887.
- Howe, T. R., and B. H. Iglewski. 1984. Isolation and characterization of alkaline protease-deficient mutants of *Pseudomonas aeruginosa* in vitro and in a mouse eye model. Infect. Immun. 43:1058-1063.
- 91. Hughes, C., P. Stanley, and V. Koronakis. 1992. E. coli

hemolysin interactions with prokaryotic and eukaryotic cell membranes. BioEssays 14:519–525.

- 92. Hui, F. M., and D. A. Morrison. 1991. Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. J. Bacteriol. 173:372–381.
- Ike, Y., D. B. Clewell, R. A. Segarra, and M. S. Gilmore. 1990. Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn917 insertional mutagenesis and cloning. J. Bacteriol. 172:155-163.
- Issartel, J.-P., V. Koronakis, and C. Hughes. 1991. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature (London) 351:759-761.
- 94a.Jansen, R., J. Briaire, E. M. Kamp, A. L. J. Gielkens, and M. A. Smits. 1993. Cloning and characterization of the Actinobacillus pleuropneumoniae RTX-toxin III (ApxIII) gene. Infect. Immun. 61:947-954.
- Jung, G. 1991. Lantibiotics—ribosomally synthesized biologically active polypeptides containing sulfide bridges and alphabeta-didehydroamino Acids. Agnew. Chem. Int. Ed. 30:1051– 1068.
- Juranka, P., F. Zhang, J. Kulpa, J. Endicott, M. Blight, I. B. Holland, and V. Ling. 1992. Characterization of the hemolysin transporter, HlyB, using an epitope insertion. J. Biol. Chem. 267:3764–3770.
- Kaletta, C., and K.-D. Entian. 1989. Nisin, a peptide antibiotic: cloning and sequencing of the *nisA* gene and posttranslational processing of its peptide product. J. Bacteriol. 171:1597–1601.
- Karow, M., and C. Georgopoulos. 1993. The essential Escherichia coli msbA gene, a multicopy suppressor of null mutations in the htrB gene, is related to the universally-conserved family of ATP-dependent translocators. Mol. Microbiol. 7:69– 79.
- 99. Karow, M., and C. Georgopoulos. 1992. Isolation and characterization of the *Escherichia coli msbB* gene, a multicopy suppressor of null mutations in the high-temperature requirement gene htrB. J. Bacteriol. 174:702-710.
- 100. Kenny, B., R. Haigh, and I. B. Holland. 1991. Analysis of the haemolysin transport process through the secretion from *Escherichia coli* of PCM, CAT or  $\beta$ -galactosidase fused to the Hly C-terminal signal sequence. Mol. Microbiol. 5:2557-2568.
- Kenny, B., S. Taylor, and I. B. Holland. 1992. Identification of individual amino acids requires for secretion within the haemolysin (HlyA) C-terminal targeting region. Mol. Microbiol. 6:1477-1489.
- 102. Kerpolla, R. E., V. K. Shyamala, P. Klebba, and G. F.-L. Ames. 1991. The membrane-bound proteins of periplasmic permeases form a complex. J. Biol. Chem. 266:9857–9865.
- 103. Klein, C., C. Kaletta, N. Schnell, and K.-D. Entian. 1992. Analysis of genes involved in biosynthesis of the lantibiotic subtilin. Appl. Environ. Microbiol. 58:132–142.
- Kolter, R., and F. Moreno. 1992. Genetics of ribosomally synthesized peptide antibiotics. Annu. Rev. Microbiol. 46:141– 163.
- 105. Koronakis, V., M. Cross, B. Senior, E. Koronakis, and C. Hughes. 1987. The secreted hemolysins of *Proteus mirabilis*, *Proteus vulgaris*, and *Morganella morganii* are genetically related to each other and to the alpha-hemolysin of *Escherichia coli*. J. Bacteriol. 169:1509–1515.
- 106. Koronakis, V., C. Hughes, and E. Koronakis. 1991. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. EMBO J. 10:3263-3272.
- 107. Koronakis, V., E. Koronakis, and C. Hughes. 1988. Comparison of the haemolysin secretion protein HlyB from *Proteus vulgaris* and *Escherichia coli*: site-directed mutagenesis causing impairment of export function. Mol. Gen. Genet. 213:551-555.
- 108. Koronakis, V., E. Koronakis, and C. Hughes. 1989. Isolation and analysis of the c-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial mem-

branes. EMBO J. 8:595-605.

- 109. Kraig, E., T. Dailey, and D. Kolodrubetz. 1990. Nucleotide sequence of the leukotoxin gene from Actinobacillus actinomycetemcomitans: homology to the alpha-hemolysin/leukotoxin gene family. Infect. Immun. 58:920–929.
- Kroll, J. S., I. Hopkins, and E. R. Moxon. 1988. Capsule loss in H. influenzae type b occurs by recombination-mediated disruption of a gene essential for polysaccharide export. Cell 53:347-356.
- 111. Kroll, J. S., B. Loynds, L. N. Brophy, and E. R. Moxon. 1990. The bex locus in encapsulated Haemophilus influenzae: a chromosomal region involved in capsule polysaccharide export. Mol. Microbiol. 4:1853–1862.
- 112. Kroncke, K.-D., G. Boulnois, I. Roberts, D. Bitter-Suermann, J. R. Golecki, B. Jann, and K. Jann. 1990. Expression of the *Escherichia coli* K5 capsular antigen: immunoelectron microscopic and biochemical studies with recombinant *E. coli*. J. Bacteriol. 172:1085–1091.
- 113. Kuchler, K., R. E. Sterne, and J. Thorner. 1989. Saccharomyces cerevisiae STE6 gene product: a novel pathway for protein export in eukaryotic cells. EMBO J. 8:3973–3984.
- 114. Ladant, D. 1988. Interaction of *Bordetella pertussis* adenylate cyclase with calmodulin. J. Biol. Chem. 263:2612-2618.
- 115. Lally, E. T., E. E. Golub, I. R. Kieba, N. S. Taichman, J. Rosenbloom, J. C. Rosenbloom, C. W. Gibson, and D. R. Demuth. 1989. Analysis of the Actinobacillus actinomycetem-comitans leukotoxin gene. J. Biol. Chem. 264:15451–15456.
- 116. Lee, C., P. Li, H. Inouye, and J. Beckwith. 1989. Genetic studies on the inability of  $\beta$ -galactosidase to be translocated across the *Escherichia coli* cytoplasmic membrane. J. Bacteriol. 171:4609-4616.
- 117. Letoffe, S., P. Delepelaire, and C. Wandersman. 1990. Protease secretion by *Erwinia chrysanthemi*: the specific secretion functions are analogous to those of *Escherichia coli* α-haemolysin. EMBO J. 9:1375–1382.
- 118. Letoffe, S., P. Delepelaire, and C. Wandersman. 1991. Cloning and expression in *Escherichia coli* of the *Serratia marcescens* metalloprotease gene: secretion of the protease from *E. coli* in the presence of the *Erwinia chrysanthemi* protease secretion functions. J. Bacteriol. 173:2160–2166.
- 119. Letoffe, S., and C. Wandersman. 1992. Secretion of CyaA-PrtB and HlyA-PrtB fusion proteins in *Escherichia coli*: involvement of the glycin-rich repeat domain of *Erwinia chrysanthemi* protease B. J. Bacteriol. 174:4920–4927.
- Lonzano, J. C. N., J. N. Meyer, K. Sletten, C. Pelaz, and I. F. Nes. 1992. Purification and amino acid sequence of a bacteriocin produced by *Pediococcus acidilactici*. J. Gen. Microbiol. 138:1985-1990.
- 121. Lory, S. 1992. Determinants of extracellular protein secretion in gram-negative bacteria. J. Bacteriol. 174:3423-3428.
- 122. Mackman, N., K. Baker, L. Gray, R. Haigh, J.-M. Nicaud, and I. B. Holland. 1987. Release of a chimeric protein into the medium from *Escherichia coli* using the C-terminal secretion signal of haemolysin. EMBO J. 6:2835-2841.
- 123. Mackman, N., J.-M. Nicaud, L. Gray, and I. B. Holland. 1985. Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*. Mol. Gen. Genet. 210:529–536.
- 124. Marugg, J. D., C. F. Gonzalez, B. S. Kunka, A. M. Ledeboer, M. J. Pucci, M. Y. Toonen, S. A. Walker, L. C. M. Zoetmulder, and P. A. Vandenbergh. 1992. Cloning, expression, and nucleotide sequence of genes involved in production of pediocin PA-1, a bacteriocin from *Pediococcus acidilactici* PAC1.0. Appl. Environ. Microbiol. 58:2360-2367.
- 125. Masure, H. R., D. C. Au, M. K. Gross, M. G. Donovan, and D. R. Storm. 1990. Secretion of the *Bordetella pertussis* adenylate cyclase from *Escherichia coli* containing the hemolysin operon. Biochemistry 29:140–145.
- 126. Masure, H. R., and D. R. Storm. 1989. Characterization of the bacterial cell associated calmodulin-sensitive adenylate cyclase from *Bordetella pertussis*. Biochemistry 28:438-442.
- 127. McGrath, J. P., and A. Varshavsky. 1989. The yeast STE6 gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. Nature (London) 340:400–404.

- 128. Monaco, J. J., S. Cho, and M. Attaya. 1990. Transport protein genes in the murine MHC: possible implications for antigen processing. Science 250:1723–1726.
- 129. Morrison, D. A., F. M. Hui, and L. Zhou. 1993. Competence for transformation in *Streptococcus pneumoniae*: regulatory genes and sequences, p. 3-11. *In* E. Balla and G. Berencsi (ed.), Genetic transformation: DNA transfer and gene expression in micro-organisms. Intercept, Ltd., Publishers, Andover, United Kingdom.
- 130. Mosser, J., A.-M. Douar, C.-O. Sarde, P. Kioschis, R. Feil, H. Moser, A.-M. Poustka, J.-L. Mandel, and P. Aubourg. 1993. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature (London) 361:726-730.
- 131. Nakahama, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Hase, and H. Matsubara. 1986. Cloning and sequencing of *Serratia* protease gene. Nucleic Acids Res. 14:5843–5855.
- 132. Niki, H., R. Imamura, T. Oruga, and S. Hiraga. 1990. Nucleotide sequence of the *tolC* gene of *Escherichia coli*. Nucleic Acids Res. 18:5547.
- 133. Nishi, K., M. Yoshida, M. Nishimura, M. Nishikawa, M. Nishiyama, S. Horinouchi, and T. Beppu. 1992. A leptomycin B resistance gene of *Schizosaccharomyces pombe* encodes a protein similar to the mammalian P-glycoproteins. Mol. Microbiol. 6:761–769.
- 134. Oropeza-Wekerle, R. L., W. Speth, B. Imhof, I. Gentschev, and W. Goebel. 1990. Translocation and compartmentalization of *Escherichia coli* hemolysin (HlyA). J. Bacteriol. 172:3711– 3717.
- 135. Ortiz, D. F., L. Kreppel, D. M. Speiser, G. Scheel, G. McDonald, and D. W. Ow. 1992. Heavy metal tolerance in the fission yeast requires an ATP-binding cassette-type vacuolar membrane transporter. EMBO J. 11:3491-3499.
- 136. Otten, S. L., K. J. Stutzman-Engwall, and C. R. Hutchinson. 1990. Cloning and expression of daunorubicin biosynthesis genes from *Streptomyces peucetius* and *S. peucetius* subsp. *caesius*. J. Bacteriol. 172:3427-3434.
- 137. Page, M. D., and S. J. Ferguson. 1989. A bacterial c-type cytochrome can be translocated to the periplasm as an apo form; biosynthesis of cytochrome cd1 (nitrate reductase) from Paracoccus denitrificans. Mol. Microbiol. 3:653-661.
- 138. Pavelka, M. S., L. F. Wright, and R. P. Silver. 1991. Identification of two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. J. Bacteriol. 173:4603–4610.
- Pugsley, A. P. 1993. The complete general protein secretory pathway in gram-negative bacteria. Microbiol. Rev. 57:50-108.
- 140. Pugsley, A. P., C. d'Enfert, I. Reyss, and M. G. Kornacker. 1990. Genetics of extracellular protein secretion by Gramnegative bacteria. Annu. Rev. Genet. 24:67–90.
- 141. Puvanesarajah, V., F. M. Schell, G. Stacey, C. J. Douglas, and E. W. Nester. 1985. Role for 2-linked-β-D-glucan in the virulence of Agrobacterium tumefaciens. J. Bacteriol. 164:102– 106.
- 142. Ralph, W. W., T. Webster, and T. F. Smith. 1987. A modified Chou and Fasman protein structure algorithm. Comp. Appl. Biosci. 3:211–216.
- 143. Ramseier, T. M., H. V. Winteler, and H. Hennecke. 1991. Discovery and sequence analysis of bacterial genes involved in the biogenesis of c-type cytochromes. J. Biol. Chem. 266: 7793-7803.
- 144. Reizer, J., A. Reizer, and M. H. Saier, Jr. 1992. A new subfamily of ABC-type transport systems catalyzing export of drugs and carbohydrates. Protein Sci. 1:1326–1332.
- 145. Riordan, J. R., J. M. Rommens, B.-S. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J.-L. Chou, M. L. Drumm, M. C. Iannuzzi, F. S. Collins, and L.-C. Tsui. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complimentary DNA. Science 245: 1066–1072.
- 146. Roberts, I. S., R. Mountford, N. High, D. Bitter-Suermann, K. Jann, K. N. Timmis, and G. J. Boulnois. 1986. Molecular cloning and analysis of genes for production of K5, K7, K12

MICROBIOL. REV.

and K92 capsular polysaccharides in *Escherichia coli*. J. Bacteriol. 168:1228–1233.

- 147. Ross, J. I., E. A. Eady, J. H. Cove, W. J. Cunliffe, S. Baumberg, and J. C. Wootton. 1990. Inducible erythromycin resistance in staphylococcus is encoded by a member of the ATP-binding transport super-gene family. Mol. Microbiol. 4:1207-1214.
- 148. Rossmann, M. G., A. Liljas, C. I. Branden, and L. J. Babaszak. 1975. Evolutionary and structural relationships among dehydrogenases, p. 62–103. *In P. D. Boyer (ed.)*, The proteins. Academic Press, Inc., New York.
- 149. Rosteck, P. R., Jr., P. A. Reynolds, and C. L. Hershberger. 1991. Homology between proteins controlling *Streptomyces fradiae* tylosin resistance and ATP-binding transport. Gene 102:27–32.
- 150. Salmond, G. P. C., and P. J. Reeves. 1993. Membrane traffic wardens and proteins secretion in Gram-negative bacteria. Trends Biochem. Sci. 18:7-12.
- 151. Schatz, P. J., and J. Beckwith. 1990. Genetic analysis of protein export in *Escherichia coli*. Annu. Rev. Genet. 24:215-248.
- 152. Scheu, A. K., A. Economou, G. F. Hong, S. Ghelani, A. W. B. Johnston, and J. A. Downie. 1992. Secretion of the *Rhizobium leguminosarum* nodulation protein NodO by haemolysin-type systems. Mol. Microbiol. 6:231–238.
- 153. Schnell, N., G. Engelke, J. Augustin, R. Rosenstein, V. Ungermann, F. Götz, and K.-D. Entian. 1992. Analysis of genes involved in the biosynthesis of lantibiotic epidermin. Eur. J. Biochem. 204:57-68.
- 154. Schnell, N., K.-D. Entian, F. Götz, T. Hörner, R. Kellner, and G. Jung. 1989. Structural gene isolation and prepeptide sequence of gallidermin, a new lanthionine containing antibiotic. FEMS Microbiol. Lett. 58:263–268.
- 155. Schnell, N., K.-D. Entian, U. Schneider, F. Götz, H. Zähner, R. Kellner, and G. Jung. 1988. Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. Nature (London) 333:276–278.
- 156. Schoner, B., M. Geistlich, P. R. Jr., R. N. Rao, E. Seno, P. Reynolds, K. Cox, S. Burgett, and C. Hershberger. 1992. Sequence similarity between macrolide-resistance determinants and ATP-binding transport proteins. Gene 115:93–96.
- 157. Schulein, R., I. Gentschev, H.-J. Mollenkopf, and W. Goebel. 1992. A topological model for the haemolysin translocator protein HlyD. Mol. Gen. Genet. 234:155–163.
- 158. Sebo, P., P. Glaser, H. Sakamoto, and A. Ullman 1991. Highlevel synthesis of active adenylate cyclase toxin of *Bordetella pertussis* in a reconstituted *Escherichia coli* system. Gene 104:19-24.
- 159. Sebo, P., and D. Ladant. Repeat sequences in the Bordetella pertussis adenylate cyclase toxin can be recognized as alternative carboxy-proximal secretion signals by the Escherichia coli  $\alpha$ -haemolysin translocator. Mol. Microbiol., in press.
- 160. Servos, J., E. Haase, and M. Brendel. 1993. Gene SNQ2 of Saccharomyces cerevisiae, which confers resistance to 4-nitroquinoline-N-oxide and other chemicals, encodes a 169 kDa protein homologous to ATP-dependent permeases. Mol. Gen. Genet. 236:214-218.
- Siegele, D. A., and R. Kolter. 1992. Life after Log. J. Bacteriol. 174:345–348.
- 162. Siegele, D. A., and R. Kolter. Isolation and characterization of an *E. coli* mutant unable to resume growth after starvation. Genes Dev., in press.
- 163. Silver, R. P., P. Annunziano, M. S. Pavelka, R. P. Pigeon, L. F. Wright, and D. E. Wunder. 1993. Genetic and molecular analysis of the polysialic acid gene cluster of *Escherichia coli* K1, p. 59–71. *In* J. Roth, U. Rutishauser, and F. A. Troy (ed.), Polysialic acid. Birkhauser Verlag, Basel.
- 164. Smith, A. N., G. J. Boulnois, and I. S. Roberts. 1990. Molecular analysis of the *Escherichia coli* K5 kps locus: identification and characterization of an inner-membrane capsular polysaccharide transport system. Mol. Microbiol. 4:1863–1869.
- 165. Smith, R. F., and T. F. Smith. 1990. Automatic generation of primary sequence patterns from sets of related protein sequences. Proc. Natl. Acad. Sci. USA 87:118–122.

- 166. Smith, R. F., and T. F. Smith. 1992. Pattern-induced multisequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for comparative protein modelling. Protein Eng. 5:35-41.
- 167. Spies, T., M. Bresnahan, S. Bahram, D. Arnold, G. Blanck, E. Mellins, D. Pious, and R. DeMars. 1990. A gene in the human major histocompatibility complex class II region controlling the class I antigen presentation pathway. Nature (London) 348:744-747.
- 168. Stanfield, S. W., L. Ielpi, D. O'Brochta, D. R. Helinski, and G. S. Ditta. 1988. The *ndvA* gene product of *Rhizobium meliloti* is required for  $\beta$ -(1-2)glucan production and has homology to the ATP-binding export protein HlyB. J. Bacteriol. 170: 3523–3530.
- 169. Stanley, P., V. Koronakis, and C. Hughes. 1991. Mutational analysis supports a role for multiple structural features in the C-terminal secretion signal of *Escherichia coli* haemolysin. Mol. Microbiol. 5:2391–2403.
- 170. Stoddard, G. W., J. P. Petzel, M. J. van Belkum, J. Kok, and L. L. McKay. 1992. Molecular analyses of the lactococcin A gene cluster from *Lactococcus lactis* subsp. *lactis* biovar diacetylactis WM4. Appl. Environ. Microbiol. 58:1952–1961.
- 171. Strathdee, C. A., and R. Y. C. Lo. 1989. Cloning, nucleotide sequence, and characterization of genes encoding the secretion function of the *Pasteurella haemolytica* leukotoxin determinant. J. Bacteriol. 171:916–928.
- 172. Suh, Y., and M. J. Benedik. 1992. Production of active Serratia marcescens metalloprotease from Escherichia coli by alphahemolysin HlyB and HlyD. J. Bacteriol. 174:2361–2366.
- 173. Thomas, W. D., S. P. Wagner, and R. A. Welch. 1992. A heterologous membrane protein domain fused to the C-terminal ATP-binding domain of HlyB can export *Escherichia coli* hemolysin. J. Bacteriol. **174**:6771–6779.
- 174. Thompson, S. A., L. L. Wang, A. West, and P. F. Sparling. 1993. *Neisseria meningitidis* produces iron-regulated proteins related to the RTX family of exoproteins. J. Bacteriol. 175: 811-818.
- 175. Tomasz, A., and J. L. Mosser. 1966. On the nature of the pneumococcal activator substance. Proc. Natl. Acad. Sci. USA 55:58-66.
- 176. Tommassen, J., A. Filloux, M. Bally, M. Murgier, and A. Lazdunski. 1992. Protein secretion in *Pseudomonas aeruginosa*. FEMS Microbiol. Rev. 103:73–90.
- 177. Tormo, A., M. Almirón, and R. Kolter. 1990. *surA*, an *Escherichia coli* gene essential for survival in stationary phase. J. Bacteriol. 172:4339-4347.
- 178. Trowsdale, J., I. Hanson, I. Mockridge, S. Beck, A. Townsend, and A. Kelly. 1990. Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters. Nature (London) 348:741-744.
- 178a. Ueda, K., K. Miyake, S. Horinouchi, and T. Beppu. 1993. A gene cluster involved in aerial mycelium formation in *Streptomyces griseus* encodes proteins similar to the response regulators of two-component regulatory systems and membrane translocators. J. Bacteriol. 175:2006–2016.
- 179. Van Belkum, M. J., B. J. Hayema, A. Geis, J. Kok, and G. Venema. 1989. Cloning of two bacteriocin genes from a lacto-coccus bacteriocin plasmid. Appl. Environ. Microbiol. 55: 1187–1191.

- Van Belkum, M. J., B. J. Hayema, R. E. Jeeninga, J. Kok, and G. Venema. 1991. Organization and nucleotide sequences of two lactococcal bacteriocin operons. Appl. Environ. Microbiol. 57:492-498.
- 180a.van der Ven, F. J. M., H. W. van der Hooven, C. W. Hilbers, and R. N. H. Konings. 1992. NMR studies of lantibiotics: the three-dimensional structure of nisin in aqueous solution, p. 435-447. *In R. James, C. Lazdunski, and F. Pattus (ed.), Bacteriocins, microcins, and lantibiotics. Springer Verlag KG, Heidelberg, Germany.*
- 181. Vazquez, M., O. Santana, and C. Quinto. 1993. The NodI and NodJ proteins from *Rhizobium* and *Bradyrhizobium* strains are similar to capsular polysaccharide secretion proteins from Gram-negative bacteria. Mol. Microbiol. 8:369–377.
- 182. Walker, J. E., M. Sarste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the  $\alpha$  and  $\beta$ -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1:945–951.
- Wandersman, C. 1989. Secretion, processing and activation of bacterial extracellular proteases. Mol. Microbiol. 3:1825–1831.
- Wandersman, C. 1992. Secretion across the bacterial outer membrane. Trends Genet. 8:317–322.
- 185. Wandersman, C., and P. Delepelaire. 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. Proc. Natl. Acad. Sci. USA 87:4776–4780.
- 186. Wandersman, C., P. Delepelaire, S. Letoffe, and M. Schwartz. 1987. Characterization of *Erwinia chrysanthemi* extracellular proteases: cloning and expression of the protease genes in *Escherichia coli*. J. Bacteriol. 169:5046–5053.
- 187. Wandersman, C., and S. Letoffe. 1993. Involvement of lipopolysaccharide in the secretion of *Escherichia coli* alphahemolysin and *Erwinia chrysanthemi* proteases. Mol. Microbiol. 7:141–150.
- 188. Wang, R., S. J. Seror, M. Blight, J. M. Pratt, J. K. Broome-Smith, and I. B. Holland. 1991. Analysis of the membrane organization of an Escherichia coli protein translocator, HlyB, a member of a large family of prokaryotic and eukaryotic surface transport proteins. J. Mol. Biol. 217:441-454.
- 189. Waters, V. L., and J. H. Crosa. 1991. Colicin V virulence plasmids. Microbiol. Rev. 55:437-450.
- 190. Welch, R. A., C. Forestier, A. Lobo, W. Thomas, and G. Rowe. 1992. The synthesis and function of the *Escherichia coli* hemolysin and related RTX-toxins. FEMS Microbiol. Immunol. 105:29-36.
- 191. Wickner, W., A. J. M. Driessen, and F.-U. Hartl. 1991. The enzymology of protein translocation across the *Escherichia coli* plasma membrane. Annu. Rev. Biochem. 60:101-124.
- 192. Wooten, J. C., and M. H. Drummond. 1989. The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. Protein Eng. 2:535–543.
- 193. Yang, C. C., and J. Konisky. 1984. Colicin V-treated Escherichia coli does not generate membrane potential. J. Bacteriol. 158:757-759.
- 194. Zevenhuizen, L. P. T., and A. R. W. van Neerven. 1983. (1,2)-Beta-D-glucan and acidic oligosaccharides produced by *Rhizobium meliloti*. Carbohydr. Res. 118:127–134.
- 195. Zevenhuizen, L. P. T. M., and H. J. Scholten-Koerselman. 1979. Surface carbohydrates of *Rhizobium*. I. Beta-1,2-glucans. Antonie Leeuwenhoek **45**:165–175.