Three Paradoxes of Ferric Enterobactin Uptake

Phillip E. Klebba*

Department of Chemistry & Biochemistry, University of Oklahoma, 620 Parrington Oval, Norman, OK 73019 USA, and Faculte de Medecine, Institut Necker Enfants Malades, 156, rue de Vaugirard 75015, Paris France . *Corresponding author: Tel: 405-325-4969, Fax: 405-325-6111, peklebba@ou.edu

Bacteria need iron for so many critical metabolic processes, including glycolysis, energy generation by electron transport, DNA synthesis, and defense against toxic reactive oxygen species, that the element is indispensable to their survival. Several decades ago this iron requirement was correlated to bacterial pathogenesis in animals and man, and research since then indelibly linked prokaryotic iron acquisition and infectious disease. Bacteria seek and acquire iron from their mammalian hosts, by secreting siderophores that capture the metal from iron containing proteins in animal tissues and by synthesizing elaborate cell envelope systems that transport either the bacterial ferric siderophores or the eucaryotic iron proteins themselves. Regardless of their method of iron accumulation, bacteria are susceptible to growth inhibition by iron deprivation, which, if it occurs in vivo, may prevent or reduce virulence. However, the multitude of specialized, sophisticated and efficient procaryotic systems to scavenge Fe⁺⁺⁺, combined with our limited knowledge of how they function, makes it difficult to use this strategy as a defense against pathogenesis. Hence the elucidation of the mechanism of iron transport through the outer membrane (OM) protein FepA, directly pertains to efforts against bacterial disease. The delineation of how the receptor protein recognizes and transports the native E. coli siderophore, ferric enterobactin (FeEnt) broadly relates to the other TonB-dependent iron acquisition systems of E. coli and other enteric bacteria, and to the discovery of new therapeutic strategies against bacterial pathogens.

In the past three years knowledge of ferric siderophore receptors exploded (Figure 1), from the

completion of the crystal structures of FepA, a ferric catecholate transporter, FhuA, a ferric hydroxamate FecA, the ferric citrate (FeCit) transporter. transporter, and the C-terminal domain of the protein that they require for functionality, TonB. Over many preceding years microbiologists, geneticists, molecular biologists and biochemists described the multiple protein components of cell envelope iron uptake systems, their energetic requirements, the dichotomy of beneficial and toxic ligands that enter the cell via their OM receptor proteins, the unique high-affinity nature of their uptake mechanism, their dependence on another cell envelope protein, TonB, channel-forming properties, and conformational dynamics in response to ligand binding. This research provided a conceptual foundation for the structural framework that the crystallography revealed.

The first siderophore receptor that was crystallized and solved, FepA, contained, as expected and previously demonstrated, the largest known \$barrel of the OM, covered on the exterior surface by large loops that bind ligands, and closed on the periplasmic side by its own N-terminus. In a fundamental sense receptors like FepA fulfill the definition of a porin: they contain a transmembrane pore, through which solutes pass into the cell. However, they are ligand-gate din that ferric siderophore binding stimulates conformational changes that activate siderophore internalization through the transmembrane channel. Furthermore, the TonB- and energy-dependence of their transport reaction distinguishes such ligand gated porins (LGP) from general and specific porins: they accumulate iron chelates against a concentration gradient, using cellular energy, and facilitated by the TonB-ExbB-ExbD complex in the inner membrane (IM). In contrast to the typical trimeric arrangement found in porins, FepA FhuA and FecA were isolated and crystallized as monomers, that contained two distinct domains: a C-terminal, 22-stranded antiparallel \$barrel (C-domain) that spans the outer membrane and projects the extracellular loops that function in ligand binding, and a globular N-terminal domain that folds into the barrel interior, blocking access to the periplasm (N-domain).

N-domain. The structurally distinct N-domain within the C-domain consists of "- and \$-structure and loops that rise to the top of the channel,

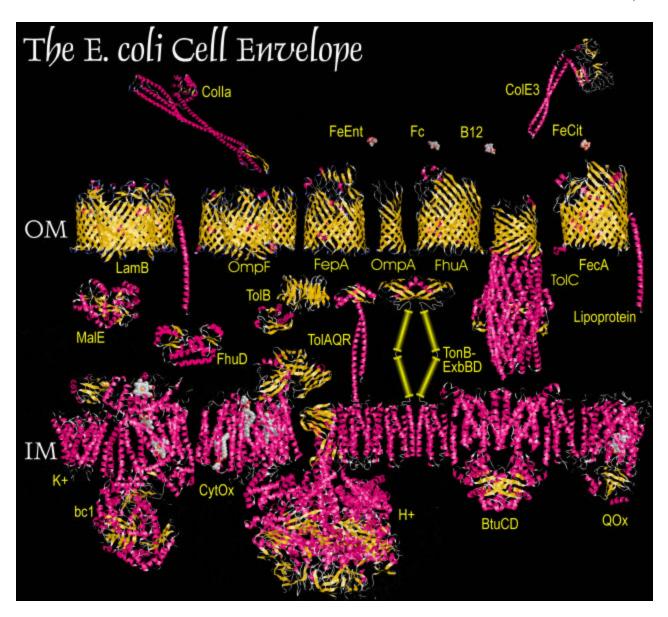


Figure 1. Proteins of the Gram-negative bacterial cell envelope. Molecular coordinates from the Protein Data Bank (http://www.rcsb.org/pdb/) and modeled using Rasmol 2.6.

directly beneath the ligand binding site, that provide a signaling pathway linking ligand recognition and transport. When ferrichrome (Fc) binds to FhuA or FeCit bind to FecA, residues in the loops undergo minor changes that propagate through the N-domain, changing the disposition of residues at the periplasmic interface of the OM. FepA was crystallized without full ligand occupancy, and comparable changes were not observed in its crystals. Although their overall topology is similar, differences

occur in the folding and composition of the N-domains of FepA, FhuA and FecA, that sometimes localizes analogous residues in different places. All the proteins contain four short, highly conserved sequences that assemble into four strands of a \$-sheet, and several short "-helices, or helical turns. Two loops project up from the "\$-structure toward the opening of the pore vestibule. However, the loops in the FepA, FhuA and FecA N-domains are distinctive: they fold differently, and those of FepA

contain a preponderance of Arg residues at the top; whereas FhuA contains aromatic residues in the same relative position. The different folding of the N domains achieves unique three-dimensional forms: that of FepA is more elongated, while those of FhuA and FecA are more compact.

C-domain. General and specific porins contain antiparallel, amphiphilic \$-sheets that circumscribe an aqueous, open transmembrane channel. Short reverse turns on their periplasmic surfaces, and large loops on their external surfaces, join the \$-strands within the sheet. FepA contains a comparable transmembrane \$-barrel, exclusively formed by its C-terminal 575 amino acids. The amphiphilicity of their component \$-strands, the nature of the loops and turns connecting them, and the delineation of position in the OM bilayer by aromatic residues at the internal and external interfaces, are conserved attributes among general, specific and ligand-gated porins. The most distinguishing differences between the barrel of FepA the other classes of porin are \$-strand length and number: the longest strands in OmpF and LamB are approximately 15 residues, but several of the LGP strands exceed twenty amino acids in length. The longer strands and large surface loops project the pore vestibule higher above the cell surface, which explains the antibody recognition of epitopes within the loops in live bacteria. The LGP barrels contain 22 strands, whereas general porins have 16, and sugarspecific porins 18. The increased number of \$strands creates a barrel with larger diameter, allowing passage of the larger siderophore ligands. Whereas the L3 (transverse) loops of general and specific porins fold inward and narrow the interior diameter of their channels, the N-domains of the LGP completely close their pores. This modified architecture of siderophore receptors demonstrates their distinctiveness among OM proteins, but also exemplifies further evolution on an existing theme: in other porins a structurally independent feature, the transverse loop, restricts channel permeability; in LGP, a novel globular assembly, the N-domain, fully regulates solute passage through the pore.

Interaction with FeEnt. Eleven loops encircle the channel on the cell surface, forming an exterior vestibule through which ferric siderophores enter. Some are expansive and participate in binding of FeEnt and the toxins that pass through FepA. The 11 loops of LGP are not homologous; they may dramatically differ in length. Both FepA and FhuA contain aromatic amino acids in the loops that

populate the mouth of the vestibule; FecA does not (Figure 2). In FepA this group of predominantly Tyr residues forms the initial adsorption site for ferric siderophores (6). Deeper in the FepA vestibule an abundance of basic residues group in a cluster at the top of the N-domain, presumably creating affinity for the triple-anionic catecholate siderophore.

FeEnt binding to FepA was thought to mainly involve the central region of the protein, because antibodies against surface epitopes in this region blocked FeEnt binding and transport. The dominant chemical properties of FeEnt, negative charge and aromaticity, suggested the involvement of basic and aromatic amino acids in recognition. These predictions were verified by site-directed mutagenesis: alanine substitutions for R286, R316 and K483, and Y260, Y272 and F329 impaired ligand binding.

Discrepancies in vitro and in vivo. Experiments with a purified, fluorescently-labeled Cys mutant protein initially revealed two kinetically distinguishable stages of FeEnt binding, intimating that the ligand moves between two distinct binding sites in the surface loops. After rapid adsorption ($k = 2 \times 10^{-2} \text{ s}^{-1}$) to the first site, FeEnt progresses more slowly ($k = 2 \times 10^{-3} \text{ s}^{-1}$) to a second site. The crystallographic data support the expectation of two potential sites in the vestibule, in that the FepA and FhuA crystals contained FeEnt and Fc in two different positions, corresponding to the proposed binding sites B1 and B2.

Conformational dynamics during ligand internalization is an inescapable feature of LGPmediated transport, because of the complete occlusion of their pores by the N-domain. Electron spin resonance (ESR) studies of nitroxides attached to four different sites (S271C, E280C, E310C; C493) in three different FepA loops, showed by three different methods (conventional, power saturation and time-domain ESR) that the purified protein undergoes loop movement when it binds FeEnt. Experiments with live E. coli expressing nitroxidelabeled residue E280C showed that additional, TonBand energy-dependent conformational changes occur during FeEnt internalization. Finally, results with both FepA (24)(25) and FecA (11) showed that in the absence of ligand, surface loops L7 (FepA) and L8 adopt an open conformation, that closes when the appropriate ferric siderophore binds. Nevertheless, the crystallographic descriptions of FhuA did not show differences in the disposition of its loops with and without bound ferrichrome.

Comparisons of equilibrium binding data derived from purified FepA, studied by extrinsic fluorescence $(K_d = 20 \text{ OM})$ and from live bacteria, studied by 59 FeEnt adsorption ($K_d = 0.2$ OM), showed a 100-fold difference in affinity of the siderophore-receptor interaction in the two conditions. The incongruity was even greater (250fold) in binding assays utilizing purified FepA, resolubilized from crystals, and the affinity of FeEnt for the isolated N-terminus of FepA is 20,000-fold lower (26). Dissimilar behavior of proteins may occur in different environments, but disparities at this level do not likely result from experimental variation or methodological differences: it is apparent that FepA exists in measurably different forms when resident in the OM bilayer, and when detergentsolubilized and purified. This difference was substantiated by measurement of the affinity of the FepA-FeEnt interaction in vivo, using fluorescent methodologies (7), and likely stems from alterations of FepA tertiary structure that occurs upon extraction from the OM bilayer. The crystallographic environment, replete with salts and/or precipitants, is considerably different from the native membraneous state, in which the loops of FepA exist in an open conformation (25)(7). This difference in environments perhaps also explains the observation that solubilization by non-ionic detergents and crystallizaton produced monomeric forms of both FepA and FhuA, while evidence exists that in vivo LGP are oligomeric or trimeric.

Mechanistic view of the FeEnt-FepA **interaction.** It is simplifying to consider transport of ferric enterobactin through FepA as a series of sequential steps, some of which are biochemically and genetically well defined, and some of which are comparatively obscure. The functions of the receptor's Nterminal globular domain, the TonB protein, and cellular energy fall into the latter category, creating three paradoxes of iron transport: how do solutes pass through a transmembrane \$barrel that is blocked by the globular N-terminal domain; how does TonB, a structurally simple protein that associates with the inner membrane, facilitate the transport activity of receptor proteins in the outer membrane; what are the energetic requirements of metal transporters like FepA, and how are they fulfilled?

What is known. The binding of FeEnt to FepA is a biphasic, high-affinity reaction. FeEnt initially adsorbs to aromatic and charged residues in the surface loops of FepA, in a site previously

designated B1. The reaction is specific, in that it is not subject to competitive inhibition by other siderophores, except those mimicing the structure of the FeEnt iron complex. After initial adsorption, concomitant with conformational changes in the loops, the iron complex moves within the vestibule to a second site, designated B2. These binding reactions are TonB- and energy-independent events, in that they occur with indistinguishable high affinity $(K_d = 0.2 \text{ nM})$ in tonB cells or tonB + cells that areenergy starved or poisoned. In FhuA, Fc binding induces movement of the TonB-box, on the periplasmic surface, to the center of the \$-barrel; whether such movements occur in FepA is not known, but similar movements were inferred from ligand binding to both FecA (11) and BtuB (20).

Ligand Selectivity. The ability of LGP to discriminate among different metal chelates is puzzling. FepA recognizes the metal center of the ferric catecholates it transports. FhuA interacts with ferrichrome in a site that complements the metal center of the chelate, lined with aromatic residues and defined by H-bonds from residues in the N-domain and surface loops. However, FhuA shows extremely broad recognition of hydroxamates, including the dihvdroxamate ferric rhodotorulate and ferrichrome analogs ferrirubin and ferrirhodin, stereoisomers with bulky substitutions to the iron center of the chelate. This specificity conflicts with the perfect complementarity between its ligand binding site and ferrichrome. In a lock-and-key model of binding, such a perfect structural match excludes siderophores with diverse structures at the iron center. Thus another means of binding occurs, akin to induced fit of the binding site to the siderophore, that accommodates molecules of different size or shape.

FepA manifests a more selective recognition pattern, accepting the tricatecholates FeEnt, FeTrencam, FeMecams and FeMyxochelin C, but rejecting the slightly different catecholates FeCorynebactin, FeAgro, and other analogs with chemical modifications to the catecholates around the iron. These results suggest a binding pocket restricted by size: all of the non-binding siderophores are larger, in one way or another, than FeEnt. Hence, although FhuA appears promiscuous in regard to ligand recognition, FepA seems opposite, fastidious to structural nuances in its ligands.

On one hand, the similar size and coordination geometry of the iron centers of ferric siderophores, which provide the major determinants,

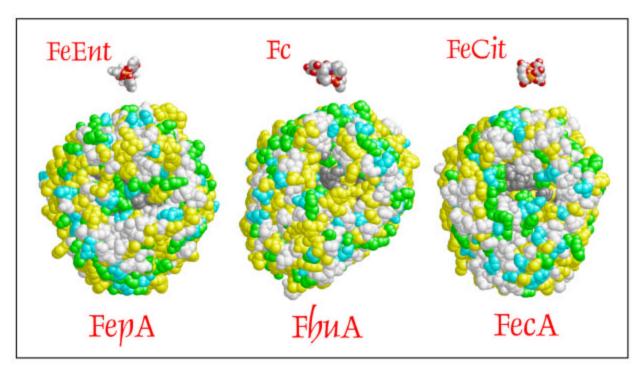


Figure 2. Comparison of the cell surface regions FhuA (left) FepA (center) and FecA (right). The three receptors are shown from a top view in space-filling format, Basic and acidic residues are green and cyan, respectively; aromatic and non-polar amino acids (Leu, Ile, Val, Met, Ala) are yellow. Note the more non-polar surfaces of FhuA and FepA, which recognize hydrophobic siderophores Fc and FeEnt, relative to those of FecA, which recognizes FeCit.

make it difficult to explain their selective adsorption to particular OM receptor proteins, whose surface structures are themselves at least superficially similar (Figure 2). On the other hand, for the most part each ferric siderophore has unique chemical properties. Consider for example, FeEnt, Fc, and FeCit. The former has a fully aromatic character, in that the metal chelated. with three-fold symmetry hexacoordinate geometry, by three catecholate groups that impart a net charge around its iron center of -3. Ferrichrome is not aromatic; its three hydroxamate chelation groups, derived from Hhydroxy ornithine, form an electrically neutral complex with Fe3⁺. The latter compound, FeCit, is also a non-aromatic, neutral complex like Fc, and it achieves these properties by dimerization: 2 citrate moieties complex two Fe 3⁺ atoms. Thus the three ferric siderophores are chemically distinct from one another, as are the majority of ferric siderophore classes.

Nevertheless, most siderophores have a

common hydrophobicity that undoubtedly plays a role in their binding to receptor proteins, and a question that remains is how does ferric siderophore discrimination proceed in the microenvironment? One possibility is that their initial binding occurs by nonspecific hydrophobic interactions with the non-polar or aromatic amino acids in the surface loops. That is, hydrophobic side chains in the surface loops of siderophore receptor proteins may non-specifically sequester siderophores, in a comparable manner to their extraction and purification from aqueous solution by organic solvents. In this case the selection of a correct ligand and rejection of others only occurs later, at a subsequent stage that precedes internalization. A second alternative is initial discrimination of the correct siderophore in the first stage of its adsorption process. The latter mechanism has more biochemical and physiological logic. If each receptor protein initially adsorbed several or many different classes of ferric siderophores with significant affinity, and only rejected inappropriate ligands at the secondary stage of the binding process, then the act of ligand selection would assume futile inefficiency in any environment populated with diverse organisms and siderophores (as for example, the vertebrate gut).

The ligand-free open state *in vivo* is not likely a static conformation with loops spread like the petals of a flower. Instead, as was illustrated by crystallography, the loops are flexible in solution,

	1	ConB-Box		N\$1	NL1
NgoFetA .	AENNANVA	L <mark>D</mark> TV <mark>TV</mark> KGDRQGSKIRTN		VTLQ.QKDESTAT	DMRELL.KEEPSI
BpuFepA .	QTATQIHDPS	SQ <mark>VQQ<mark>M</mark>ATVQVLGTAEEE</mark>		IK.E.SL <mark>GVSVITAEEI</mark> A	RR.PPTN <mark>DL</mark> S <mark>DL</mark> I.RRE <mark>PGV</mark>
XciFepA.		S <mark>LDTV</mark> H <mark>VTA</mark> AQIA		R.QA.LGTSVITAEDIE	RR.PPVNDLSQLL.RSMPGV
PaeFepA	AGQGDGSVIE	L <mark>GE QTVV</mark> ATAQEE		.TTK.QAP.GVSIITAEDIA	K <mark>R</mark> .PPSN <mark>DL</mark> SQ <mark>I</mark> I. <mark>R</mark> TM <mark>PGV</mark>
SenIroN .	AAESTDD	DNG <mark>ETIVV</mark> ESTAEQV	••••••	LK.QQP.GVSIITRDDI.	QKNPPVNDLADII.RKMPGV
StyFepA .	EDKTDSAALT	N <mark>EDTIVVTA</mark> AQQN	••••••	LQAP. <mark>GVSTITADEI</mark> .	RKNPPARDVSEII.RTMPGV
EcoFepA .	QEPTDTPVS	H <mark>DDTIVVTA</mark> AEQN	•••••••	LQAP. <mark>GVSTITADEI</mark> .	RKNPVARDVSKII.RTMPGV
EcoCir		DO <mark>ETMVVTA</mark> SSVEQN	•••••••	LKD.APAS ISVITQEDL.	Q <mark>R</mark> KP <mark>V QN . LKDV</mark> L . <mark>K</mark> EV <mark>PGV</mark>
EcoFecA .	LQQLLDGSGLQVKPLGNNSWTLEPAPAPK	E <mark>D</mark> A <mark>LTVV</mark> GDWLGDAREND		VFEHAGA.RDVIRREDFA	<mark>K</mark> TGATTM <mark>REV</mark> L.NRI <mark>PGV</mark>
EcoBtuB .		SP <mark>DTLVVTA</mark> NRFEQPR	ST	VLAPTTVVTRQDI.	DRWQSTSVNDVL.RRLPGV
EcoIutA .		D <mark>DETFVVSA</mark> NRSNR	Т	VAEMAQTTWVIEN.AE	LEQQ <mark>I</mark> QGGK <mark>ELKD</mark> ALAQLI <mark>PGL</mark>
EcoFhuA .		K <mark>EDTITVTA</mark> APAPQES	AWGPAATIAARQSATGTKTD	TPIQK.VPQS <mark>ISVVTAEEM</mark> A	L <mark>H</mark> QPKS <mark>VKE</mark> AL.SYT <mark>PGV</mark>
EcoFhuE .		T <mark>EETVIV</mark> EGSATAPDDGEND	.YSVTSTSAGTKMQ	MT.QRDIPQS <mark>VTIVS</mark> .QQRM	EDQQ <mark>L</mark> QT <mark>LGEV</mark> M.ENTL <mark>GI</mark>
		DETIVVTA		G <mark>VSVIT</mark> A EDI A	RKNPVDLRDII.RRMPGV
NmeLbpA .	AQAGGATPDAAQTQS	. <mark>L</mark> KE <mark>ITV</mark> R <mark>A</mark> AKV.		RRSKEATGL <mark>GKIVKT</mark> S <mark>ETL</mark> N	. <mark>K</mark> EQ <mark>V</mark> LG <mark>IRDL</mark> T.RYD <mark>PGV</mark>
	AENVQAGQAQEKQ				
NmeTbpA		. <mark>LDTIQV</mark> K <mark>A</mark> KKQK		RRDNEVTGL <mark>GKLVKS</mark> S <mark>DTLS</mark>	.KEQVLNIRDLT.RYDPGI
	AETQSIKDTKE.AISSEVDTQSTEDSE				
AplTbpA .	~~~~~~~~~~~~~ QAV	Q <mark>L</mark> ND <mark>V</mark> Y <mark>VT</mark> GTKKKA		KKENEVTGL <mark>GKVVKT</mark> P <mark>DTLS</mark>	.KEQVLGIRDLT.RYDPGI
PhaTbpA .	ENKKIEENNDLA	~ V <mark>LDE</mark> VI <mark>VT</mark> .ESHYAH.ERQ.		NEVTGLGKVVKNYHEMS	.KNQILGIRDLT.RYDPGI
McaTbpA I	ALSLGLLNITOVALANTTADKAEATDKTNLVV	V <mark>LD</mark> ETV <mark>VTA</mark> KKÑ.		rkanevtgl <mark>gkvvkt</mark> a <mark>eti</mark> n	.KEOVLNIRDLT.RYDPGI
MaaT.bp*		DAI <mark>TVT</mark> AAPSAPV.		RFDTDVTGL <mark>GKTVKT</mark> A <mark>DTL</mark> A	<mark>K</mark> EOVOGIRDLV.RYETGV
исанира.	ENTOTDANSDAKDIKTPVVYL				~ ~
мсапоря .	ENTQTDANSDARDTRTPVVYL	LDTITVTA		GKVVKTAETLS	.KEQVLNIRDLT .RYDPGI
мсанора .	ENTQTDANSDAKDTKTPVVYLI			GKVVKTAETLS	.KEQVLNIRDLT.RYDPGI
нсапори .	NII.		NL2	GKVVKTAETLS N\$3	.KEQVLNIRDLT.RYDPGI N\$4
-	DF.GG. GNGTSOFLTLRGMG.ONS.	N\$2	NL2	N\$3	N\$4
- NgoFetA	NL1	N\$2VDIK.VDNAYS		N\$3 IVDPA.L <mark>V</mark> KV <mark>V</mark> SVQKG <mark>A</mark>	
- NgoFetA BpuFepA	NLI DF.GGGNGTSOFLTLRGMG.QNS.	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG	WNGDRDTRGDTNW	N\$3 IVDPA.L <mark>V</mark> KV <mark>V</mark> SVQKG <mark>A</mark> VPAEEVERIEVIRGP <mark>AAA</mark>	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI
NgoFetA BpuFepA XciFepA	DF.GGGNGTSOFLTLRGMG.QNS.NLTGNSASGARGN.SRQ.VDIRGMGPEN	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG	WN <mark>GDRDTRGDT</mark> NW RS <mark>GER</mark> NTRGDT <mark>N</mark> W	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVLRGPAAA	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI
NgoFetA BpuFepA XciFepA PaeFepA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG	WNGDRDTRGDTNW RSGERNTRGDT <mark>N</mark> W WRGERDSRGDTNW	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKQ.AGAI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN	DF.GGENGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSSGQRGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLVLIDGVPVTSRNSVRYS .TLILIDGKPVTSRNSVRYS	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTAW	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGNGAAG.GVVNIITKQ.AGAI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGOYGN.NRQ.IDLRGMGPEN NLTGNSSSGORGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLVLIDGVPVTSRNSVRYS .TLILIDGKPVTSRNSVRYS	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTAW	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGNGAAG.GVVNIITKQ.AGAI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRYS .TLILIDGKPVSSRNSVRLG .TLILIDGKPVSSRNSVRQG YTLILVDGKRVNSRNAV	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTAW WRGERDT RGDTSW WRGERDT RGDTSW FR. HNDF DLNW	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPAAA IPVDSIERIEVVRGPMSS	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKGSG.I
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoCir	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSSGQRGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILIVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .TLILIDGKPVSSRNSVRLG .TLILIDGKPVSSRNSVRQG YTLILVDGKRVNSRNAV	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTAW WRGERDT RGDTSW WRGERDT RGDTSW FR. HNDF DLNW	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPAAA IPVDSIERIEVVRGPMSS	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKGSG.I
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoCir EcoFecA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRYS .TLILIDGKPVSSRNSVRLG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSRNSVRQG .TLILVDGKRVNSRNAV RSTVLMDGIPVPFAPYG	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTAW WRGERDT RGDTSW FR. HNDF DLNW QPQLSL	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVLRGPAAA VPPE.IERIEVLRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGAV	N\$4GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTII RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKGSG.I LYGSDALG.GVVNIITKK.IGQI RYGPQSVG.GVVNFVTRA.IPQI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoCir EcoFecA EcoBtuB	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTSNSASGTRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASS	N\$2VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGVPVTSRNSVRYG .TLVLIDGVPVTSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILVDGKRVNSRNAV RSTVLMDG IPVPFAPYG	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW FR. HNDF DLNW QPQLSL SGSADLSQ:	N\$3 IVDPA.LVKVVSVQKGA V.PAEEVERIEVIRGPAAA V.PADQVERIEVIRGPAAA V.PPEMVERIEMIRGPAAA V.PPEMIERIEVIRGPAAA V.PPE.IERIEVIRGPAAA V.PPE.IERIEVIRGPARA I.PVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGAV F.PIALVQRVEYIRGPRSA	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTDI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKGSG.I LYGSDALG.GVVNIITKK.IGQI RYGPQSVG.GVVNFVTRA.IPQI VYGSDAIG.GVVNIITTR.DEPG
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoGir EcoFecA EcoBtuB EcoIutA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNS ASGARGN.SRQ.VDIRGMGPEN NLTGNS ASGQYGN.NRQ.IDLRGMGPEN NLTGNS SSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.ANN.FGIRGLNPRLASI DITQNGGSGQL.SSIFIRGTNAS	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILVDGKRVNSRNAVSTVLMDG IPVPFAPYGHVLVLIDGVRLNIAGVLVVLVDGVRLNISSRTD	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGNGAGG.GVVNIITKK.IGQI RYGPQSVG.GVVNFVTRA.IPQI RYGPQSVG.GVVNIITTR.DEPC
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoGir EcoFecA EcoBtuB EcoIutA EcoFhuA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNGGSGQL.SSIFLRGTNAS DVSSRSRTNYGMNVRGRP	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .TLILIDGKPVSSRNSVRQG	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGSDALG.GVVNIITKK.IGQI RYGSDALG.GVVNIITTR.DEPCI LYGSGTG.GVVNIITTR.CPCI LYGGGSTG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoGir EcoFecA EcoBtuB EcoIutA EcoFhuA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SS.IFIRGTN.AS DVSSRSRTNYGMNVRGRP SVGTR.GASNT.YDH.LLIRGFAAEGQS	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILVDGKRVNSRNAV R STVIMDG IPV LVVLVDGVRLNLAGVLVVLVDGVRLNSSRTD	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGSDALG.GVVNIITKK.IGQI RYGSDALG.GVVNIITTR.DEPCI LYGSGTG.GVVNIITTR.CPCI LYGGGSTG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuE	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SS.IFIRGTNAS DVSSRSRTNY.GMNVRGRP SVGTR.GASNT.YDH.LIIRGFAAEGQS SKSQADSDRALYYSRGFQI	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLVLIDGVPVTSRNSVRYG .TLILIDGKPVTSRNSVRLG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILVDGKVNSRNAV RSTVLMDG IPVPFAPYGLVVLVDGVRLNSSRTD .LVVLVDGVRLNSSRTD .LVVLVDGVRLNSSRTD .DNYMV.DGIPTYFES TLILVDG	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW WRGERDT RGDTSW FR HNDF DLNW SGSADLSQ SRQ LDS I RWNLGDALS	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PADI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKR.IDEPQI LYGGGSTG.GLINIVTKK.GQPI LYGKGSTG.GLINIVTKK.GQPI LYGKSSPG.GLINWYSKR.PTTI LYGKSSPG.GLINWYSKR.PTTI MTGTGNPS.AAINWYKHATSRI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuE	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.TDLRGMGPEN NLTGNSSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN QLT.NE.GD.NRKGYSLRGL.DS.S. SAPENNGTGSHDL.ANN.FGLRGLNPRLASI DITQNG.GSGQL.SS.IFLRGTN.AS DVSSRSRTNY.GMNVRGRPSVGTR.GASNT.YDH.LLLRGFAAEGQS SKSQADSDRALYYSRGFQT. RGNYNRQ.IDLRGM	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNAV R STVLMDG IPVPFAPYGLVVLVDGVRLNLAGVLVVLVDGVRLNSSRTD .DNYMV.DGIPTYFES DNYMV.DGIPTYFES TLILVDG .VAVSVDGVAQIQAFTVQGS	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW FR.HNDF DLNW SGSADLSQ SRQ.LDS I RWNLGDALS NW LSGYGGR GGSGAI NE	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVLRGPAAA VPPE.IERIEVLRGPAAA VPPE.IERIEVLRGPAAA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPFN.MHHIEVIFGATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSS	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKK.GQPI LYGGGSTG.GVVNIITKK.GQPI LYGKSSPG.GLINIVTKK.GQPI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuA NgoTbpA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNS ASGARGN.SRQ.VDLRGMGPEN NLTGNS ASGQYGN.NRQ.TDLRGMGPEN NLTGNS SSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN NLTGNS TSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSLRGL.DS.S' SAPENNGTGSHDL.ANN.FGLRGLNPRLASI DITQNG.GSGQL.SSIFLRGTNAS DVSSRSRTNYGMNVRGRP SVGTRGASNT.YDH.LLLRGFAAEGQS SKSQADSDRALYYSRGFQL' RGNYNRQ.IDLRGM AVVEQGNGA.SG.GYSLRGVDKNR.	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLILUDGKPVTSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILVDGKVNSRNAV RSTVLMDG IPVPFAPYGLVVLVDGVRLNLAGVLVVLVDGVRLNSSRTD .QNNYL.NGLKL.GNFYNDNYMV.DGIPTYFES TLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGLAQIQSYTAQAA	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEJERIEVVRGPMSS APVSLGNMDAIDVVRGGAV FPIALVQVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSSSI	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKR.IPQI VYGSDAIG.GVVNIITTR.DEPC LYGGGSTG.GLINIVTKK.GQPI LYGKSPG.GLINMVSKR.PTTI MTGTGNPS.AAINMVRKHATSRI YGSGAAG.GVVNIITR DHGSGALG.GAVAFRTKEAAI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuA NgoTbpA	DF.GGGNGTSOFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDIRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSSGQRGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGINPRLAS! DITQNG.GSGQL.SS.IFIRGTN.AS DVSSRSRTNY.GMNVRGRP SVGTRGASNT.YDH.LIIRGFAAEGQS SKSQADSDRALYYSRGFQI: RGNYNRQ.IDLRGMAVVEQGRGA.SG.GYSIRGVDKNR	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVGSRDAVRMG .TLILVDGKPVSSRNSVRYG .TLVLIDGVPVTSRNSVRYG .TLVLIDGKPVSSRNSVRYG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNAV R STVLMDG IPVPFAPYGHVLVLIDGVRLNLAGVLVVLVDGVRLNSSRTD .QNNYL.NGLKL.GNFYNDNYMV.DGIPTYFESTLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGLAQIQSYTAQAA .VSLTVDGVSQIQSYTAQAA	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW FR HNDF DLNW SGSADLSQ SRQ LDS I RWNLGDALSS WWLLSGYGGR GGSGAI NE LGGTR T AGSSGAI NE LGGTR T AGSSGAI NE	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEJERIEVVRGPAS IPVDSIERIEVVRGPAS APVSLGNMDAIDVVRGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSS: IEYENVKAVEISKGSNSV: IEYENVKAVEISKGSNSV	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKR.IPQI VYGSDAIG.GVVNIITTR.DEPCI LYGSGSTG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINMVSKR.PTTI MYGTGNPS.AAINMVRRATSRI YGSGAAG.GVVNIITR DHGSGALG.GAVAFRTKEAAI EQGSGALA.GSVAFQTKTADI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoGir EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuA HinTbpA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SS.IFLRGTNAS DVSSRSRTNYGMNVRGRP SVGTR.GASNT.YDH.LLIRGFAAEGQS SKSQADSDRALYYSRGFQII RGNYNRQ.IDLRGM AVVEQGRGA.SS.GYSIRGMDKNR AVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .VSTUMDG IPV.PFAPYGLVVLVDGVRLNLAGVLVVLVDGVRLNLAGVLVVLVDGVRLNSSRTD .CQNNYL.NGLKL.GNFYNDNYMV.DGIPTYFESTLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VALLVDGLPQTQSYVVQSP .GLALDGLPQIQSYVSQYSRS	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSS: IEYENVKAVEISKGSNSS IEYENVKAVEISKGSSSS .EYENLRSIQISKGASSSEF	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGSDALG.GVVNIITKK.IGQI RYGSDALG.GVVNIITTR.DEPO LYGSDAIG.GVVNIITTR.DEPO LYGSDAIG.GVVNIITTR.GVPI LYGKSSPG.GLLNMVSKR.PTTI MTGTGNPS.AAINMVRKHATSRI YGSGAAG.GVVNIITR DHGSGALG.GSVVNIITR EQGSGALA.GSVAFTKEAAI EVGNGALA.GSVAFQTKTADI EVGNGALA.GSVAFQTKTADI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoGir EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuA HinTbpA HinTbpA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SS.IFLRGTNAS DVSSRSRTNYGMNVRGRP SVGTR.GASNT.YDH.LLIRGFAAEGQS SKSQADSDRALYYSRGFQII RGNYNRQ.IDLRGM AVVEQGRGA.SS.GYSIRGMDKNR AVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .VSTUMDG IPV.PFAPYGLVVLVDGVRLNLAGVLVVLVDGVRLNLAGVLVVLVDGVRLNSSRTD .CQNNYL.NGLKL.GNFYNDNYMV.DGIPTYFESTLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VALLVDGLPQTQSYVVQSP .GLALDGLPQIQSYVSQYSRS	MNGDRDT	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGPMSS APVSLGNMDAIDVVRGGAV FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSS: IEYENVKAVEISKGSNSS IEYENVKAVEISKGSSSS .EYENLRSIQISKGASSSEF	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKK.IGQI RYGSDALG.GVVNIITKK.IGQI RYGSDALG.GVVNIITTR.DEPO LYGSDAIG.GVVNIITTR.DEPO LYGSDAIG.GVVNIITTR.GVPI LYGKSSPG.GLLNMVSKR.PTTI MTGTGNPS.AAINMVRKHATSRI YGSGAAG.GVVNIITR DHGSGALG.GSVVNIITR EQGSGALA.GSVAFTKEAAI EVGNGALA.GSVAFQTKTADI EVGNGALA.GSVAFQTKTADI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuE NmeLbpA NgoTbpA NmeTbpA HinTbpA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.TDLRGMGPEN NLTGNSSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN NLTGNSTSGQRGN.NRQ.TDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SSIFLRGTNAS DVSSRSRTNYGMNVRGRP SVGTR.GASNT.YDH.LLIRGFAAEGQS SKSQADSDRALYYSRGFQI RGNYNRQ.TDLRGM AVVEQGRGA.SS.GYSIRGMDKNR AVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.TT.GYSRGWDRNRV	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYS .TLILIDGKPVTSRNSVRLG .LVVLVDGVTSRNSVRLG .CTLILIDGKPVTSRNSVRLG .CTLILVDGVTLNLAGVLVVLVDGVTLNLAGVLVVLVDGVTLNSSRTD .CQNNYL.NGLKL.GNFYNDNYMV.DGIPTYFESTLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VSLTVDGLAQIQSYVSQYSRS .VSLLVDGLPQAHSYHTLGS	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW WRGERDT RGDTSW WRGERDT RGDTSW FR. HNDF DLNW SGSADLSQ SRQ LDS I RWNLGDALSI NW LSGYGGR GGSGAINE LGGTR. T AGSSGAINE LGGTR. T AGSSGAINE LV. AR. S GYSGTGAINE SSGAINE I DANGGAINE	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVVRGPMSS APVSLGNMDAIDVVRGGGVS FPIALVQRVEYIRGPRSA DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVLRGP IEYENISTVEIDKGAGSS: IEYENVKAVEISKGSNSS IEYENVKAVEISKGGSSS SEYENLRSIQISKGASSSE IEYENIRSIELSKGASSSE IEYENIRSIELSKGASSS	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTGI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKG.GSI RYGNGAAG.GVVNIITKKGSG.I LYGSDALG.GVVNIITKK.IGQI RYGSDALG.GVVNIITTR.DEPO LYGSDAIG.GVVNIITTR.DEPO LYGSGAIG.GVVNIITTR.GVED LYGKSSPG.GLLN.MVSKR.PTTI MTGTGNPS.AAIN.MVRKHATSRI YGSGAAG.GVVNIITR DHGSGALG.GSVAFTKEAAI EVGNGALA.GSVAFQTKTADI EVGNGALA.GSVAFQTKTADI GSGSLG.GSVQFTKSAAI GSGSLG.GSVQFTKSAAI
NgoFetA BpuFepA XciFepA PaeFepA SenIroN StyFepA EcoFepA EcoFecA EcoBtuB EcoIutA EcoFhuA EcoFhuA HinTbpA HinTbpA PhaTbpA	DF.GGGNGTSQFLTLRGMG.QNS. NLTGNSASGARGN.SRQ.VDLRGMGPEN NLTGNSASGQYGN.NRQ.IDLRGMGPEN NLTGNSSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN NLTGNSTSGQRGN.NRQ.IDLRGMGPEN QLT.NE.GD.NRKGVSIRGL.DS.S' SAPENNGTGSHDL.AMN.FGIRGLNPRLASI DITQNG.GSGQL.SS.IFLRGTNAS DVSSRSRTNYGMNVRGRP SVGTR.GASNT.YDH.LLIRGFAAEGQS SKSQADSDRALYYSRGFQII RGNYNRQ.IDLRGM AVVEQGRGA.SS.GYSIRGMDKNR AVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR SVVEQGRGA.SS.GYSIRGMDKNR	N\$2 VDIK.VDNAYS .TLILIDGKPVTSRNAVRYG .TLILVDGKPVSSRNAVRYG .TLILVDGKPVSSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVTSRNSVRYG .TLILIDGKPVTSRNSVRQG .TLILIDGKPVTSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .TLILIDGKPVSSRNSVRQG .VTLILVDGVSLNLAGVLVVLVDGVRLNSSRTD .GQNNYL.NGLKL.GNFYNDNYMV.DGIPTYFESTLILVDG .VAVSVDGVAQIQAFTVQGS .VSLTVDGVAQIQSYTAQAA .VSLTVDGLAQIQSYTAQAA .VALLVDGLPQIQSYVSQYSRS .VSLLVDGLPQIQSYVSQYSRS .VSLLVDGLPQAHSYHTLGS .VAVLVDGINQAQHYALQGP	WNGDRDT RGDTNW RSGERNT RGDTNW WRGERDS RGDTNW WRGERDT RGDTNW WRGERDT RGDTSW WRGERDT RGDTSW WRGERDT RGDTSW WRGERDT RGDTSW WRGERDT RGDTSW FR HNDF DLNW SGSADLSQ SRQ LDS I DAVI RWNLGDALS: NW LSGYGGR GGSGAI NE LGGTR T AGSSGAI NE LGGTR T AGSSGAI NE LU AR S GYSGTGAI NE SSGAINE I DANGGAI NE UVAGKN Y AA GGAI NE	N\$3 IVDPA.LVKVVSVQKGA VPAEEVERIEVIRGPAAA VPADAVERIEVIRGPAAA VPADQVERIEVIRGPAAA VPPEMVERIEMIRGPAAA VPPEMIERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPAAA VPPE.IERIEVIRGPARA IPVDSIERIEVVRGGRAV DPFN.MHHIEVIFG.ATS DPY.MLERAEIMRGPVSV DMA.LFERVEVVRGATGLI VPAEMVERIEVIRGP IEYENISTVEIDKGAGSS: IEYENVKAVEISKGSNSS IEYENVKAVEISKGSSSSE IEYENVKAVEISKGASSSE IEYENIRSIQISKGASSSE IEYENIRSIELSKGASSA IEYENVRSVEISKGANSS	N\$4 GSASAGIGATNGAIIAKTVDAQI RYGSGAMG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PADI RYGSGASG.GVVNIITKR.PTDI RYGNGAAG.GVVNIITKR.PTNI RYGNGAAG.GVVNIITKK.GSG.I RYGNGAAG.GVVNIITKK.GSG.I LYGSDALG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITKK.IGQI RYGPQSVG.GVVNIITRK.IGQI LYGSDAIG.GVVNIITRK.IGQI LYGSGAIG.GVVNIITRK.GVDI LYGKGSTG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKSSPG.GLINIVTKK.GQPI LYGKGSTG.GLINIVTKK.GQPI LYGKGSTG.GLINIVTKK.GQPI LYGKGSTG.GLINIVTKK.GQPI LYGKGSTG.GVVNIITR DHGSGALG.GVVNIITR DHGSGALG.GSVVNIITR DHGSGALG.GSVVNIITR DHGSGALG.GSVVNIITR DHGSGALG.GSVVNIITK LEVGNGALA.GSVTFQTKTADI LEVGNGALA.GSVTFQTKSAAI GSGSLG.GSVQFTTKEVSD LEVGSGALS.GSVVAFVTKTADI

Figure 3. Sequence alignment of N-termini of FepA homologs Basic residues are colored green, acidic residues blue, aromatic residues magenta. Significantly conserved regions among the 21 proteins are highlighted yellow, and the consensus sequences below indicate moderately (black), highly (maroon) and very highly (red) conserved residue

imparting an overall form and motion akin to the tentacles of a the sea anemone. During its diffusion near the membrane surface FeEnt encounters charged {K483; (25)} and aromatic {F329, Y272; residues at the loop extremities that entrap the siderophore in a network of non-covalent interactions that constitute the first binding stage. As the multiple determinants, within multiple loops, converge on the ferric siderophore, the natural affinities of the individual association reactions close the loops around metal complex, in effect creating the secondary interactions that occur with charged and aromatic residues deep within the now vestibule, W101, Y260 and R316. Hence the loops of the barrel ultimately select the correct ferric siderophore (24), and the binding reactions require neither energy nor TonB to reach completion, or maximal affinity (5). Other ferric siderophores do not mistakenly adsorb to FepA because the chemistry of their metal complexes do not properly configure with the appropriate side chains in the loops of FepA. An implicit benefit of this selection mechanism is that the initial binding sites do not become erroneously occupied, and blocked, by encounters with inappropriate metal complexes.

What is unknown. The sub-reactions of FeEnt uptake that occur subsequent to binding are less resolved. The ferric siderophore begins the transport phase localized at the top of the N-domain, and through an unknown sequence of energy- and TonB-dependent events that involve further conformational changes in the loops (5), and unavoidably, movement of or in the NDomain as well, it traverses FepA channel and enters the periplasm. Three preeminent questions remain about the mechanism of metal transport, that derive from three paradoxes of existing data:

- 1. How does the N-domain regulate pore activity?
 - 2. What is the function of TonB?
 - 3. How is metal transport energized?

The crystal structures of FepA, FhuA, FecA and the Cterminal domains of TonB and TolA completely changed the study of ferric siderophore transport, superannuating the structural guessing games that preceded them. Nevertheless, mechanistic uncertainties persevere, as illustrated by the unexpected disposition of the N-domains within LGP channels. The latter two conundrums are historical (circa 1970) artifacts, that still confound ferric siderophore transport mechanisms.

Paradox 1: Transport through a closed channel: blockage of the transmembrane \$-barrel by the N-domain. The position of the N-domain inside

the C-domain leaves no opening, gap, or pore through which FeEnt may pass, insisting that structural changes must occur in FepA during transport. FeEnt binds to FepA with a sub-nanomolar K_d, that translates into a dissociation half-life of over a minute, and this calculation conflicts with the receptor's experimentally observed, 20 second turnover time. Lastly, LGP monomers apparently bind only a single molecule at a time, and transport it against a their concentration gradient. so uptake thermodynamics differ from those of general or specific porins, which transport by mass action. Hence internalization of FeEnt requires a driving force. These points argue for protein conformational changes that undermine the affinity of the siderophorereceptor binding interaction, create a pathway to the periplasm, and propel the metal complex through the

The N-domain, composed of approximately 150 amino acids that fold into a compact "\$-structure, lodges into the periplasmic outlet of the barrel domain. Protein sequence and structure analyzes reveal several important features of the globular Nterminus. It is predicated on a 4-stranded \$-sheet that is remarkably similar among them (Figure 3). comparisons among such homologs of FepA, some of which transport FeEnt and some of which transport other ferric siderophores, lactoferrin or transferrin. show the conservation of individual \$ strands within the N-domain sheet, but also a variety of primarily basic residues that distribute on its surface, at the aqueous interface with the barrel walls (Figure 4). The conservation of residues that occurs in the barrel and N-terminal domains of FepA and its structural relatives, relative to the diversity seen in their surface loops, suggests that once bound, siderophores pass through LGP channels by a common mechanism that involves the N-domain. The similarities further establish that the mechanism of iron acquisition from eukaryotic iron binding proteins, by the transferrin and lactoferrin binding proteins of Gram-negative bacteria, is fundamentally similar to that of ferric siderophore receptors.

The identification of common amino acids in the transporters of *different* ligands is the primary consideration underlying the search for mechanistically important residues in LGP. FepA, Cir, FecA, BtuB, IutA, FhuA and FhuE (Fig. 3) recognize two catecholate, one carboxylate, one corrinoid, and three hydroxamate metal complexes. Three of the chelates are uncharged, one has a net charge of -1, and three have net charges of -3. Therefore, conserved amino acids among the seven do

P.E.Klebba

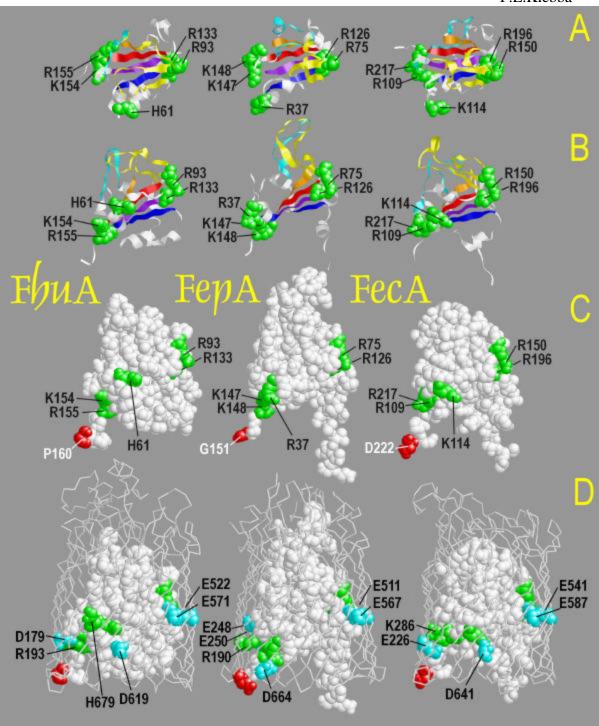


Figure 4. N-domain structure, and proposed Ion-Pair interactions. Top (**A**) and side (**B**) views of the \$-sheets within the N-domains of FhuA, FepA and FecA, shown in ribbon format. The colors of the individual strands within the sheet correspond to those depicted in Fig. 3. Conserved basic residues (green) project from the sheet, and are shown in space filling format. These amino acids reside on the surface of the N-domains (**C**), distributed around the conserved \$-sheet, within ion-pair interaction distance to acidic (cyan) and basic residues (green) on the barrel wall (**D**; white, shown in backbone format). The last residues of the N-domains, that connect to the first \$-strand of the barrel, are colored red.

not likely derive from ligand recognition properties, but they may originate from shared mechanistic features. Secondly, because different receptors transport ligands with different efficiencies, in any individual transporter particular residues may be more or less conserved than others. Thirdly, aligned residues in the sequences of the solved proteins do not always locate to comparable positions in their tertiary structures. In their N-domains, for example, diverse positions in sequence may fold to homologous locations in tertiary structure (FepA 147, 148, FhuA 154, 155 vs FecA 109, 196). Similarly, in the barrel, non-homologous amino acids sometimes localize to comparable sites (e.g., FepA E248, E250 vs FhuA E163, D179). These considerations suggest that mechanistic residues may show less than identity in aligned LGP sequences (Fig.3).

Ion-pairs. Beneath the vestibule a common feature appears in FepA, FhuA and FecA: ionic interactions between the N-domain and the \$ barrel wall. Basic residues in the N-domain and acidic and basic residues on the barrel appropriately converge in positions to form ionic bonds that stabilize the N-domain within the C-domain (Fig 4). The basic residues that exist on the N-domain surface are highly conserved in sequence, and derive from regions of structure within or at the extremities of the \$-sheet strands; these pair with Asp and Glu residues on the interior barrel walls, also at equivalent positions in all three transporters.

The ion pairs are best observed by rendering the barrel transparent and viewing residues of interest from the exterior (Fig 4). In FepA, on the side of the protein where the N-domain connects to the barrel (the "hinge" side), basic residues in the N-domain K147, K148) exist within ionic bond distance (<3.5 Å) to two acidic residues (E248, E250) on the barrel wall. On the opposite side of the globular domain (the "lock" side) other potential ion pairs exist, joining R75 and R126 to E511 and E567. These two sets of apposed charges portray the N-domain as a door, with a hinge and a lock. Analogous amino acids exist in comparable locations within FhuA and FecA alignment. Certain of the charged residues are the

alignment. Certain of the charged residues are the most conserved residues in LGP sequence. On the mechanistic side, they reside on the exterior surface of the N-domain and internal surface of the barrel, matched with residues of opposite charge; any meaningful changes in N-domain structure likely require the dissolution of these interactions.

The location, composition and arrangement

of stabilizing ion-pairs insinuate a mechanism for motion of the N-domain. The susceptibility of FeEnt transport to PMF inhibitors raises the possibility that protons, routed to the barrel interior, disrupt the ionic bonds between these residues. If the aqueous milieu within the pore drops to a pH below 4, then protonation of acidic side chains on the inside of the barrel will eliminate the salt bridges. The existence of additional basic residues on the hinge-side of the barrel wall intimates that in such conditions charge repulsions will occur with basic residues on the Ndomain surface, causing movement, perhaps expulsion, of the N-domain from the pore. So its structure suggests that the N-domain acts as a door, hinged to the barrel on one side and locked closed by salt bridges around its circumference. A decrease in pH in the barrel may both unlock the door and actuate an electrostatic force that initiates its opening.

Models of channel gating. The crystallographic data from FepA, FhuA and FecA resolved their structural organization, but not their transport mechanisms. It is worthwhile to consider the implications of two mechanistic extremes of the transport process, "sequential" and "concerted" transport reactions.

Sequential transport. Ample precedent exists for the stepwise movement of small molecules through membrane channels, including the transport of ions through the acetylcholine receptor, the potassium and chloride channels, bacteriorhodopsin and the proton ATPase, and the passage of small solutes through general porins and sugars through specific FeEnt passage through FepA may occur by similar sequential transport, that passes the acidic siderophore along a series of basic residues located on the interior of the \$-barrel, and/or within the Ndomain. Although LGP did not contain any pores or gaps in their interiors that might allow passage of a ferric siderophore, variations from these structures in vivo, combined with conformational changes, may transiently open a path to the periplasm. Maltodextrins traverse maltoporin through a very small pore, of 6 Å diameter at its constriction point, so narrow that it forces de-solvation of the sugars during transit. Ferric siderophores are considerably larger than a hexose, implying that a transient pore through FepA must acquire a minimum diameter of 15 - 20 Å. The existence of successive sites through the channel domain, with increasing affinity for the solute, is another prerequisite of a the process. Potential basic candidate residues align across the FepA channel, but

the high initial affinity of FeEnt binding in the surface loops creates a criterion of avidity that subsequent sites must supercede. Furthermore, general and specific porin transport systems are driven by mass action, a crucial difference between passage of those solutes and ferric siderophores, which are present in small concentrations and actively accumulated: this requirement constitutes the primary argument against a sequential transport mechanism for FepA. Thus in a sequential transport process the input of energy must accomplish several independent conformational rearrangements in FepA, the disruption of the initial binding sites, and the simultaneous creation or exposure of subsequent binding sites of increasing affinity within an appropriately-sized, nascent channel, that lead the ferric siderophore through the protein interior to the periplasm. The ability of FhuA to internalize a variety of ferric siderophores of with properties divergent chemical and (ferrichrome, ferrirhubin, ferrirhodin), and also the antibiotic rifamycin, is another argument against such a sequential transport mechanism, because it presupposes that the mechanistic binding sites within the channel are themselves highly promiscuous.

Concerted transport. The N-terminal domain is a unique protein fold among all other solved protein structures, and it singularity raises the possibility that it also functions in an unprecedented way: an all-or-none transition from a structure that occludes the FepA pore, to a structure that promotes movement of FeEnt into the periplasm. The short \$sheet, "-helices and connecting loops within the Ndomain may lend themselves to such a rearrangement According to this view the binding of FeEnt "loads" the N-domain into an activated form, that is then recognized and "triggered" to transport by cellular At least two kinds of conformational rearrangements may accomplish transport, a global alteration in the N-domain that reduces its size and narrows its shape within the barrel, or ejection of the entire N-domain from the barrel. Either action achieves continuity with the periplasm, and accomplishes transport if the surface loops simultaneously close to prevent back diffusion. However, it is difficult to envision a rearrangement of the N-domain to diminish its already densely and I will not consider this compacted shape, mechanism further.

At the solution of the FepA and FhuA crystal structures, structuralists argued against the notion of N-domain exit from the \$-barrel, because of the existence of over 50 potential hydrogen bonds that

presumably hold it in place (3, 10, 19), between its surface residues and residues on the barrel walls. However, on both theoretical and experimental bases, as originally cited and reported in Scott et al., (24) we endorse the concept of Ndomain expulsion during transport of FeEnt through FepA. The novel structure of the N-domain may portend a novel mechanism, and no experimental data yet exist to assess the feasibility of globular domain exit from the barrel. Presumably it enters the pore, because the \$-barrel correctly assembles in its absence. Evidence exists that the presence of the N-terminus in the channel also optimizes or activates the motions of the surface loops for binding, and charge interactions between the Nand C-domains may mediate this action. Pore closure in the absence of ligand constitutes a final function of the globular domain, that prevents influx of natural detergents that disrupt the inner membrane. Such maintains an absolute physiological requirement, the selective OM permeability barrier.

The apriori bases for the N-domain-exit (Ball and Chain) mechanism are substantial. regarding the noted intra-protein Hoond objection, discussed above, several points are germane. interior of the barrel is hydrophilic, a general porin characteristic that persevered in siderophore receptors. Not only charged amino acids, but also polar, uncharged residues cover the interior walls of FepA, FhuA and FecA. So in biological environments the channel is likely water-filled. This layer of water separating the N- and C-domains, which was seen in the FepA crystal structure, but perhaps inadequately relative to the aqueous environment in vivo, showed that many of the potential H-bonds between the Ndomain and the barrel are bridged by water molecules. The existence of 50 intra-protein H-bonds does not preclude exit of the Nterminus from the barrel, as long as majority are re-formed as the opening occurs. The reformation of the H-bonds may occur between the side-chains and water, with other side chains, or H-bond acceptors/donors on the siderophore, without any energetic penalty. strength of H-bonds in proteins is 1.5-1.8 kCal/Mol, so even the energetic equivalent of hydrolysis of a single ATP (7.5 kCal/M) is sufficient to account for the breaking of 5 H-bonds without reformation. present, the energy requirements of the transport are unknown; according to considerations 10 ATP are required to compensate the breakage of 50 H-bonds, although this upper limit is not likely necessary because the N-domain departs into the aqueous periplasm, where H-bonds may immediately re-form with water. Those that do not re-form perhaps justify the need for energy to internalize the solute. Finally, unstated in the argument against N-domain exit was the fact that a sequential mechanism also requires extensive dissolution of hydrogen bonds, in this case between the strands of the \$-sheet, and presumably also, the N-domain surface and the barrel walls. The dissolution of a 4-stranded \$-sheet will require a significant energy contribution, and no obvious mechanism exists to achieve it.

The strong conservation of charged residues on the interacting surfaces of the N and C-domains suggest that not only the amino acids, but also the surfaces themselves, are intimately involved in the transport process. The model mechanism is simple and readily rationalized. Protons enter the barrel from the periplasmic side, neutralize negative charges on the barrel wall and unmask positive charges that either initiate expulsion of the N-domain from the channel. TonB may participate in these reactions in several ways at several stages. Movement of the N-domain out creates negative hydrostatic pressure in the pore, that pulls the surface loops closed, altering the ligand binding site to release the ferric siderophore into the channel, now continuous with the periplasmic space.

The strongest data supporting this so-called Ball-and-Chain mechanism are hybrid receptor proteins that encode the N-domain of FepA, and the Cdomain of FhuA (FepNFhu\$) (24). These constructs, which correctly bind and transport ferrichrome, demonstrate that the primary recognition specificity of the receptor protein resides in its loops, and that the Ndomain plays a non-specific role in the transport process. Secondly, genetic constructs that completely delete the N-terminus, leaving only the empty \$barrels of FepA and FhuA, still showed residual, TonB-dependent activity transport for appropriate ferric siderophores. These data intimate that the empty barrel is in fact a transport intermediate. The suggestion was made that the activity of such constructs originates from complementation with a cryptic FepN-domain in the host bacterium (27). If these data are valid, they verify the feasibility of the independent movement of the N-domain into the pore. N-domain exit that also achieves loop closure conceptually solves several kinetic, thermodynamic and physiological problems. As a channel to the periplas m arises, loop closure may simultaneously collapse the ligand-receptor binding interaction and prevent solute escape to the exterior, eliminating the need for a driving force of transport. The ferric

siderophore will enter the periplasm by diffusion, perhaps enhanced by low-level affinity for the N-domain itself, and accumulate within by adsorption to its binding protein.

TonB: facilitation of Paradox 2: siderophore transport by a minor cell envelope component. In spite of intense interest, the 239 residue E. coli TonB protein remains one of the mysteries of the bacterial cell envelope. connection to energy metabolism, postulated since the nearly simultaneous discovery of the TonB- and energy-dependence of iron transport by Wang and Newton, still has an inherent logic. Only a few exceptions exist to the almost ubiquitous need for TonB mediation of LGP transport; these include the penetration of T5 through FhuA, Ecolicins through BtuB, and cloacin DF13 through IutA. The uptake of all metal complexes requires both TonB and energy, leading to the assumption that the two prerequisites are fulfilled by one molecular entity. I do not object to this conclusion on a conceptual basis, I object to the idea that the concept is proven (16)(15): although an abundance of inferential data exist for the connection between TonB and energy, and for specific physical interactions between TonB and ferric siderophore receptors, no unequivocal proof exists for either Nevertheless, considerable progress has theory. occurred in the understanding of at least TonB, including the important demonstration that its Cterminus exists in close proximity to OM proteins in the cell envelope, as evinced by the extensive crosslinking studies of Postle and colleagues.

Initial genetic and sequence data identified an 7-residue, moderately conserved approximately sequence, called the TonB-box, near the N-terminus of siderophore receptors, that was proposed physically interact with TonB residue O160. sequence conservation itself is only tangential to a possible relationship with TonB, as illustrated by the fact that crystallography revealed proposed TonBboxes 1.5 and 2, described by Kadner(21), as two \$strands of the Ndomain. In fact, no true identity exists in the TonB-boxes of LGP (Fig 3) that might portend a specific biochemical interaction with a single site on a single protein. This realization argues against it's proposed binding to site Q160 in TonB. The crystal structure of the TonB C-domain reinforced this inference, because its overall structure is smooth and contains no obvious binding clefts, and because Q160 is apparently blocked from access to the inner OM surface by the 70-residue C-domain dimer. Later TonB-crosslinking studies, either by activation with

formaldehyde or by introduction of single Cys residues in the TonB-box region of the receptor at or near TonB residue 160, were interpreted as evidence for the proposed role of TonB as an energy transducer in metal transport reactions (18)(4). But proximity is not energy transduction. The experiments established that TonB is close enough to OM proteins in the cell envelope to form crosslinks with them. Similar interactions occur between TonB and FepA proteins lacking a TonB-box (23), and the non-TonB-dependent OM protein OmpA (23)(14); TonB associates with the OM even when siderophore receptors are not present within it (14), underderscoring the lack of specificity in the affiliation. Furthermore, the TonB-dependent uptake of ferric siderophores by receptor proteins lacking the complete N-domain, including the TonB-box (2)(24) questioned the proposed significance of this region to iron transport¹.

If the TonB-box of FepA does not specifically interact with TonB, then what is its function in FeEnt uptake? The region shows moderate conservation among iron transporters of Gramnegative bacteria, though less than that observed for the \$-strands in the N-domain \$-sheet. The allegedly specific interaction between the two proteins, which received apparent support from the crystallographic finding that ligand binding relocates the TonB-box regions of FhuA and FecA from a position adjacent to the barrel wall, to the center of the channel, neglect the existence of between 6 and 33 (82 if one includes the special case of FecA) residues upstream, that manifest an overall negative charge, flexibility, and according to the model, are juxtaposed between the TonB C terminus and the TonB-box. Although the movement of the TonB-box in response to ligand binding was seen as a signal of occupancy to TonB, it is alternatively possible that the interaction between the TonB box and the barrel wall is another association that physically holds the N-terminus in place, through strong hydrophobic bonds (Fig. 5). In FepA (and FecA) this association assumes nearly perfect complementarity between the non-polar surfaces of the TonB-box and the barrel wall. This idea concurs with

¹Subsequent work, nevertheless, raised the possibility that the *E. coli* strains harboring the N-terminal deletions of FhuA and FepA contain cryptic fragments of the FhuA and/or FhuA N-termini, that complemented the empty \$-barrels by inserting within them and facilitating TonB-dependent colicin sensitivity (27).

the only moderate overall conservation of TonB-box sequences: their true specificity may not lie in interaction with a single, energy-transducing protein, but rather for residues on their individual barrel walls. It's relevant that in FepA and FecA this interaction with the barrel also physically shields the ion pair at the lock site (in FepA, between K75-R126 and E511-E567). Movement of the TonB box to the center of the channel permits access to this site. According to this model, ligand binding breaks the association between the TonB-box and the barrel wall, not as a signal to TonB, but as a means of releasing the N-domain to swing out of the channel.

The biochemical function of TonB is the second major uncertainty of ferric enterobactin transport through FepA. The possibility exists that current perceptions of its role in transport facilitation, especially its ability to mechanically transmit energy, are fundamentally incorrect. In spite of the intuitive the relationship between TonB and energy transfer, at present no experimental data directly implicates TonB in energy metabolism. Although purportedly involved in PMF-facilitation of transport (4, 18), TonB is not known to either create a proton gradient or to utilize one; although postulated to induce inner and outer membrane fusion, or to physically jump from the inner membrane to the outer membrane, carrying with it potential energy stored in a unique conformation, neither the origin of the protein's ability to leave one membrane and enter another, nor the nature of the energy-transducing conformation, are demonstrated. The most significant conceptual problem with current ideas is that they do not presuppose well demonstrated biochemical mechanisms. To summarize the major findings implicating TonB in an OM-associated activity: (i) various genetic experiments that show partial suppression between mutations in siderophore receptors and mutations in TonB, suggesting physical contact between the proteins; (ii) several studies with crosslinking reagents indicate the proximity of TonB to proteins in the inner and outer membranes; (iii) TonB is an absolute requirement of OM iron transport, and such systems also require energy, for the apriori reason that they concentrate solutes against a concentration gradient, and the aposteriori reason that energy starvation and energy poisons eliminate iron transport. However, none of these observations and correlations constitute evidence that TonB transduces energy in the cell envelope. This is not to say that the participation of TonB in the energetic aspects of OM transport is inconceivable, but rather, that the conclusion that it is an energy transducer by some kind

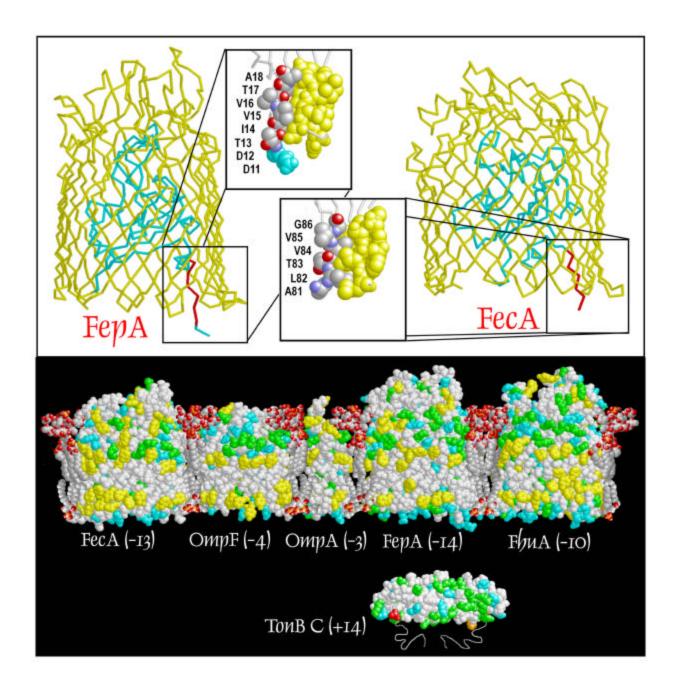


Figure 5. The TonB-box and TonB. (Top) A backbone representation of the N-termini of FepA and FecA (cyan), shown within their respective \$-barrels (yellow), illustrates the disposition of the TonB-box (red) region. In both protein, the TonB-box (inset: seen in space filling form with CPK colors) packs tightly against the barrel wall, held by interactions between hydrophobic residues on both surfaces. (**Bottom**) The overall negative charge of the periplasmic surface of the OM, from phospholipid head groups (CPK colors) and acidic residues of OM proteins (cyan), likely explains the non-specific affinity of the TonB C-terminus, which shows an overall positive charge as a result of the presence of a preponderance of basic residues (green). The first solved residue of the domain (165) is red, and the last solved residue (238) is orange.

of mechanical process, based on existing crosslinking and genetic data, is inherently flawed. For example, FepA chemically crosslinks to OmpA and OmpF (25), but this does not imply that these proteins participate in the FepA transport process.

If TonB does not itself charge with energy, diffuse across the periplasm, and discharge the energy to proteins in the OM (18), or achieve some approximation of this action, then what is its function in the cell envelope? Along these lines, new data are relevant.

- 1. TonB is present at low levels in the cell envelope, and forms a dimer (9), making the number of active, TonB-containing assemblies approximately 150 500 copies per cell (14).
- 2. As proposed (16) the TonB C-terminus forms a \$"\$ structure (9), that structurally

resembles the C-terminus of TolA (28), a periplasmic protein that associates with PAL, a peptidoglycan-associated lipoprotein. Although the complete structure of TonB is unknown, its overall domain organization bears similarity to that of TolA. The latter protein, which associates with TolQ and TolR, performs a structural role in the cell envelope, and its integrity and ability to crosslink to peptidoglycan-associated lipoprotein (PAL) are apparently PMF-dependent (8). The protein was recently implicated in LPS biosynthesis (12).

3. TonB's reported affinity for OM proteins extends not only to LGP, but to the major OM protein OmpA, and lipoprotein as well. Immobilized OmpA adsorbs the TonB C-terminus from solution (23), and OmpA and Lpp are two of the cell envelope proteins that may crosslink to TonB after formaldehyde activation (Figure 6). Finally, TonB associates with the OM even in the absence of any siderophore receptors therein (14). This finding reiterates the nonspecificity of the

associations that TonB experiences in the cell envelope, which was previously seen in

immunoblots of crosslinking reactions involving TonB, but was interpreted as a specific, ligand-potentiated affiliation between TonB and siderophore receptor proteins. Nevertheless under the conditions employed to study this phenomenon, TonB crosslinks to at least 15 cell envelope proteins {(27); Fig 6}.

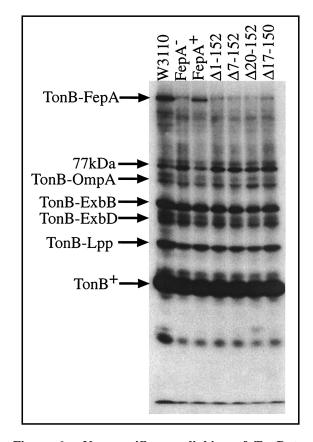


Figure 6. Non-specific crosslinking of TonB to other cell envelope proteins. Immunoblot analysis of formaldehyde-generated crosslinking, developed with anti-TonB serum. (From (27) reprinted with permission of the American Society for Microbiology). Approximately 50% of the total TonB present in the sample crosslinked to 15-18 cell envelope proteins, including lipoprotein and OmpA.

4. The TonB-C terminus is necessary for its activity. Expression of the cloned C-terminal 69 residues of TonB inhibits the activity, and the isolated C-terminus shows the tendency to spontaneously insert into phosphatidyl choline liposomes (23).

From these and previous findings, a picture of the cell envelope that emerges portrays associations of OM proteins in complexes that include or are nucleated by the TolAQR, Pal and Lpp proteins, or the TonB, ExbBD and OM proteins. In the former case, existing evidence links the TolA system to the structural

integrity of the cell envelope and LPS secretion; in the latter case, evidence links the TonB system to OM metal transporters. Together these data suggest that both TonB and TolA span from the IM, where they associate with ExbBD and TolQR, respectively, to the OM, where they associate with integral or peripheral proteins, including lipoproteins, OmpA, FepA, etc. The exact natures of these affiliations are unknown, but presumably involve non-covalent interactions with either the periplasmic interfaces of OM proteins, or the lipid bilayer itself. Thus the notion again arises that TonB (and TolA) bridge the two bilayers across the periplasm. A theoretical requirement likely exists for molecular trafficing between the two membranes, at least for biosynthetic reasons, and although the exact mechanisms of neither protein nor lipid insertion into the OM are known, evidence of zones of adhesion between inner and outer membranes does exist. Furthermore, studies of LPS biosynthesis revealed that export to the OM is blocked in TolA mutants (12). These important data impart a previously lacking functionality to the inter-membrane assembly that TolA participates in: it acts in the movement of LPS molecules from the IM to the OM.

It is inaccurate and potentially misleading to imply that TonB-mediated energy transduction was already demonstrated, or that its mechanism is understood. For example, the statement that "upon forming a complex with an outer membrane receptor, TonB releases stored energy, possibly in the form of mechanical force, and assumes the discharged conformation," (11) is not fact, but speculation. From the same data other consistent structural and mechanistic models may arise. One of these, for example, permanently anchors the N-termini of TonB and TolA in the IM by their transmembrane hydrophobic helices, and non-covalently, nonspecifically associates their Ctermini with the OM. This affiliation with the OM is not transient nor stimulated by energy, it is chemical in nature, involving ionic or hydrophobic bonds.

Although the FepA, FhuA and FecA \$-barrels contain many Arg, Lys, His, Asp and Glu residues, they distribute with the same general pattern in all three proteins: the exterior surfaces of the vestibules are primarily basic, and the periplasmic surfaces of the barrels are acidic (Figure 5). The TonB C-domain, on the other hand, has a basic surface, suggesting its association with the acidic periplasmic surfaces of the lipids and proteins of the OM. In the case of the TolA C-terminus, the associations may primarily occur with peptidoglycan-associated, OM-

imbedded lipoproteins. In this way TonB and TolA, complexed with their accessories ExbBD and TolQR, respectively, may span the bridge the cell envelope. Stoichiometric discrepancies (14, 16) and the much reduced fluidity of the OM bilayer make it unlikely that ferric siderophore

receptors localize at these proposed transport zones containing the TonB and TolA. However, preferential association with TonB is unnecessary in the model, because the proposed inter-membrane bridges have inherent random mobility, by virtue of their residence in the fluid IM bilayer, and their non-specific interactions with the internal surface of the OM. Such protein-mediated membrane connections may perform a variety of physiological functions, including biosynthesis (suggested by the requirement of TolA in the LPS export process) and metal transport (illustrated by the requirement for TonB in the FeEnt uptake reaction). In this context, one of their physiological roles is to provide a structural framework that may enable, in the case TolA, a pathway for passage of the strongly hydrophilic LPS O-antigen through a membrane bilayer, and in the case of TonB, a pathway for energy transfer that allows siderophore-mediated iron uptake. This postulate does not, however, resolves neither the nature nor the mode of delivery of the intra-membrane energetic

Paradox 3: Active transport in a membrane that cannot sustain an ion gradient. In light of the biochemical attributes of metal transport, low extracellular concentrations of ferric siderophores that bind to OM receptor proteins with high affinity, a convincing logic exists for their active transport. How the cell accomplishes this feat in the OM, which contains more than 10^5 10 Å holes, is a formidable Certainly an unusual, undiscovered energetic system exists that acts on ligand-bound receptor proteins, stimulating them to ligand internalization. In this regard, it is difficult conceive of TonB as a molecule that transduces the energy. Relative to other proteins involved in energy metabolism, TonB manifests stark differences. It is a small protein with undistinguished structural features (the sole exception is the central Glu-Pro, Lys-Pro rod region, that is presumably dispensable to TonB function), present at low amounts in the cell envelope. The usual pathways of energy production and utilization, through oxidation of carbon sources, reduction of sequential electron carriers that pump protons, and dissipation of PMF to generate ATP, involve large multiprotein complexes that

contain chromophores. The F_0F_1 ATPase, a relevant prototypic membrane protein that utilizes PMF to phosphorylate ADP, is thought to exist as a complex of 22 subunits of 8 different proteins. So if TonB-ExbBD delivers energy to the OM, it does so by a unique, novel mechanism. Therein lies the fascination with the energy transfer process, but also the need for conclusive proof.

Bioenergetics. Only a modest amount of data exists on the energetics of metal acquisition. Much of what is known originates from only a few papers on transport of vitamin B12 (1), FhuA (13) and FeEnt (22). In the case of vitamin B12, uptake of the metal complex by wild type E. coli is insensitive to Inhibition of B12 uptake by cyanide only occurred if the bacteria were devoid of an ATPsynthase, suggesting that even if electron transport stops (and thereby also the normal production of PMF), transport still continues if ATP is utilized to generate a proton gradient. This result was observed in two separate sets of experiments, that considered both overall uptake to the cytoplasm, and uptake across the OM (1). The latter experiments also found susceptibility of the OM transport stage to CCCP, but that cyanide actually stimulated the B12 uptake process. For ferric siderophore transporters, however, a different pattern of inhibition occurs, in that their activities are blocked by cyanide, and also inhibitors of phosphorylation. Hancock and Braun (13) reached this conclusion when studying the irreversible adsorption of bacteriophages T1 and N80 to FhuA. Furthermore, like FhuA activity, inhibitors of electron transport and phosphorylation block the uptake of FeEnt by E. coli, as do agents that deplete These initial experiments did not PMF (22). specifically isolate the OM transport stage of the ferric siderophore, only the overall uptake process into the cytoplasm. But experiments with an in vivo fluorescence system that focused on FeEnt uptake through FepA corroborated the inhibitory effect of both cyanide and arsenate, as well as susceptibility to PMF-depletion (7).

These data portray differences among the energy requirements of OM metal transporters, hinting that PMF may not play an exclusive central role. The main evidence for the energization of siderophore receptors by PMF derives from Bradbeer's studies of the isolated OM transport phase of B12 uptake: these data were extrapolated to iron uptake systems, under the assumption that they functioned identically. However, Bradbeer's characterizations of the cyanide-independence of BtuB reveals a critical difference to

the FeEnt acquisition system. This discrepancy is unexpected for two transport proteins that are so similar in structure, but other differences also exist between them. For example, the transport of B12 by BtuB requires TonB, while the penetration of E colicins through BtuB as requires the Tol system. In the case of FepA, the uptake of both FeEnt and colicins B and D is TonB-dependent.

Functional insight from sequence. Homology exists between ExbB and ExbD and two proteins of the bacterial flagellar system, MotA and MotB (17). The exact role of the former proteins, which presumably exist in complex with TonB, is uncertain, but the latter two proteins are cytoplasmic membrane components of the flagellar motor that form a proton-conducting, MotA₄MobB₂ multimer; eight such complexes surround the flagellar rotor, and the MotAB complex is envisioned as a stator that torque proton conduction-driven imparts by conformational changes (17). The underlying explanation for the relatedness of MotAB to ExbBD is sequence homology in proposed transmembrane strands 3 and 4 of MotA, and proposed transmembrane strands 2 and 3 of ExbB. The two potential hydrophobic helices manifest 20% identity but extensive homology along their length. Furthermore, mutation of a conserved proline (P713) between MotA helices 3 and 4, presumably located at the cytoplasmic interface, has strong effects on the flagellar rotation, ExbB contains a comparable conserved, functionally important Pro (P141) between its suggested helices 2 and 3. These and other similarities raise the possibility of a rotational mechanism of TonB action, in which proton passage through ExbBD turns TonB in the periplasm. On the other hand, the apparent stoichiometry of the E. coli TonB:ExbB:ExbD complex (1:8:2; (14)) differs from that of MotA₄MobB₂, and also unlike the Ton system, the rotational ability of the flagellar motor originates from a conglomeration of many proteins that create a rotor housed within a proteinaceous architecture that spans the bacterial cell envelope. No such architecture is known to exist and facilitate TonB action. Nevertheless, the relatedness between Mot and Exb is considerable, making the likelihood biochemical similarity probable.

As a conclusion, I emphasize the consistency of alternative interpretations of existing data on TonB and energy transduction, and postulate two such contrasting mechanisms. These originate from the following findings on ferric siderophore transport: (i) ligand binding releases the TonB-box from its normal

association with the interior of the \$-barrel wall and relocalizes the negatively charged N-terminus; (ii) the N-domain is held in the barrel at least in part by a group of ionic bonds between its surface and the barrel wall, and evidence exists that the N-domain may exit the \$-barrel as FeEnt traverses the channel; (iii) the Cterminus of TonB, which is required for its functionality, has a general affinity for the OM and its proteins, from ionic interactons between its predominantly basic surface, and the largely acidic inner surface of the OM lipids and proteins; (iv) TonB and TolA show structural similarity in their C-termini, and the latter protein also spans the periplasm and is needed in LPS biogenesis; (v) The sequence relatedness of MotAB and ExbBD suggest that the latter protein complex may cause the rotation of TonB in the IM.

A variety of interpretations may reconcile these observations, among which I will consider two (Figure 7). Model 1. Metal transport by rotational motion of TonB. According to this view TonB, ExbB and ExbD form a PMF-utilizing rotational complex in the IM. TonB is the rotor, and the dissipation of the proton gradient across the IM, through a proton channel within ExbBD, the stator, promotes rotation. In light of the general affinity of the TonB--C-domain for OM bilayer, the model suggest that it crawls, twirls, or spins across the periplasmic surface of the OM, facilitating transport in the process. If, for example, the negatively charged residues upstream of the TonB box of ligand-bound receptors are sufficiently attracted to, or bind to the Cdomain, then its motion may physically pull their Ndomain from the channel, accomplishing transport in the process. The rotational motion may also provide a mechanism of random movement across the inner OM surface. Encounters with ligand-free receptors will not interfere with the process, because their Ndomains remain locked in place by the interactions of the TonB-box and ion pairs with the barrel wall. Exclusive energization by PMF is not a prerequisite of this

proposal, as the rotatory mechanism is conceivably driven by other energy sources, perhaps explaining the susceptibility of FeEnt transport to inhibitors of phosphorylation and electron transport. Structural continuities between IM and OM In this postulate proteins that span the periplasm, including TonB and TolA, create or isolate a localized zone in which the energized state of the IM may directly impact upon the OM. Such inter-membrane connections may form a tunnel, analogous to that of TolC, between the two bilayers, which directs protons to the periplasmic surface of the OM. Whereas TolC anchors in the OM and floats above the surface of the IM, a TonB-TolA tunnel may anchor in the IM, and move along the underside of the OM. In this case the involvement of protons in the activation of transport is fundamental to the mechanism. Alternatively, the structure of the zone may intimately relate to the trafficing of molecules from IM to OM, and only utilized by metal transport processes as a fortuitous source of energization. Such transport zones likely contain other proteins that comprise the structure and the mechanism, whose essential role in cell membrane physiology prevents, under normal circumstances, the isolation of mutations in their structural genes.

Acknowledgements

The author thanks Salete M. C. Newton for scientific collaborations, helpful discussions and critical reading of the manuscript, and his colleagues at the University of Oklahoma, Raj Annamalai, Jennifer Allen, Matt Bauler, Zhenghua Cao, Xunqing Jiang, Bo Jin, Wallace Kaserer, Marjorie Montague, Zengbiao Qi, Daniel C. Scott, Padma Thulasiraman and Paul Warfel for their experimental contributions. This work was supported by OCAST grant 00072, NSF grant MCB9709418, and NIH grants GM53836 and 1P20RR1182.

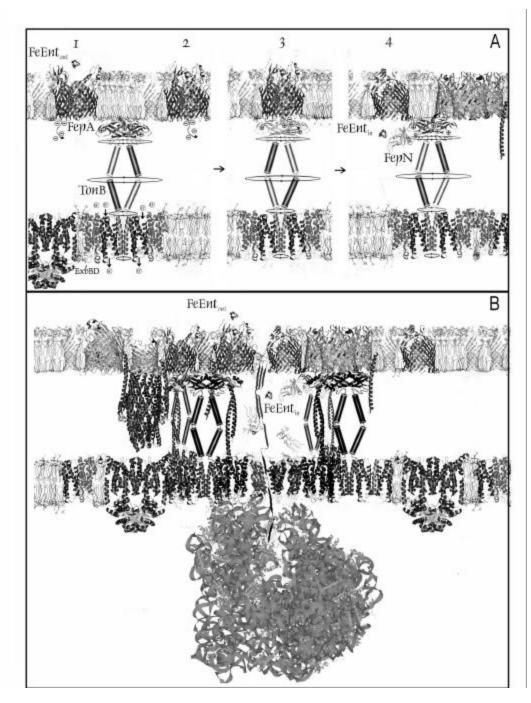


Figure 7. Two models of TonB association with the OM. In the top panel, the rotation of TonB moves it along the inner surface of the OM bilayer, where its positively-charged C-terminal domain adsorbs the negatively-charged N-termini of ligand-bound siderophore receptor proteins, pulling it from inside their \$-barrel and concomitantly internalizing the ferric siderophores. In the bottom panel, TonB-ExbBD and TolA-TolQR form complexes that span the periplasmic space, creating a localized secretion zone that also manifests transport functions, by virtue of its ability to utilize energy sources in the IM.

Literature Cited

- 1. Bradbeer, C. 1993. The proton motive force drives the outer membrane transport of cobalamin in Escherichia coli J Bacteriol. 175:3146-50.
- 2. Braun, M., H. Killmann, and V. Braun 1999. The beta-barrel domain of FhuADelta5-160 is sufficient for TonB-dependent FhuA activities of Escherichia coli Mol Microbiol. 33:1037-49.
- 3. Buchanan, S. K., B. S. Smith, L. Venkatramani, D. Xia, L. Esser, M. Palnitkar, R. Chakraborty, D. van der Helm, and J Deisenhofer 1999. Crystal structure of the outer membrane active transporter FepA from Escherichia coli [see comments] Nat Struct Biol. 6:56-63.
- 4. Cadieux, N., and R. J. Kadner 1999. Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter Proc Natl Acad Sci U S A. 96:10673-8.
- 5. Cao, Z., and P. E. Klebba 2002. Mechanisms of colicin binding and transport through outer membrane porins Biochimie. 84:399-412.
- 6. Cao, Z., Z. Qi, C. Sprencel, S. M. Newton, and P. E. Klebba 2000. Aromatic components of two ferric enterobactin binding sites in escherichia coli fepA [In Process Citation] Mol Microbiol. 37:1306-17.
- 7. Cao, Z., P. Warfel, S. M. Newton, and P. E. Klebba 2002. Spectroscopic observations of ferric enterobactin transport J Biol Chem. 29:29.
- 8. Cascales, E., R. Lloubes, and J. N. Sturgis 2001. The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB Mol Microbiol. 42:795-807.
- 9. Chang, C., A. Mooser, A. Pluckthun, and A. Wlodawer 2001. Crystal structure of the dimeric C-terminal domain of TonB reveals a novel fold J Biol Chem. 276:27535-40.
- 10. Ferguson, A. D., J. Breed, K. Diederichs, W. Welte, and J. W. Coulton 1998. An internal affinity-tag for purification and crystallization of the siderophore receptor FhuA, integral outer membrane protein from Escherichia coli K-12 Protein Sci. 7:1636-8.
- 11. Ferguson, A. D., R. Chakraborty, B. S. Smith, L. Esser, D. van der Helm, and J. Deisenhofer 2002. Structural basis of gating by the outer membrane transporter FecA Science. 295:1715-9
 - 12. Gaspar, J. A., J. A. Thomas, C. L.

- Marolda, and M. A. Valvano 2000. Surface expression of Ospecific lipopolysaccharide in Escherichia coli requires the function of the TolA protein Mol Microbiol. 38:262-75.
- 13. Hancock, R. W., and V. Braun 1976. Nature of the energy requirement for the irreversible adsorption of bacteriophages T1 and phi80 to Escherichia coli J Bacteriol. 125:409-15.
- 14. Higgs, P. I., R. A. Larsen, and K. Postle 2002. Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA Mol Microbiol. 44:271-81.
- 15. Klebba, P. E., and S. M. Newton 1998. Mechanisms of solute transport through outer membrane porins: burning down the house Curr Opin Microbiol. 1:238-247.
- 16. Klebba, P. E., J. M. Rutz, J. Liu, and C. K. Murphy 1993. Mechanisms of TonB-catalyzed iron transport through the enteric bacterial cell envelope J Bioenerg Biomembr. 25:603-11.
- 17. Kojima, S., and D. F. Blair 2001. Conformational change in the stator of the bacterial flagellar motor Biochemistry. 40:13041-50.
- 18. Letain, T. E., and K. Postle 1997. TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in Escherichia coli [published erratum appears in Mol Microbiol 1997 Aug;25(3):617] Mol Microbiol. 24:271-83.
- 19. Locher, K. P., B. Rees, R. Koebnik, A. Mitschler, L. Moulinier, J. P. Rosenbusch, and D. Moras 1998. Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes Cell. 95:771-8.
- 20. Merianos, H. J., N. Cadieux, C. H. Lin, R. J. Kadner, and D. S. Cafiso 2000. Substrate-induced exposure of an energy-coupling motif of a membrane transporter Nat Struct Biol. 7:205-9.
- 21. Phan, P. G., N. Cadieux, and R. J. Kadner 1999. Site-directed mutagenesis of BtuB TonB boxes and its effect on Vitamin B12 uptake in Escherichia coli, p. 427. 99th General Meeting of the American Society for Microbiology. ASM, Chicago, Illinois.
- 22. Pugsley, A. P., and P. Reeves 1977. Uptake of ferrienterochelin by Escherichia coli: energy dependent stage of uptake J Bacteriol. 130:26-36.
- 23. Scott, D. C. 2001. PhD. University of Oklahoma.

- 24. Scott, D. C., Z. Cao, Z. Qi, M. Bauler, J. D. Igo, S. M. Newton, and P. E. Klebba 2001. Exchangeability of N termini in the ligand-gated porins of Escherichia coli J Biol Chem. 276:13025-33.
- 25. Scott, D. C., S. M. C. Newton, and P. E. Klebba 2002. A crosslinking analysis of surface loop motion in FepA J Bacteriol. Submitted.
- 26. Usher, K. C., E. Ozkan, K. H. Gardner, and J. Deisenhofer 2001. The plug domain of FepA, a TonB-dependent transport protein from Escherichia coli, binds its siderophore in the absence
- of the transmembrane barrel domain Proc Natl Acad Sci U S A. 98:10676-81.
- 27. Vakharia, H. L., and K. Postle 2002. FepA with globular domain deletions lacks activity J Bacteriol. 184:5508-12.
- 28. Witty, M., C. Sanz, A. Shah, J. G. Grossmann, K. Mizuguchi, R. N. Perham, and B. Luisi 2002. Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: evidence for an evolutionary relationship with the TonB transporter protein Embo J. 21:4207-18.