AN ABSTRACT OF THE THESIS OF

Frank C. Kuhn for the Master of Science in Biology

Concentration in Molecular and Microbial Biology

presented on November 23, 1999

Title: Cloning and functional homology of SecD within *Enterobacteriaceae*

Abstract approved: Rodney J. Sobieski

The *secD* gene in *Escherichia coli* produces a subunit of the preprotein general secretion pathway, which transports preproteins across the inner membrane into the periplasmic space. Polymerase Chain Reaction (PCR) was performed on the isolated chromosome of *Salmonella choleraesuis* with primers designed from the *E. coli* secD gene. Cloning was confirmed with DNA sequencing which was performed in triplicate to construct the *S. choleraesuis secD* gene. Comparison of the *S. choleraesuis* sequence to *E. coli secD* shows 87% homology in genes that are both 1,848bp in size and code for a protein that is 615 amino acids in length. Translated amino acid sequence comparison shows 95% identity and 98% similar residues. Analysis of the secondary structure shows strong transmembrane similarity to *E. coli* SecD. PCR products were then cloned, and sequences compared in various ways from other *Enterobacteriaceae.* They were found to have an average homology was 87%. Primary stuctures of these products were all above 90% homologous to *E. coli.* Protein expression produced a 54kDa product at the same migration position as that of *E. coli* SecD. The cloned *S. choleraesuis secD* gene was tested for functionality in a cold sensitive mutant with successful results. Conclusions drawn are the *E. coli* and *S. choleraesuis secD* genes are analogous and other sequenced *secD* genes are very conserved and homologous to the *E. coli secD* gene.

CLONING AND FUNCTIONAL HOMOLOGY OF SECD WITHIN THE ENTEROBACTERIACEAE

A Thesis
Presented to
The Division of Biological Science
EMPORIA STATE UNIVERSITY

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

---

by
Frank C. Kuhn
December 1999

Approved by Major Advisor

Approved by Committee Member

Approved by Committee Member

Approved for Division of Biology

Approved for Graduate Council

ACKNOWLEDGMENTS

I would like to thank Dr. Sobieski for his guidance, understanding, and support throughout my research. I would also like to thank Dr. Crupper for the opportunity to work in his lab and his help. Dr. Keck for many useful discussions and supplies on short notice. In addition, I would like to thank my fellow graduate students Ching Wen (Dolores) Tseng and Ngoni Kunonga for their help in my time here at Emporia State University.

PREFACE

This thesis was written in the style of the Journal Cell.

TABLE OF CONTENTS

PAGE

LIST OF FIGURES

PAGE

LIST OF TABLES

PAGE

Direction of Research

The primary focus of this research involves a subunit of the bacterial general secretion pathway, SecD. The *secD* gene of *Escherichia coli* has been sequenced and its functions in *E. coli* have been studied in some detail. Its activities include stabilization and regulation of SecA insertion into the membrane, maintenance of the proton motive force, inhibition of unproductive backwards movement of the preprotein in the membrane, and releasing the secreted protein into the periplasmic space in a competent state. The question posed was, does SecD occur in other members of the *Enterobacteriaceae* or is it only found in *E. coli*?

To answer this question several pathogenic members of this family were obtained including species from *Enterobacter*, *Edwardsiella*, *Klebsiella*, *Proteus*, *Salmonella*, *Serratia*, and *Yersinia*. Chromosomal DNA was isolated from each, and primers for PCR were designed with the aid of a BLAST search using the *E. coli secD* gene as a template. This primer set was used in PCR in an attempt to amplify the *secD* gene from the various species. PCR products of the appropriate size were agarose gel extracted, purified, and ligated into the plasmid cloning vector pT7Blue-3. Several recombinant pT7Blue-3:*secD* clones were selected and colony purified. Multiple restriction digestions were performed on each colony to isolate clones with the expected insert, then those containing the properly sized insert were sequenced. The *secD* gene sequence for each species was constructed using 3- 4 sequencing runs to determine each nucleotide position.

Each sequence was compared to the *E. coli secD* gene for DNA homology. Sequence homology is important because it is a very good indicator of the possible protein structure and function. Comparing new sequences to a known sequence with a known function may allow prediction of possible structure and function in the membrane. Each sequence was translated to its amino acid sequence, and then all protein sequences were compared using

BLASTp and CLUSTALw. Transmembrane regions were predicted using Tmpred and an Eisenburg Hydropathy plot for each sequence, then compared to *E. coli* SecD. The positions of the putative *E. coli* SecD transmembrane regions and large periplasmic domain were found to be conserved in each sequence. Complementation studies were performed using a cold sensitive *secD* mutant strain complemented with the expression vector pQE30 containing the *Salmonella choleraesuis secD* gene. This recombinant expression vector construct was also used in attempts to produce the *S. choleraesuis* SecD protein for comparison to *E. coli* SecD on SDS-PAGE. Before starting a discussion of the materials and methods used in the research, a discussion of various protein secretion systems is needed.

Protein Secretion Systems are Related

The need to move proteins across a membrane is ubiquitous in nature. All cells, whether they are prokaryotes, yeast, or mammalian eukaryotes must move proteins across a membrane to survive. Many organelles such as mitochondria and chloroplasts possess mechanisms which fulfill this function as well, and each of these mechanisms share at least some similarly with the others (Settles and Martienssen, 1998). The bacterial general secretion (Sec) pathway bears similarity to protein translocation in mitochondria, the endoplasmic reticulum, and chloroplasts (Dalbey and Robinson, 1999). Nascent proteins that are targeted for export across the bacterial inner membrane and the endoplasmic reticulum also share similarity in the hydrophobic region of their signal sequences (Martoglio and Dobberstein, 1998). Some subunits of the Sec-dependent translocon are even homologous to subunits of these other systems, namely SecY which is homologous to Sec61α in mammalian cells and to Sec61p in yeast (Jungnickel et al., 1994 and Meyer et al., 1998). The SecE subunit also shares homology with Sec61γ in mammalian cells and Sss1p in yeast, although

the 20% amino acid identity is low (Meyer et al., 1998). Unlike these other subunits of the Sec-dependent pathway, the SecD subunit does not have an analog in eukaryotic cells; this makes it a good potential target for antimicrobial intervention because small molecule inhibitors are less likely to have a deleterious affect on the host.

Prokaryotic Secretion

There are four main types of extracellular protein secretion in prokaryotes, called Types I-IV. This research deals with the inner membrane portion of the pilus assembly and type II protein secretion pathway, called the Sec-dependent or general protein secretion pathway (Stephens and Shapiro, 1997). This pathway is a posttranslational transporter of proteins into the periplasmic space (Wickner and Leonard, 1996). For secretion of extracellular proteins through the outer membrane, the pathway may also use one or more of the Out proteins. The Sec system moves a protein from the cytoplasm through the plasma membrane; it may involve as many as 12- 15 of the gene products, *outC-M* and *outO*, to move a protein from the periplasm, through the outer membrane to the exterior of the cell.

Sec-Dependent Pathway Background and General Characteristics

The Sec-dependent pathway of preprotein translocation in *Escherichia coli* is composed of the proteins SecA, SecB, SecD, SecE, SecF, SecG, yajC, and SecY. These proteins form a transient holoenzyme (also called a translocon) which moves preproteins through the cytoplasmic membrane into the periplasmic space or out of the cell (Duong and Wickner, 1997B). Other factors that affect protein translocation are the leader sequence of the preprotein, the proton electrochemical gradient, and ATP availability in the cell (Arkowitz and Bassilana, 1994). Translocation is a complex process and involves some of the Sec components more than others; some are not strictly needed at all for a low

level of translocation activity *in vitro*. However, for efficient translocation all of the Sec-proteins must be present.

The basic mechanism of Sec-dependent protein translocation starts with the nascent preprotein in the cytoplasm. The preprotein is immediately bound by the cytoplasmic chaperone protein SecB (Collier et al., 1988 and Economou et al., 1995). SecB and the positively charged amino-terminus of the preprotein target the SecB-preprotein complex to the cytoplasmic membrane (Arkowitz and Bassilana, 1994). At the cytoplasmic membrane SecB binds to SecA, which binds the preprotein near the amino-terminal region of the leader sequence (Hartl et al., 1990). SecA is peripheral to the membrane at the time of preprotein binding; but as it binds, SecA inserts into the cytoplasmic membrane along with the leader sequence of the preprotein (Economou et al., 1995). SecA-preprotein insertion into the membrane portion of the translocon causes SecB to release from the preprotein (Duong and Wickner, 1997A).

After membrane insertion by SecA, the leader sequence hydrophobic region is curved like a hairpin in the membrane; then the hydrophilic region is exposed to the periplasmic space and cleaved by the periplasmic protease (Kim, et al., 1994). Cleavage leaves the preprotein inserted in the membrane in close association with SecY, in the SecYEG complex insertion target. It is this integral membrane-bound heterotrimer that forms the initial contact site for the preprotein (Duong and Wickner, 1997A). The SecYEG heterotrimer along with the SecDFyajC complex forms the integral membrane holoenzyme which the preprotein uses to pass through the cytoplasmic membrane into the periplasmic space (Duong and Wickner, 1997B). Insertion of SecA-preprotein into the membrane complex is driven by ATP hydrolysis and is stabilized by SecDFyajC (Duong and Wickner, 1997A).

Repeated insertion and deinsertion of the SecA-preprotein complex into the membrane is continued by several cycles of ATP binding and hydrolysis (Economou et al., 1995). This allows the preprotein to pass through the

holoenzyme in a ratchet type mechanism (Arkowitz and Wickner, 1994). The repeated cycles of insertion and deinsertion occur to insert a sufficient portion of the preprotein to allow the translocation to be rapidly completed by the proton motive force (Schiebel et al., 1991). The proton motive force is stabilized by the SecDFyajC complex, which also allows complete translocation and release of the protein from the membrane (Duong and Wickner, 1997A). The translocated protein may either remain in the periplasmic space, insert into the outer membrane, or continue through the outer membrane, depending on its site of activity (Danese and Silhavy, 1998).

Components of the Sec-Dependent Pathway

A combination of genetic and biochemical studies have led to the identification of the Sec genes involved in the Sec-dependent protein secretion pathway of *E. coli*. These genes are *secA*, *secB*, *secD*, *secE*, *secF*, *secG*, *yajC*, and *secY* (Doung and Wickner, 1997B). The most conserved portion of the translocon is the integral membrane complex SecYE in association with the peripheral protein secA; these parts of the translocation mechanism are strictly required for cell viability (Tokuda, 1994). Characteristics of the preproteins themselves also affect the function of the Sec-dependent translocation mechanism.

Preproteins

Before beginning a discussion of each Sec gene product and its function in protein translocation, a discussion of the preprotein characteristics that affect translocation is needed. Preproteins consist of a signal sequence at the amino-terminus consisting of a positively charged anterior, followed by a hydrophobic region, and a polar posterior region with the future mature domain that follows (Hayashi and Wu, 1990). This leader sequence is cleaved during translocation,

leaving the mature domain, which folds into its active conformation in the periplasmic space (Arkowitz and Bassilana, 1994).

The leader sequence is usually 15 to 30 amino acids long, with positively charged amino acids at the first three positions of the amino-terminal end (Arkowitz and Bassilana, 1994). This is followed by a hydrophobic region of 10 to 15 amino acids that is essential for secretion. This hydrophobic region is inserted into the membrane to expose the posterior cleavage site in the hydrophilic region of the leader sequence to the periplasmic protease. The carboxy-terminal end of the leader sequence is a hydrophilic region, the last three amino acids of which are used as a target site for cleavage of the leader sequence during translocation (Arkowitz and Bassilana, 1994). After cleavage, the signal sequence is degraded by the cytosolic oligopeptidase (Martoglio and Dobberstein, 1998). The future mature domain portion of the preprotein are bound by several molecules of the SecB chaperone proteins for transport to the cytoplasmic membrane (Randall and Hardy, 1995 and Randall et al., 1998).

The secB Gene

The SecB gene product forms a homotetramer that acts as a export-dedicated chaperone for nascent preproteins in the cytoplasm (Tokuda, 1994 and Danese and Silhavy, 1998). The nascent preprotein emerges from the ribosome into the cytoplasm and as about 150 amino acid residues extend (Danese and Silhavy, 1998), the mature portion of the preprotein is immediately bound by SecB (Collier et al., 1988 and Economou et al., 1995). This initial binding to basic residues of the preprotein causes a conformational change in SecB which allows it to bind to additional basic or apolar regions of the preprotein (Wickner and Leonard, 1996). As translation continues, several molecules of SecB bind to sites along the nascent preprotein (Arkowitz and Bassilana, 1994). The arrangement of

several molecules of SecB along the mature portion of the preprotein prevents folding and aggregation of the preprotein (Hartl et al., 1990).

In general, SecB functions to inhibit the folding and assembly of the nascent preprotein by binding to its unfolded state. It keeps the preprotein in a translocationally stable unfolded form which allows it to move into the membrane bound Sec translocon instead of folding inappropriately in the cytoplasm (Arkowitz and Bassilana, 1994). SecB keeps the preprotein from folding before it can be transported through the cytoplasmic membrane, so it is said to keep the preprotein in a translocation-competent state (Hartl et al., 1990 and Kosic et al., 1993). SecB and the positively charged amino-terminus of the preprotein leader sequence target the preprotein to the cytoplasmic membrane moving it there in a translocationally-competent form (Arkowitz and Bassilana, 1994). Then the SecB SecA affinity enables preprotein association with SecA for insertion into the cell membrane translocon (Wickner and Leonard, 1996).

The secA Gene

SecA is a 102 kDa protein that functions as a homodimer (Tokuda, 1994 and Danese and Silhavy, 1998) and has been shown to be conserved in gram negative bacteria (Arkowitz and Bassilana, 1994). SecA is also the only Sec protein to exhibit ATP hydrolysis activity, so it is termed the translocation ATPase (Lill et al., 1990). It is the most diverse of the Sec proteins, it can be found as a cytoplasmic, peripheral, and integral membrane protein (Kim et al., 1994). SecA in the cytoplasm has two functions. It can bind to preproteins targeting them to the inner membrane and it can act as its own translational repressor by binding to *secA* mRNA to inhibit protein translation (Danese and Silhavy, 1998).

Peripheral SecA-ADP interacts with the SecB-preprotein complex as it moves in close proximity to the inner membrane (Schiebel et al., 1991). SecA acts as the receptor for this complex due to its high affinity for SecB, the preprotein

leader sequence, and mature domains (Lill et al., 1990 and Hartl et al., 1990). SecY and anionic (acidic) phospholipids act as a high affinity binding site for the SecA-SecB-preprotein complex at the inner leaflet of the cytoplasmic membrane (Hartl et al., 1990, Lill et al., 1990, and Ulbrandt et al., 1992). SecA binds to both the leader and mature sequences of the nascent preprotein, as well as, to acidic phospholipids and SecYEG; these associations allow it to release the ADP and bind ATP at NBD1 (nucleotide binding domain) and begin the process of translocation (Duong and Wickner, 1997A and Economou et al., 1995).

As the SecA-preprotein-SecB complex interacts with the inner membrane, ATP binds, inducing SecA to undergo a conformational change prompting a 30 kDa domain to insert deeply into the membrane as it partially unfolds (Ulbrandt et al., 1992 and Economou et al., 1995). SecA insertion using ATP binding energy causes 20 to 30 amino acids of the preprotein to insert into the membrane (Schiebel et al., 1991) in close association with SecY (Meyer et al., 1999) and exposes a portion of SecA to the periplasm (Kim et al., 1994). Movement of SecB-preprotein to the membrane complex brings SecB in association with SecA; they bind and SecB is removed from the preprotein and is no longer involved in translocation (Duong and Wickner, 1997B). This activates the ATPase activity of SecA (Lill et al., 1990) and as ATP hydrolysis occurs the preprotein is released from SecA (Schiebel, et al., 1991), then SecA deinserts from the membrane (Economou et al., 1995). SecA membrane insertions and deinsertions are driven by ATP and stabilized by the SecDFyajC complex, this occurs until a sufficient portion of the preprotein is inserted into the membrane for the membrane electrochemical potential to complete translocation (Economou et al., 1995, Kim et al., 1994, and Schiebel, et al., 1991). The proton motive force rapidly completes translocation when the preprotein is no longer associated with SecA (Economou et al., 1995, Kim et al., 1994, and Schiebel, et al., 1991).

The secYEG Gene Complex

The prlA (*secY*) gene product was identified in 1984 and the gene which encodes it was named *secY* (Ito, 1984). It was the first translocation protein identified in either prokaryotes or eukaryotes as an integral membrane protein, and it shares homology with the gram positive *Bacillus subtilis* SecY (Ito, 1992). Recently, electron microscopy demonstrated the formation of ring-like structures by *B. subtilis* SecY in the membrane similar to the Sec61p complex (Meyer et al., 1999). *E. coli* SecY has a mass of 49 kDa and must be visualized by 20% SDS-PAGE due to its excessive hydrophobic interaction with SDS, and it will aggregate if heated to 100°C before loading (Ito, 1984). SecY contains 10 transmembrane regions which are very evident in its hyropathy profile, and its amino- and carboxy- terminal regions are cytoplasmic (Akiyama and Ito, 1987). Its structure looks similar to other solute transport membrane proteins and appears to be the most conserved of the Sec proteins (Tokuda, 1994).

The *secE* gene was discovered by mapping a mutation which caused a cold sensitive protein export defect (Riggs et al., 1988). SecE has a mass of 14 kDa and spans the cytoplasmic membrane three times with the amino-terminus cytoplasmic (Tokuda, 1994 and Yang et al., 1997). The main role of SecE is to stabilize SecY (Taura et al., 1993). Its biological activity appears to reside in its third transmembrane segment (Tokuda, 1994). It is also known through sequencing studies that the SecE homolog in *B. subtilis* only retains one transmembrane segment, although many amino acids are conserved in the pre-cytoplasmic region (Tokuda, 1994). Recently, membrane vesicles containing only 2% of the wild type *E. coli* SecE were shown to be active in protein translocation, indicating SecE may not be essential *in vitro* (Yang et al., 1997).

Direct involvement of SecY and SecE in protein translocation has been demonstrated in proteoliposome reconstitution studies, and they are essential for translocation *in vivo* (Taura and Ito, 1994). The SecYE complex forms the pathway for the preprotein to pass through the inner lipid bilayer into the

periplasmic space (Arkowitz and Bassilana, 1994). SecY and SecE act and have always been isolated as a complex, of the integral membrane subunits these alone can reconstitute membrane translocation (Brundage et al., 1990) . This has been shown with immuno-precipitation using only anti-SecY antibodies that also contain SecE (and band 1) upon analysis, and affinity chromatography studies that also contain both proteins (Brundage et al., 1990). It was shown band 1 which co-precipitates with the SecYE complex was in fact SecG (Brundage et al., 1990 and Tokuda, 1994).

While the SecYE complex does have activity alone, SecG has been shown to increase preprotein translocation *in vitro* and *in vivo* (Duong and Wickner, 1997B). Genetic studies showed the p12 gene product was directly involved in protein translocation and was renamed SecG (Nishiyama et al., 1994). SecG increases translocation by stabilizing the SecYE complex (Kim et al., 1994) after SecA SecY interaction (Hanada et al., 1994). SecG does this by undergoing a topology inversion as SecA inserts into the membrane and reverts upon SecA deinsertion (Nishiyama et al., 1996 and Wickner and Leonard, 1996). SecG significantly increases the efficiency of protein translocation over the SecYE complex alone, and is especially important at low temperature or in the absence of the proton motive force (Hanada et al., 1994 and Hanada et al., 1996).

The SecYEG heterotrimer and SecA are the most conserved subunits of the translocation machinery. These components do have some translocation activity alone, but efficient translocation requires the SecD operon gene products (Duong and Wickner, 1997B). Together these products form the SecYEGDFyajC holoenzyme for preprotein translocation across the membrane (Duong and Wickner, 1997A).

The secDF Gene Complex

SecDFyajC, the strictly prokaryotic subunits of the Sec translocon, work together to promote preprotein translocation through the SecYE protein-

conducting channel of the translocation machinery (Duong and Wickner, 1997A and Meyer et al., 1999). These SecD operon encoded proteins form an integral membrane heterotrimeric complex composed of SecD, SecF, and yajC proteins called the SecDFyajC complex (Duong and Wickner, 1997B). This SecDFyajC complex associates with the SecYEG complex to form SecYEGDFyajC, the hexameric integral membrane translocase holoenzyme, also called the translocon (Duong and Wickner, 1997A).

The *secD* gene was discovered in 1987 through mutations which led to impaired growth, cold sensitivity (Cs), and reduced protein export (Gardel et al., 1987). These severe defects in protein translocation and growth were analyzed and mapped to the *secD* locus (Gardel et al., 1987 and Gardel et al., 1990). The *secD* locus was sequenced and the *secD* and *secF* genes were identified as part of an operon which also contained yajC (Gardel et al., 1990 and Pogliano and Beckwith, 1992). This was the first identified co-transcription of genes involved in protein translocation (Gardel et al., 1990). The *secD* and *secF* genes share homology and both code for integral cytoplasmic membrane proteins which are involved in later steps of preprotein translocation after preprotein insertion (Gardel et al., 1990 and Matsuyama et al., 1992 and Sugai and Wu., 1992).

Kyte-Doolittle hydropathy plots SecD and SecF are very similar and the C-terminal 176 amino acids share 26% identity with 48% similarity in these proteins (Gardel et al., 1990). The SecD and SecF proteins migrate to 54kDa and 27kDa respectively by 10% SDS-PAGE, this differs from the theoretical calculated from the gene sequences. This is caused by the hydrophobic nature of SecD and SecF in a manner similar to SecY anomalous migration in low percent gels (Matsuyama et al., 1992). When the acrylamide concentration was increased to 20% SecD migrated to 67kDa and SecF to 35kDa, approximately the predicted mass of the proteins (Matsuyama et al., 1992). Topology of SecD in the membrane consists of six transmembrane regions and a large periplasmic domain, with both the N- and C- terminus cytoplasmic (Pogliano and Beckwith, 1992).

SecF also has a large periplasmic domain between two transmembrane stretches (Pogliano and Beckwith, 1994).

The direct participation of SecD in translocation was shown using spheroplasts and anti-SecD antibodies. Treatment of spheroplasts with antibody resulted in the inhibition of the protein translocation of several proteins, as shown by the accumulation of both preprotein and mature forms in the spheroplasts (Tokuda, 1994). It was also demonstrated that the protein is not released from the cell surface in SecDF depleted cells, shown by the sensitivity of a protein to trypsin which occurs when this protein is not folded into its mature form (Arkowitz and Bassilana, 1994). These results indicate SecD plays a role in release of the mature protein at the periplasmic surface of the cytoplasmic membrane (Arkowitz and Wickner, 1994).

ATP driven SecA cycles of insertion and de-insertion into the membrane SecYEG complex are regulated and stabilized by SecDFyajC during translocation (Kim et al., 1994 Economou et al., 1995 and Duong and Wickner, 1997B). This interaction with SecA has been shown in studies where the SecDF locus is overexpressed leading to enhanced SecA insertion and by SecDF depletion that caused instability in membrane-inserted SecA (Economou et al., 1995 and Duong and Wickner, 1997A). The SecDFyajC complex slows both the forward and reverse translocation by stabilizing the SecA insertion state and maintaining the proton motive force (PMF). This prevents unproductive backwards movement of the protein in the translocon which would otherwise be associated with SecA deinsertion (Duong and Wickner, 1997A).

SecDF directly couples the PMF to translocation. Cells depleted of SecDF are unable to maintain the proton electrochemical gradient and show a striking reduction of mature domain translocation (Arkowitz and Wickner, 1994). This was shown in inverted membrane vesicles (IMV's) which cannot maintain a stable PMF when SecDF depleted, so they cannot support translocation in ATP depleted vesicles (Arkowitz and Wickner, 1994). Although the loss of PMF

causes only a small decrease in overall translocation, it is very costly to the cell in ATP utilized (Duong and Wickner, 1997B).

Duong and Wickner in 1997 reported finding the translocon holoenzyme and confirmed the direct involvement of SecDF in the translocation of preproteins across the membrane. This holoenzyme was shown to involve both the SecYEG and SecDFyajC subunits embedded in the cytoplasmic membrane. While it is true that the SecYE complex in the presence of ATP can activate SecA-preprotein insertion into the membrane and its ATPase activity; to have efficient PMF assisted translocation and SecA membrane cycling the cell must also have SecDFyajC activity stabilize this interaction (Duong and Wickner, 1997A and 1997B).

Energy Requirements for Protein Translocation

Both ATP and proton motive force (PMF) are used as energy sources for protein translocation in *E. coli* (Arkowitz and Wickner, 1994). ATP is essential for the initiation of translocation, but the proton motive force is used to drive protein translocation to completion after insertion of the preprotein into the membrane (Arkowitz and Wickner, 1994 and Tokuda, 1994). The insertion of the SecA-preprotein is mechanistically different from the ongoing protein translocation that moves the protein though the membrane in a racheting type mechanism (Tokuda, 1994). It is ATP that drives the insertion of the SecA-preprotein deeply into the inner lipid bilayer; its hydrolysis transfers the preprotein from SecA to the SecYEG complex and then releases SecA from the membrane (Economou et al., 1995). To translocate the entire preprotein through the cytoplasmic membrane this SecA and ATP binding happens in several cycles (Tokuda, 1994). It has been shown that each of these cycles translocates a short portion, 20- 30 residues of the preprotein as SecA inserts due to the ATP binding (Economou et al., 1995).

The proton motive force is composed of the electrical gradient and the chemical proton gradient and causes a significant stimulation of translocation across the membrane (Arkowitz, and Bassilana, 1994). It appears that the translocation machinery is in an energized state in the presence of the proton motive force and it has a higher affinity for ATP (Economou et al., 1995). Proton transfer through the membrane is coupled with the protein translocation and is stabilized by SecD and SecF gene products (Arkowitz and Wickner, 1994). SecD and SecF are also required for the release of the translocated protein from the membrane outer surface in a competent state (Tokuda, 1994).

Materials and Methods

The following methods were used to demonstrate the presence of SecD in several members of the *Enterobacteriaceae* family. Standard procedures and protocols used in this research can be found in Current Protocols in Molecular Biology (Altshul et al., 1995) or Molecular Cloning (Sambrook et al., 1989). The *secD* gene of this family was amplified, isolated, cloned, and sequenced to compare its homology with the *E. coli secD* gene. Lastly, the *secD* gene of *S. choleraesuis* was cloned into an expression vector for protein production and complementation of a cold sensitive (Cs) *E. coli* mutant, then compared to the SecD protein of *E. coli*.

I. Bacteria and Propagation Conditions

All bacterial strains used were members of the *Enterobacteriaceae* family of gram negative bacteria. Table 1 lists the bacteria used, their source, propagation media, and other specifics about each strain. Most strains used were confirmed with standard biochemical tests, and routinely propagated on Luria-Bertani (LB) or Brain Heart Infusion (BHI) plates incubated at 37°C. The *E. coli* K12 KJ173 SecD mutant from Dr. Beckwith of Harvard University was obtained through a request to the *E. coli* stock culture collection (http://cgsc.biology.yale.edu/) and routinely propagated at 40°C on BHI plates containing 200 μg/ml Streptomycin (Strep). Long term storage of bacterial cultures was in BHI/20% glycerol at -70°C.

II. Plasmids and Host Cell Propagation Conditions

The host cell for the plasmids was *E. coli* DH5α propagated on LB plates containing 100 μg/ml Ampicillin (Amp) or 25 μg/ml Kanamycin (Kan) with incubation at 37°C. Table 2 shows the plasmids used, their propagation conditions, and their source. Recombinant pQE16, pQE30, pQE31, and pQE32

TABLE 1. The bacteria used in this research. The bacterial source and the propagation media are shown for each species used.

Table 1. Bacteria used in research

Genus-Species	Strain	Source	Propagation Media	Comments
Escherichia coli	DH5alpha	Dr. Crupper	LB	
Escherichia coli	K12 KJ173 SecD mutant	E. coli stock culture collection	BHI 200ug/ml Strep	From Dr. Beckwith at Harvard University
Salmonella choleraesuis	Kunzendorf Porcine 210	A KSU veterinary isolate	BHI	From KSU College of Veterinary Medicine
Klebsiella pneumoniae		A KSU equine veterinary isolate	BHI	From KSU College of Veterinary Medicine
Enterobacter aerogenes		Newman Hospital clinical isolate	BHI	
Citrobacter freundii		Newman Hospital clinical isolate	BHI	
Yersinia enterocolitica		Presque Isle	LB	
Serratia marcescens		Dr. Sue Katz, Midwestern University	BHI	19555 N. 59 Ave. Glendale, AZ 85308
Proteus vulgaris		Newman Hospital clinical isolate	BHI	
Proteus mirabilis		Newman Hospital clinical isolate	BHI	
Edwardsiella tarda		Presque Isle	BHI	

TABLE 2. The plasmids used in this research. The plasmid, its source, and the propagation media are shown for each host species used.

Table 2. Plasmids and Host Bacteria

Plasmid	Host Cell	Propagation	Comments
pBluescript SK+/-	Escherichia coli DH5alpha	LB 100ug/ml Amp at 37°C	Stratagene
pT7Blue-3	Escherichia coli DH5alpha	LB 100ug/ml Amp at 37°C	Novagen Madison, WI
pQE30	Escherichia coli DH5alpha	LB 100ug/ml Amp at 37°C	QIAGEN Chatsworth,MA
pQE31	Escherichia coli DH5alpha	LB 100ug/ml Amp at 37°C	QIAGEN Chatsworth,MA
pQE32	Escherichia coli DH5alpha	LB 100ug/ml Amp at 37°C	QIAGEN Chatsworth,MA
pREP4	Escherichia coli DH5alpha	LB 25ug/ml Kan at 37°C	QIAGEN Chatsworth,MA

plasmids for protein expression were propagated in *E. coli* DH5α containing the repressor plasmid pREP4 grown on LB plates containing 100 μg/ml Amp and 25 μg/ml Kan incubated at 37°C.

III. Chromosomal DNA Isolation, Purification, and Quantitation

The chromosomal isolation performed is a modification of an established protocol from Current Protocols in Molecular Biology (Altshul et al., 1995). *E. coli* DH5α, *S. choleraesuis*, *Klebsiella pneumoniae*, *Serratia marcescens*, *Proteus vulgaris*, *P. mirabilis*, *Enterobacter aerogenes*, or *Edwardsiella tarda* cultures were incubated overnight in 25 ml of BHI broth at 37°C with shaking at 250 rpm in the New Brunswick Series 25 incubator. The overnight cultures were transferred to 50 ml disposable conical tubes and the cells pelleted by centrifugation at 4,000 rpm for 5 minutes in the Sorvall General Laboratory Centrifuge 1. The supernatant was removed and the pellet was resuspended in 5 ml TE (10 mM Tris, 1 mM EDTA at pH 8.0). The cells were lysed by adding 250 μl of 20% Sodium Dodecyl Sulfate (SDS), mixed by inversion for one minute, then incubated at 60°C for 15 minutes. The cleared solution received 50 μl of RNAse (10 μg/ml) and mixed by inversion for 15 seconds. The solution was incubated at 37°C for 20 minutes. Proteins in the solution were digested by adding 75 μl of Proteinase K (10 μg/ml), mixed by inversion for 15 seconds, and incubated at 60°C for one hour. To precipitate the SDS, lipids, and denatured proteins from the solution, 6 ml of molecular biology grade saturated phenol was added to each tube and mixed by inversion for 3 minutes, followed by centrifugation at 4000 rpm for 5 minutes. The upper (aqueous) layer was transferred to a new 15 ml conical tube and the phenol wash was repeated. An equal volume chloroform/isoamyl alcohol (24:1) was added to the solution, mixed by inversion for 3 minutes, then centrifuged for 5 minutes at 4,000 rpm. The aqueous (upper) layer was placed

into a new 15 ml conical tube, then the chloroform /isoamyl alcohol wash was repeated to continue to remove any contaminating traces of phenol. The aqueous layer was placed into a 50 ml conical tube, filled to volume with ice-cold 95% ethanol and mixed by inversion to precipitate the chromosomal DNA. A hooked pasteur pipette, made by flaming the end of the pipette, was used to remove the chromosomal DNA from the tube. The DNA was placed into the bottom of a 1.5 ml microcentrifuge tube, then a pipet was used to remove excess ethanol before spreading the DNA in the tube to dry. The DNA was dried in a 37°C incubator for 30 minutes, and resuspended in 1 ml of TE for storage at 4°C.

The DNA concentration and purity was determined using a Beckman DU 530 spectrophotometer. The DNA purity was measured using the 260/280 O.D. with a ratio between 1.8 and 1.95 considered acceptable. A 1:100 dilution was used to determine the DNA concentration, using the formula: (260 O.D. reading)(Dilution factor)(the double stranded factor: 50) = DNA concentration in μg/ml (1998 Promega catalog).

IV. Polymerase Chain Reaction (PCR)

All PCR reactions were performed using a Minicycler series Thermocycler (MJ Research, Watertown, MA). Primers were designed from the *E. coli secD* gene sequence (Gardel et al., 1990) obtained from a Basic Local Alignment Search Tool (BLAST) search (Altschul et al., 1997). The primer set used in cloning vector experiments consisted of the forward primer 5' GTG TTA AAC CGT TAT CCT TTG TGG AAG 3' with the reverse complement primer 5' TCA GAT TGA CAG CTT CTT GAC GCG C 3' from Integrated DNA technologies (Coralville, IA) which were delivered desalted, but not cartridge purified. The 1μl of purified chromosomal DNA from *E. coli* DH5α, *S. choleraesuis*, *K. pneumoniae*, *Y. enterocolitica*, *S. marcescens*, *P. vulgaris*, *P. mirabilis*, *C. freundii*, *E. aerogenes*, or *E. tarda* was used at 500 ng to 1 μg (quantitation as previously

described) per reaction in the following PCR protocols.

A master mix was made containing 1X Vent PCR buffer, 2 mM $MgSO_4$, ddH_2O, dNTPs, and 2U Vent polymerase (New England Biolabs). The dNTP's from Promega (Madison, WI) were used at a final concentration of 2 mM consisting of dATP (2'-deoxyadenosine-5'-triphosphate, sodium salt), dTTP (2'-deoxythymidine-5'-triphosphate, sodium salt), dGTP (2'deoxyguanosine-5'-triphosphate, sodium salt), and dCTP (2'-deoxycytidine 5' -triphosphate, sodium salt) in a 25mM stock solution. Each reaction was performed in a 0.5 ml microfuge tube at 50 µl per reaction with one drop of filter sterilized mineral oil as an overlay and also added to each well of the thermocycler.

The PCR program consisted of 5 minutes of denaturation at 94°C, followed by 29 cycles of denaturation at 94°C for 1 minute, with primer annealing for 30 seconds each at 53°C, 54°C, 55°C, and elongation at 72°C for 2 minutes. The program finished with an additional elongation step at 72°C for 5 minutes; the tubes were held at 4°C until removed and stored at 4°C.

PCR conditions were developed using the *E. coli* DH5α chromosome to optimize the magnesium concentration, then these conditions were used in subsequent reactions with the *E. coli* DH5α used as a positive control. A positive PCR reaction was considered a band at approximately 1.85kDa when electrophoresed in an agarose gel with the appropriate molecular weight markers.

V. Agarose Gel Electrophoresis

Agarose gel elecrophoresis was carried out in a E-C MiniCell EC370M Electrophoretic Gel System (Fisher Biotech) according to the standard protocol (Sambrook et al., 1989). The DNA sample mixed with 1X loading buffer (0.25% bromophenol blue, 30% glycerol) was loaded into a submerged 1.2% molecular biology grade agarose (Fisher) gel in 1X TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA), and electrophoresed at 80 to 120 volts for 40 to 90 minutes. Gel

images were captured with a UV camera (Ultra-Lum, model CCD) and processed using the SCIcon image program from NIH.

VI. Agarose Gel DNA Purification

PCR product amplified from each chromosome was loaded into a 1.2% agarose gel after adding 70 µl of PCR product into 14 µl of 6X blue loading dye. The DNA was separated by electrophoresis for 80 minutes at 98 volts. The expected DNA band at 1.85 Kb was observed under a Transilluminator FBTIV-88 UV light (Fisher Scientific) and excised using a new razor blade, then placed in a 1.5 ml microfuge tube. This DNA was purified from the agarose gel using the QIAEX II Agarose Gel Extraction kit (QIAGEN, Chatsworth, MA) and protocols according to the manufacturers recommendations.

VII. Ligation

The gel purified putative *secD* gene from *S. choleraesuis, K. pneumoniae, Y. enterocolitica, C. freundii, E. aerogenes*, and *E. tarda* was ligated into the blunt end cloning vector pT7BIue-3 (Novagen) using T4 DNA ligase (Promega). The DNA fragment of approximately 1.85Kb was inserted at the *EcoR*V location of the multiple cloning site (MCS) of the pT7BIue-3 in the following manner.

A 0.5 ml microfuge tube received 3µl of ddH_2O and 2 µl of the purified DNA fragment. The end conversion mix (Novagen) was retrieved from storage at -70°C and thawed on ice for 10 minutes, 5µl was added to the tube and flicked several times to mix. The solution was incubated for 15 minutes at 22°C, heat inactivated at 75°C for 5 minutes using the thermocycler program called "CLONING". The solution was then incubated on ice for 2 minutes. The pT7Blue-3 plasmid was added to the solution at 50ng (1µl) and 1µl T4 DNA ligase was added last. The tube was flicked several times to mix and incubated for 2 hours at 22°C using the thermocycler program called "LIGREAC".

Recombinant plasmids were stored on ice until transformed into *E. coli* DH5α competent cells.

VIII. Transformation

Competent cells were retrieved from storage at -70°C and thawed on ice for 10 minutes. A chilled 0.5 ml microfuge tube was removed from the -20°C freezer and 100 µl of competent cells were added using a chilled pipet tip. Approximately 500 ng of plasmid DNA was added, and the tube was flicked several times to mix. The mixture was incubated at 4°C for 15 minutes then heat shocked at 42°C for 90 seconds, followed by incubation at 4°C for 5 minutes using the thermocycler program "BAD". LB (600µl) was added to a 1.5 ml microfuge tube and the competent cell solution was added to the media, 150 µl was spread plated onto a LB-100 µg/ml Amp plate and incubated overnight at 37°C.

IX. Colony Selection and Purification

The *E. coli* DH5α competent cells containing the recombinant pT7Blue-3 were grown on LB-100 µg/ml Amp agar containing 5-bromo-4-chloro-3-indolyl-B-D-galactoside (X-gal). This provided blue/white screening to select colonies that contained plasmid carrying a DNA insert. Eight white colonies were selected for isolation along with a blue colony as a negative control in each gene insertion attempt. The selected colonies were streaked onto BHI 100 µg/ml Amp X-gal plates for colony purification, then streaked to BHI-100µg/ml plates to grow pure cultures for plasmid isolation. The selected colonies were stored in BHI/20% glycerol at -70°C for long term storage.

X. Plasmid Miniprep

The plasmid miniprep was performed as described by in Molecular Cloning (Sambrook et al., 1989) with modification. A colony was placed into 3 ml of LB broth and centrifuged for 20 seconds, then a loopful of this mixture was added to a 3 ml LB-100 μg/ml Amp broth and grown for 16 hours at 37°C with shaking at 250 rpm in a New Brunswick Series 25 Incubator. The cells were harvested by placing half the suspension in a 1.5 ml microfuge tube and centrifuging for 1 minute at maximum speed in a Beckman Microfuge lite centrifuge. The supernatant was decanted and the pellet resuspended with the remaining overnight cell suspension then centrifuged again for 1 minute. The supernatant was decanted and the pellet was resuspended by adding 200 μl of 4°C Solution I (50mM sucrose, 25mM Tris pH8.0, 10mM EDTA pH8.0) and pipetting. Solution II (0.2 N NaOH, 1% SDS) was made fresh and 200 μl was added to the tube and mixed by inverting the tube 10 times. The mixture was incubated at room temperature for 10 minutes or until the the solution became clear. Then 150 μl of Solution III (5M potassium acetate, 11.5% acetic acid) was added to the tube, mixed by inversion, and incubated on ice for 10 minutes. The microfuge tube was centrifuged for 5 minutes, and the supernatant was transferred to a new 1.5 ml tube. Ten microliters of RNAse (1 μg/ml) was added to the tube and incubated at room temperature for 15 minutes. Then 500 μl of 1.6 M NaCl-13% PEG (polyethylene glycol) was added and mixed by inversion to selectively precipitate the plasmid DNA from the solution. The tube was centrifuged for 5 minutes and the supernatant was decanted, then the pellet was resuspended in 400 μl TE buffer (10mM Tris, 1 mM EDTA pH 8.0) by vortexing. Phenol extraction was performed by adding 400 μl of phenol to the tube which was mixed by inversion for 1 minute, followed by centrifugation at maximum rpm for 4 minutes. The upper (aqueous) layer was transferred to a new 1.5 ml tube

and chloroform/isoamyl alcohol (24:1) was added to the solution at equal volume, mixed by inversion for 1 minute, then centrifuged for 4 minutes at maximum rpm. The 1.5 ml tube was filled to volume with ice cold 95% ethanol, incubated on ice for 10 minutes, and centrifuged at 4°C for 5 minutes at maximum speed. The supernatant was decanted and 500 µl of ice cold 70% ethanol was added to the tube and vortexed, then centrifuged for 5 minutes at maximum speed at 4°C. The supernatant was decanted, then the pellet was resuspended in 30 µl ddH_2O. Restriction enzyme digestion was performed on 3 µl of the plasmid DNA, with the rest stored at -20°C.

XI. Recombinant Plasmid Restriction Digestion

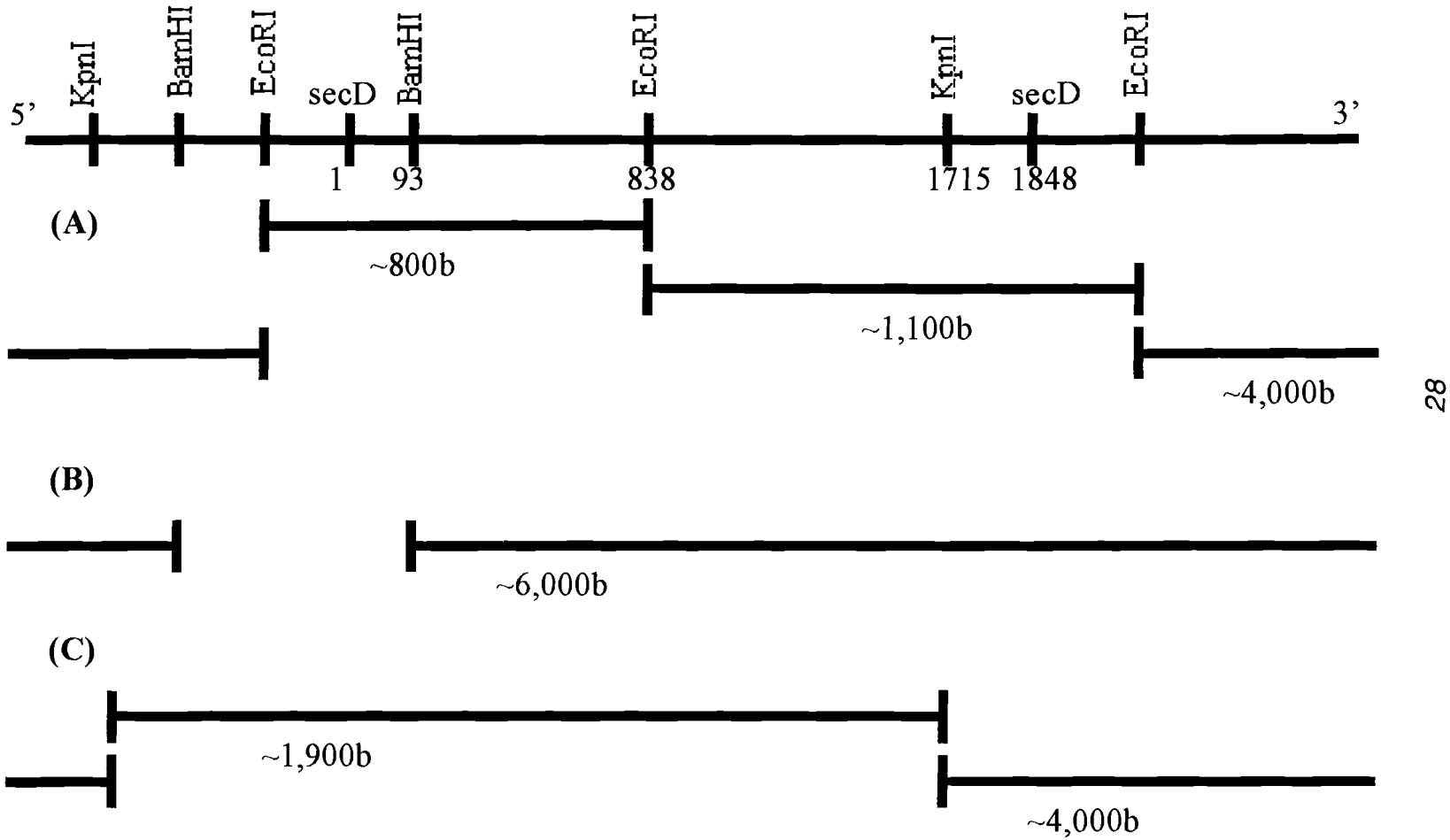
Restriction digestion with *EcoR*I, *BamH*I, or *Kpn*I was performed to check the DNA fragment insert size to show proper insertion into the plasmid. The *EcoR*I restriction site flanks the *EcoR*V location on either side (figure 1a) into which the DNA fragment was inserted. *EcoR*I cuts out the inserted DNA so its size can be ascertained. The *BamH*I and *Kpn*I enzymes also determine the orientation of insertion of the putative *secD* gene, as shown by figures 1b and 1c. These restriction digestions were electrophoresed using 1.2% agarose gels, as previously described, and the insert size was compared to the gel purified *S. choleraesuis secD* gene, and to standard molecular weight markers (Minnesota Molecular). Two of the colonies that appeared to contain plasmids with the appropriate insert were QIAGEN kit plasmid purified and sequenced.

XII. Competent Cell Production

E. coli DH5α and *E. coli* DH5α containing the repressor plasmid pREP4 were used in the preparation of competent cells. These cells were used for routine cloning and protein expression experiments respectively.

FIGURE 1. The *secD* gene was inserted into the pT7Blue-3 MCS at the *EcoR*V restriction site. To examine the recombinant clones restriction digestion was performed with various enzymes, the result was examined on an agarose gel for the proper sized insert in the proper orientation. (A) Restriction digestion with *EcoR*I leaves a 800, 1,100, and 4,000 base fragments because it cuts out the DNA insert and cuts the secD gene at position 838. (B) Restriction digestion with *BamH*I leaves a 6,000 base band if the gene is properly inserted. Improper insertion leaves bands at 1,900 and 4,000. (C) Restriction digestion with *Kpn*I leaves bands at 1,900 and 4,000 if the gene is properly inserted, improper insertion leaves a single band at 6,000.

Figure 1. Restriction digestion of *secD* gene.



The *E. coli* DH5α was incubated overnight in 2 ml of 2X LB broth at 30°C with shaking at 280 rpm in a New Brunswick Series 25 incubator. Then 500 µl of the overnight culture was added to 200 ml of preheated LB broth in a 1 liter flask and the incubation was continued at 30°C with shaking at 280 rpm. After approximately 5 hours the culture absorbance at $O.D._{600}$ reached 0.3, and 4 ml of 1 M $MgCl_2$ was added; incubation was continued for approximately 1 hour until the absorbance at $O.D._{600}$ reached 0.45. The culture was placed into 50 ml tubes and incubated on ice for 2 hours. The cells were pelleted by centrifugation at 3,000 x G for 5 minutes at 4°C in a Beckman J2-HS centrifuge with a Beckman JA-20 rotor. The pellet was resuspended in 100 ml of 4°C $CaCl_2$ media and incubated on ice for 40 minutes. The cells were pelleted by centrifugation at 3,000 x G for 5 minutes at 4°C then resuspended in 10 ml of ice-cold $CaCl_2$-10% glycerol media. The cells were placed into 1.5 ml microfuge tubes in 200 µl aliquotes and stored at -70°C.

The *E. coli* KJ173 SecD mutant cells were prepared according to the following protocol similar to that found in Molecular Cloning (Sambrook et al.,1989). A loopful of the mutant cells stored at -70°C was inoculated into 3 ml LB, vortexed for 10 seconds, then a loopful was inoculated into 3 ml LB 200µg/ml Strep and grown for 14 hours at 37°C with shaking at 250 rpm. Prewarmed 50ml 2XLB200µg/ml Strep with 1ml 1M $MgCl_2$ (added just before warming) was inoculated with 500µl of the overnight culture. A 1 ml sample was removed and stored at 4°C for use as a blank. The culture was grown to early log phase, 0.45 $O.D._{600}$ checking the O.D. every 30 minutes. The cells were transferred to a prechilled 50ml conical tube and incubated on ice for 15 minutes. The cells were centrifuged at 4°C for 7 minutes at 4,000 x G, then as much of the supernatant was removed as possible. The pellet was gently resuspended in 25 ml (1/2 volume)

0.1M ice-cold $CaCl_2$ and incubated in a ice water bath for 2 hours, then centrifuged at 4°C for 7 minutes at 4,000 x G. The cells were resuspended in 5 ml (1/10 volume) with 0.1M ice-cold $CaCl_2$-glycerol and incubated in a ice water bath for 10 minutes. Then aliquoted at 100µl to prechilled 0.5 ml microfuge tubes using chilled tips, with the cells kept on ice at all times. The tubes were incubated on ice for 30 minutes, then quickly stored at -70°C.

XIII. Recombinant Plasmid Insert Sequencing and Gene Construction

Recombinant pT7Blue-3 plasmids containing an insert of the appropriate size, as determined by restriction analysis as described above, were harvested using the QIAprep plasmid miniprep kit (QIAGEN). Host cells containing the desired recombinant plasmid were inoculated into 3 ml LB-100 µg/ml Amp broth and grown for 14 to 16 hours. The cells were pelleted by centrifugation in a Beckman Microfuge lite centrifuge at maximum speed for 1 minute. The pellet was resuspended and the plasmid was purified according to the manufacturers recommendations.

The DNA concentration and purity was determined using a Beckman DU 530 spectrophotometer. The DNA purity was measured using the 260/280 OD ratio, and any sample with a ratio between 1.8 and 1.95 was considered acceptable, then 1 µg was speedvac dried (Savant Integrated Speedvac System ISS110) in a 0.5 ml tube and sent for sequencing. Generally, a 1:100 dilution was used to determine the DNA concentration. This number was converted to µg/µl, then divided into 1 to determine the number of microliters to lyophilize.

Sequencing was performed at the University of Arkansas Health Sciences Center and the University of Kansas Medical Center. Cycle sequencing was performed with the dRhodamine Dye Terminator Cycle Sequencing Kit (Perkin Elmer/ABI) and gel (6M Urea, 9.5% Long ranger) electrophoresed on a model 377

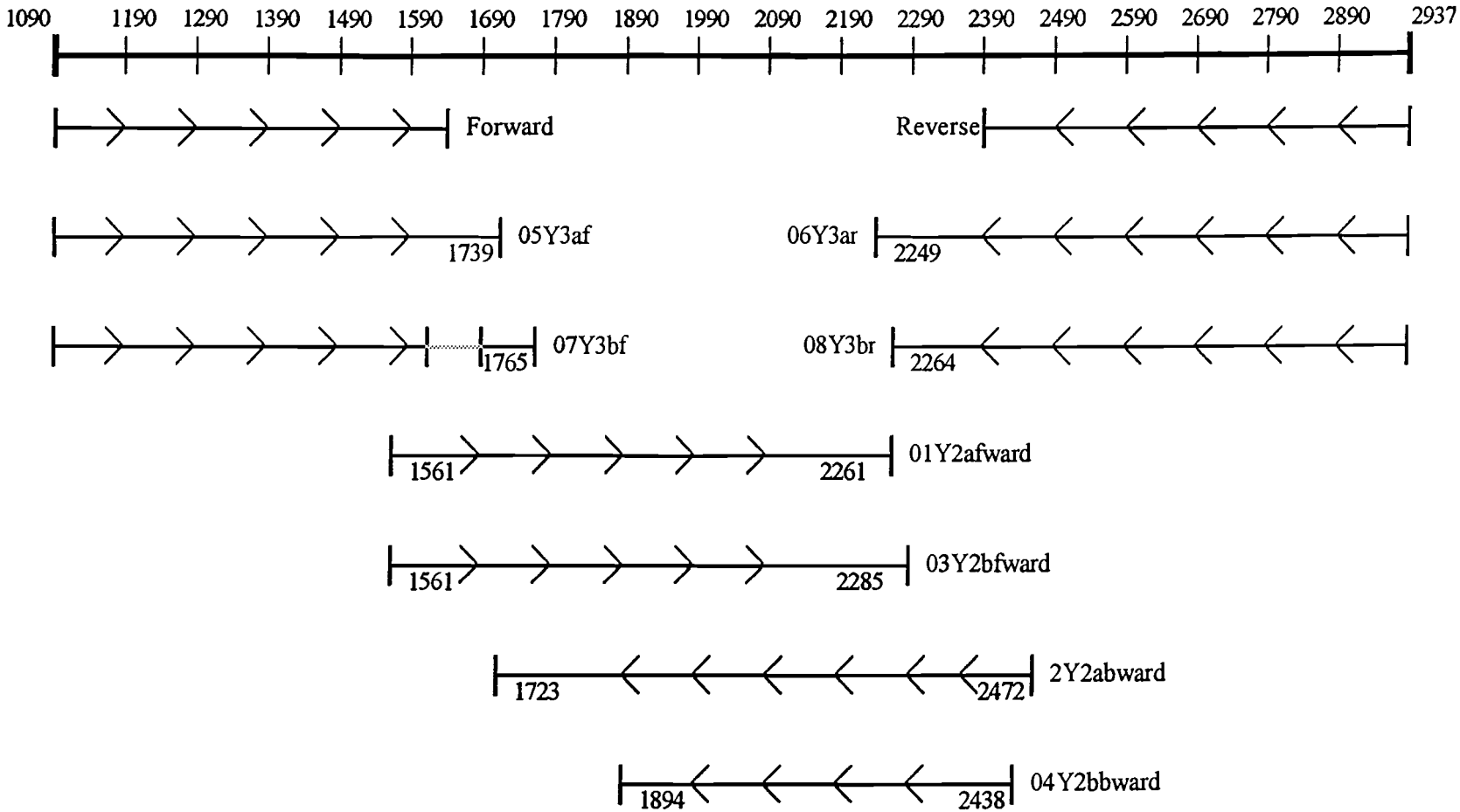
DNA sequencer (Perkin Elmer/ABI). Multiple sequencing runs on two different clones were used, as described below, to determine the consensus sequence for each *secD* gene.

XIV. *S. choleraesuis* Gene Construction

The *S. choleraesuis* gene was constructed and shown in figure 2. Briefly, sequencing was performed in triplicate on the N, and C, terminal regions of the gene. Following multiple sequencing runs for each nucleotide position, each sequence was compared to the other sequencing runs at each position to construct the gene using the NIH web site BLAST software "compare two gene sequences" (Altschul et al., 1997). The reverse complement sequence conversion was performed using the now defunct FASTA server (http://watson.genes.nig.ac.jp:8080/homology/fasta-e.shtml). These sequences were analyzed using the BLAST program "blastn" to remove the plasmid vector sequence portions at the end of the gene sequence (Altschul et al., 1997). The average number of bases which could be accurately read was approximately 650 for each primer. This meant only approximately 1,300 bases of the estimated 1,847 bases of the gene were read, 650 from the N-terminus and 650 from the C-terminus of the gene. To construct the full gene nested primers were designed from the N- and C- terminal sequences proximal to the missing region. The nested primer set consisted of the FWARD primer 5' CTC CAG GAA CAA AAT ATC GAT AGC C 3' and the reverse complement BWARD primer 5' AAT CAG ACC AAA CTT CTT ATA GAA GA 3'. These were used to sequence the interior of the gene using two sequencing runs from each nested primer, giving the forward and reverse complement. Each sequencing run was compared to the *E. coli secD* gene to establish its place in the sequence and to every other sequence with which it corresponded or overlapped to ensure proper gene construction. The gene comparison shown in results section figure 10 was translated into FASTA and formatted with CLUSTALw at the Leiden internet

FIGURE 2. The *S. choleraesuis secD* consensus gene sequence was constructed by comparing at least three single strand sequencing runs for each base position. The forward 05Y3af and 07Y3bf sequences were compared to produce the 1090 to 1723 gene sequence. The reverse 06Y3ar and 08Y3br sequences were compared to produce the 2937 to 2477 gene sequence. The missing middle sequence was produced by comparing the 01Y2afward, 03Y2bfward, 2y2abward, and 04Y2bbward sequences. Where the middle sequence overlapped the constructed terminal sequences, the two constructs were compared and this resulted in 100% homology. The gray section of sequence 07Y3bf was an unreadable section; both coding and non-coding strands were sequenced, the arrows on each strand indicate which strand was sequenced.

S. choleraesuis SecD gene sequence analysis:



site. It was constructed using the ISREC Boxshade 3.21 server (http://www.isrec.isb-sib.ch:8080/software/ BOX_form.html) with code written by K. Hofmann and M. Baron. The following format was used: Input sequence other, Output format RTF_New, Consensus line symbols, Font 12, Fraction sequences for shading 0.5. This boxshade output was converted to a RTF file and processed in Word 2000.

XV. *secD* Gene Construction from other *Enterobacteriaceae*

All other secD genes were constructed in a manner similar to that shown in Figure 2. The other genes were compared to the *S. choleraesuis* gene because they were more similar to it than to the *E. coli secD* gene. Sequencing of the internal region of genes with mismatches in nested primer to template used an increased primer concentration of 10pmol/µl and addition of DMSO to many of the internal DNA sequencing runs to allow the nested primers to work. The Leiden Genome Technology Centre site (http://ruly70.medfac.leidenuniv.nl/~gtc/links.html) was used to translate the gene sequences into FASTA format and to translate the reverse complement sequences. ClustalW was used to compare the analogous segments from this site and the European Bioinformatics Institute EMBL Outstation (http://www2.ebi.ac.uk), instead of comparing the sequences two at a time, a base at a time, as was done for the *S. choleraesuis* gene.

XVI. Amino Acid Sequence Translation

Further, each constructed gene sequence was translated to its amino acid sequence using the ExPASy site internet software from the Swiss Institute of Bioinformatics (SIB) (Bairoch and Apweiler, 1998). This software translates the nucleotide sequence in all 6 frames to its amino acid sequence. There are a total of six frames in the DNA duplex, three on each strand; the amino acid sequence

was determined by translating the DNA sequence in each frame. The triplet codons of each strand were translated in all three frames into possible amino acid sequences. To find the proper reading frame all translated sequences were examined for start and stop codons. The stop codons in the DNA are TAA, TAG, and TGA. The open reading frame that codes for the protein should not have a stop codon until the end of the sequence, whereas the other five frames usually have stop codons throughout the sequence. The translated sequences are examined and the longest open reading frame is generally assumed to be the correct reading frame. This proper reading frame is converted to FASTA format. The amino acid sequence was compared to the *E. coli* SecD amino acid sequence using the NIH BLAST software "blastp" (Altschul et al., 1997), to assess the amino acid sequence homology and to search the database for other similar proteins.

Transmembrane regions for each protein were predicted using multiple prediction servers, the results were compared, and a consensus was constructed segment by segment for each protein. Two servers were relied on for prediction and compared to additional programs. HMMTOP prediction of transmembrane helices and topology of proteins (Tusnady and Simon., 1998) based on the maximum divergence of amino acid distributions in five defined structural parts of the protein using the hidden Markov model. The authors state initial tests showed correct transmembrane prediction was 90% (143/158 proteins) and membrane spans plus topology was correct 85% (135/158 proteins) of the time (Tusnady and Simon., 1998). The second was PHDhtm which predicts transmembrane helices and topology by multiple sequence comparison to build a consensus prediction. Two servers from the Stockholm University theoretical chemistry protein prediction servers platform were used in transmembrane predictions: TopPred 2 topology prediction of membrane proteins and the Dense Alignment Surface method (DAS) for prediction of transmembrane alpha-helices in prokaryotic membrane proteins. TMpred prediction of transmembrane regions

and orientation and TMHMM prediction of transmembrane helices were also used to increase confidence in the overall prediction for each protein.

XVII. Expression Vector Construction

PCR was performed on the pT7Blue-3 recombinant plasmid which had already been sequenced, as previously described in section IV. The forward primer *Sac*I-SecD 5' CCC GAG CTC GTG TTA AAC CGT TAT CCT TTG TGG AAG 3' was designed using the *E. coli secD* forward primer with a *Sac*I restriction site added at the 5 prime end. The reverse primer was the standard *E. coli secD* reverse primer used on the *S. choleraesuis* chromosome. The amplified product at approximately 1.85Kb was gel purified, cloned into pT7Blue-3, transformed into DH5α, colony purified, restriction enzyme digested, plasmid miniprep performed, and the concentration determined as previously described.

The recombinant pT7Blue-3 (3.25 μg) was added to a 1.5 ml microfuge tube and restriction digested with *Sac*I in a 50μl reaction overnight as previously described. The restriction enzyme was inactivated by heating at 65°C for 20 minutes. The tube was filled with 400μl ice cold 95% ethanol and centrifuged for 10 minutes at 4°C. The ethanol was decanted and the pellet was speedvac dried. The pellet was resuspended in 12μl of ddH_2O, then 1.5μl of 10X restriction digestion buffer and 1.5μl of *Hind*III was added. The tube was mixed and incubated overnight at 37°C, then heated at 65°C for 20 minutes. The tube was filled to volume with ice cold 70% ethanol, incubated on ice for 10 minutes, and centrifuged for 10 minutes at 4°C. The ethanol was decanted and the pellet was speedvac dried. The pellet representing approximately 1μg of the *Sac*I-*secD* plasmid segment was resuspended in 20μl of ddH_2O.

This product was cloned into a histidine expression vector for overexpression of the *S. choleraesuis* SecD protein fused to six histidine

residues at its amino terminal end. The histidine tag was used to isolate the protein by affinity column chromatography using a nickel column to collect the protein and to detect the fusion protein on SDS-PAGE. The plasmid expression vector used is dependent on maintaining the proper reading frame of the gene. The expression vector is supplied in all three reading frames to facilitate this by the inclusion of one, two, or three bases anterior to the his tag. Selecting the proper vector is essential to producing a functional protein.

Overexpression of the fusion protein must be repressed and this is done by the pREP4 repressor plasmid which is a low copy number plasmid that constitutively expresses high levels of the *lac* repressor protein from the *lac*I gene. It was used because it represses the pQE series expression plasmids very high transcription rate from the T5 promoter. The *lac* repressor protein binds to the operator sequences and tightly represses the recombinant proteins expression. Protein production is induced by isopropyl-B-D-thiogalactoside (IPTG), it binds to and inactivates the repressor protein (Qiagen, 1999).

The pQE30 plasmid (1µg) was restriction digested with *Sac*I in a 30µl reaction, and *Hind*III as described above. The gene was ligated into the cut pQE30 according to the following protocol. A total of five microcentrifuge tubes were used, two negative controls and three dilutions of the experimental with 15µl total reactions. All of the tubes received 1µl of the digested pQE30 plasmid and 1.5µl of ligation buffer. In addition, negative control 1 received 0µl of the *secD* gene, 12µl of ddH_2O, and 0.5µl of ligase. Negative control 2 received 1µl of the *secD* gene, 11.5µl of ddH_2O, and 0µl of ligase. Experimental dilution 1 received 1µl of the *secD* gene, 11µl of ddH_2O, and 0.5µl of ligase. Dilution 2 received 3µl of the *secD* gene, 9µl of ddH_2O, and 0.5µl of ligase. Dilution 3 received 5µl of the *secD* gene, 7µl of ddH_2O, and 0.5µl of ligase. All tubes were incubated for 3 hours at 22°C using the thermocycler program "Cloning". The

recombinant plasmids were stored on ice until transformed into *E. coli* DH5α competent cells. Following transformation into DH5α, the cells were spread plated and grown overnight, colony purified by restreaking for isolation and grown overnight, plasmid miniprep was performed, and restriction enzyme digested with *BamH*I and *Hind*III to check for proper insert size as previously described. Two plasmids containing the proper insert were transformed, as previously described, into DH5α containing the repressor plasmid pREP4.

The transformed cells were spread plated onto LB-100μg/ml Amp-25μg/ml Kan plates and grown overnight at 37°C. The 5 plates of each set were compared and the second dilution (3μl *secD*) was determined to be the most efficient, so this dilution was used in all subsequent pQE30 ligations.

Eight colonies were selected and colony purified on LB-100μg/ml Amp-25μg/ml Kan plates. They were inoculated into 3ml LB-100μg/ml Amp-25μg/ml Kan broth and plasmid miniprep was performed. To check for proper insertion the plasmids were restriction digested with *EcoR*I, *BamH*I, and *Hind*III as stated above and shown in the results (figure 5). *Hind*III cuts the pREP4 plasmid causing a band at approximately 3.8Kb. Six recombinant plasmids were retained for protein expression.

The first attempt to overexpress the *S. choleraesuis* SecD protein, used the complete gene including the leader sequence. This approach did not yield good results, possibly because the histidine tag was cleaved along with the proteins' leader sequence from a majority of the protein expressed. The next approach attempted used the *S. choleraesuis secD* gene truncated by cleaving off the N- terminal 93 bases (including the leader sequence) with *BamH*I. This gene fragment was gel purified and cloned into pT7Blue-3, restriction digested to check for proper insertion as described, then the plasmid was restriction digested with *BamH*I and *Hind*III. This truncated *secD* gene was cloned into all three

frames pQE30, pQE31, and pQE32 of the expression vector plasmid, then protein expression was attempted with all of the recombinant clones.

XVIII. Protein Expression

Six pQE30-SecD clones were analyzed for protein expression labeled as: 3-3, 3-3b, 3-4, 3-4x, 2-5, and 12-5. Tubes containing 3ml LB100μg/ml Amp-25μg/ml Kan broth were inoculated and incubated for 14 hours at 37°C with shaking at 250 rpm. Then 500μl of the overnight culture was inoculated into a prewarmed (30 minutes) 125ml flask containing 12.5 ml LB100μg/ml Amp-25μg/ml Kan broth after 0.5ml was removed to use as a blank for subsequent spectrograph measurements. The solution was incubated until an $O.D._{600}$ of 0.5 was reached (~3 hours), checking the O.D. reading every 30 minutes. The cells were induced by adding IPTG to a final concentration of 1μM, after 4 ml was removed to use as an un induced control. Incubation continued for both the experimental and control for 4 hours at 37°C with shaking at 250 rpm. The cells were pelleted at 4,000 rpm in a Beckman tabletop centrifuge, the supernatant was removed, and the pellet was resuspended in lysing buffer. The cells were frozen and thawed for 2 cycles, then sonicated for 8 to 10 minutes (or until clear) in an alcohol-ice bath. Each sample was centrifuged at 4°C for 30 minutes to remove cell debris, then stored at -20°C until needed. The samples were tested for fusion protein production using protocols number 4 and 16 described in The QIAexpressionist: A handbook for high-level expression and purification of 6xHis-tagged proteins (QIAGEN, March 1999).

XIX. Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Analysis of the putative SecD protein of *S. choleraesuis* was carried out under both native and denaturing conditions. All samples were heated at 95°C for 5 minutes before being loaded into the gel. The EC120 Mini vertical gel

system (E-C Apparatus Corporation, Holbrook, NY) was used according to the manufacturer's recommendations for a 1mm thick 10% discontinuous polyacrylamide gel.

XX. SecD Mutant Complementation

The cold sensitive (Cs) SecD mutant KJ173 was complemented with the recombinant expression vector pQE30:12-5 and pQE30:3-3. The experimental protocol was, 1μg of plasmid DNA added to each 0.5 ml tube containing 100μl of SecD mutant competent cells except the SecD mutant only tube. The control tubes received no plasmid, 1μg pQE30 only, or 1μg pBluescript into 100μl of SecD mutant competent cells. The competent cells were transformed as described in the thermocycler using the "BAD" program, then 1 mM IPTG (final concentration) and 1000μl LB broth was added to each tube. The tubes were incubated for 6 hours at 22°C, and 100μl of each was spread plated onto 2 plates containing the appropriate antibiotic: no plasmid on BHI 200μg/ml Strep, with plasmid on BHI 200μg/ml Strep 100μg/ml Amp. These were grown in a low temperature incubator (Precision model 815, GCA corporation) for 18 hours at 22°C. The pQE30:12-5 construct showed better complementation and was used in the broth study below.

The Cs SecD mutant KJ173 was also complemented with pQE30:12-5 in the following broth study. The negative control plasmids were pQE31 only and recombinant pQE16. The recombinant plasmid pQE16 produces a soluble protein unrelated to protein secretion, it was isolated as described above. Plasmid DNA (1μg) was added to 0.5 ml tubes containing 100μl of SecD mutant competent cells each. Three tubes received the pQE16 plasmid, three tubes received the pQE31 plasmid, and four tubes received the pQE30:12-5 plasmid. The competent cells were transformed as described in the thermocycler using the "BAD" program,

then 1 mM IPTG (final concentration) and 1 ml LB broth was added to each tube. The tubes were incubated for 6 hours at 21°C with shaking at speed four. After this incubation 3µl of transformed cells were inoculated into 3ml LB 200µg/ml Strep 100µg/ml Amp tubes and these were incubated at 21°C with shaking at speed four for 48 hours. The absorbance at $O.D._{600}$ for each tube was measured by removing a 1ml sample from each tube and taking two readings, then averaging the result if they differed. The pQE31 negative control results were averaged and subtracted from each experimental result, then that number was divided by the mean of the negative control to give the percent increase in growth over the negative control.

Results

I. PCR Optimization

Initial PCRs were performed on the *E. coli* DH5α chromosome to develop a standard protocol for use in subsequent attempts to amplify the *secD* gene from other members of the *Enterobacteriaceae*. The Mg^{++} concentration of the reaction was optimized by testing from 1 to 4 mM. The optimum concentration was found to be 2mM. Taq and Vent polymerase were compared, and Vent performed better as shown in figure 3.

II. Restriction Digest of the Recombinant pT7Blue-3 Plasmid

Proper insert of the putative *secD* gene was visualized using a 1.2% agarose gel. Isolated recombinant plasmids were restriction digested with *EcoR*I (figure 4), *Kpn*I, and *BamH*I (figure 5) to show proper insertion and proper size of the insert, as shown in figure 1. Recombinant clones with the proper gene insert were sequenced and frozen at -70°C for long term storage. A representative restriction digest is shown for each in figures 4 and 5.

III. *Enterobacteriaceae* Cloned *secD* Genes

The constructed *secD* gene sequences of *S. choleraesuis*, *K. pneumoniae*, *C. freundii*, and *E. aerogenes* are shown in figures 6 to 9. The sequenced genes have an average homology of about 87%, as shown in table 3.

IV. Genetic Comparison of *S. choleraesuis secD* to *E. coli secD*

The sequenced *S. choleraesuis secD* gene was compared to the *E. coli secD* gene to determine if the two genes were similar. This sequence comparison was done using the BLAST software "blastn" from NIH on the internet (Altschul et al., 1997) and shown in table 3. The gene comparison results shown in figure

FIGURE 3. PCR of *S. choleraesuis secD* using the *E. coli secD* primer set. Lane 11 is the negative control. Lanes 8 and 10 are Taq polymerase plus 2mM magnesium. Lanes 7 and 9 are Taq plus 3mM magnesium. Lane 6 is empty. Lane 3 and 5 are Vent polymerase plus 2mM magnesium. Lanes 2 and 4 are Vent plus 3mM magnesium. Lane 1 is empty.

Figure 3. PCR of S. choleraesuis SecD

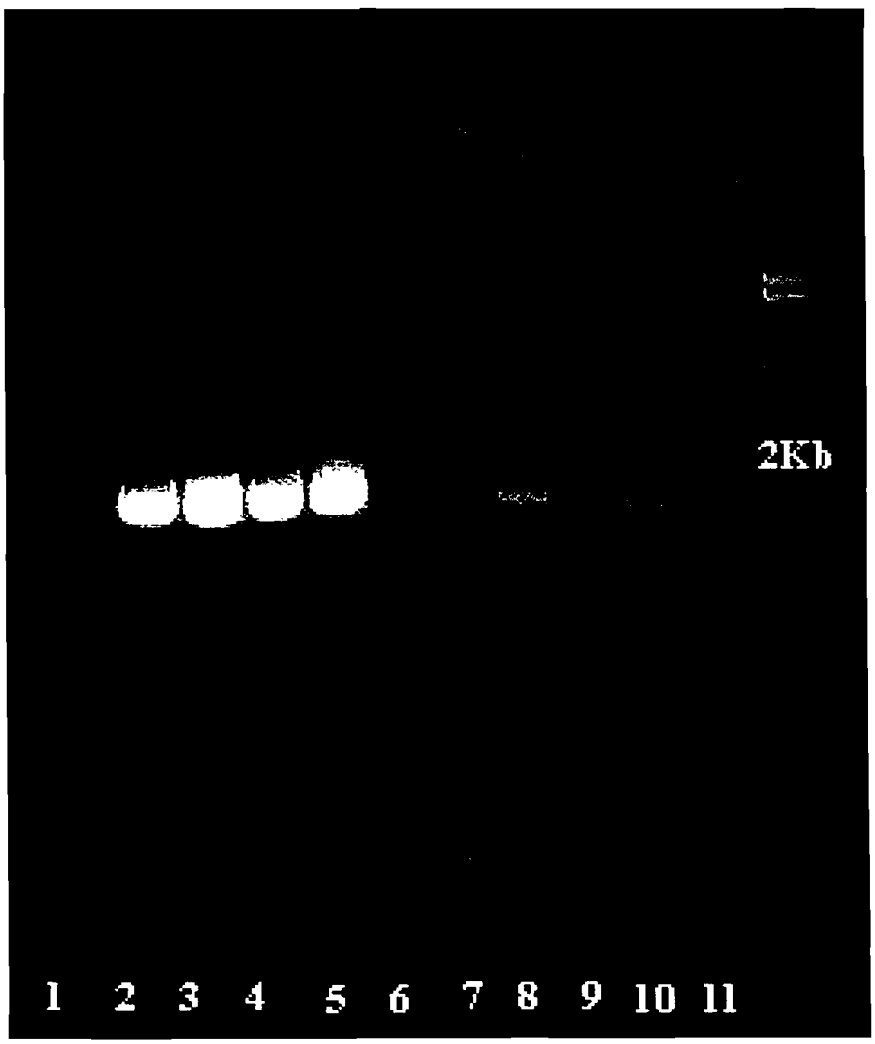


FIGURE 4. Restriction digestion of recombinant pT7Blue-3:secD clones. Lanes 1 through 4 are empty. Lane 5 contains the recombinant plasmid isolated from colony 8 digested with *EcoR*I. Lane 6 contains the recombinant plasmid isolated from colony 1 digested with *EcoR*I. Lane 7 contains the pT7Blue-3 plasmid isolated from a blue colony digested with *EcoR*I.

Figure 4. Restriction digestion of recombinant clones

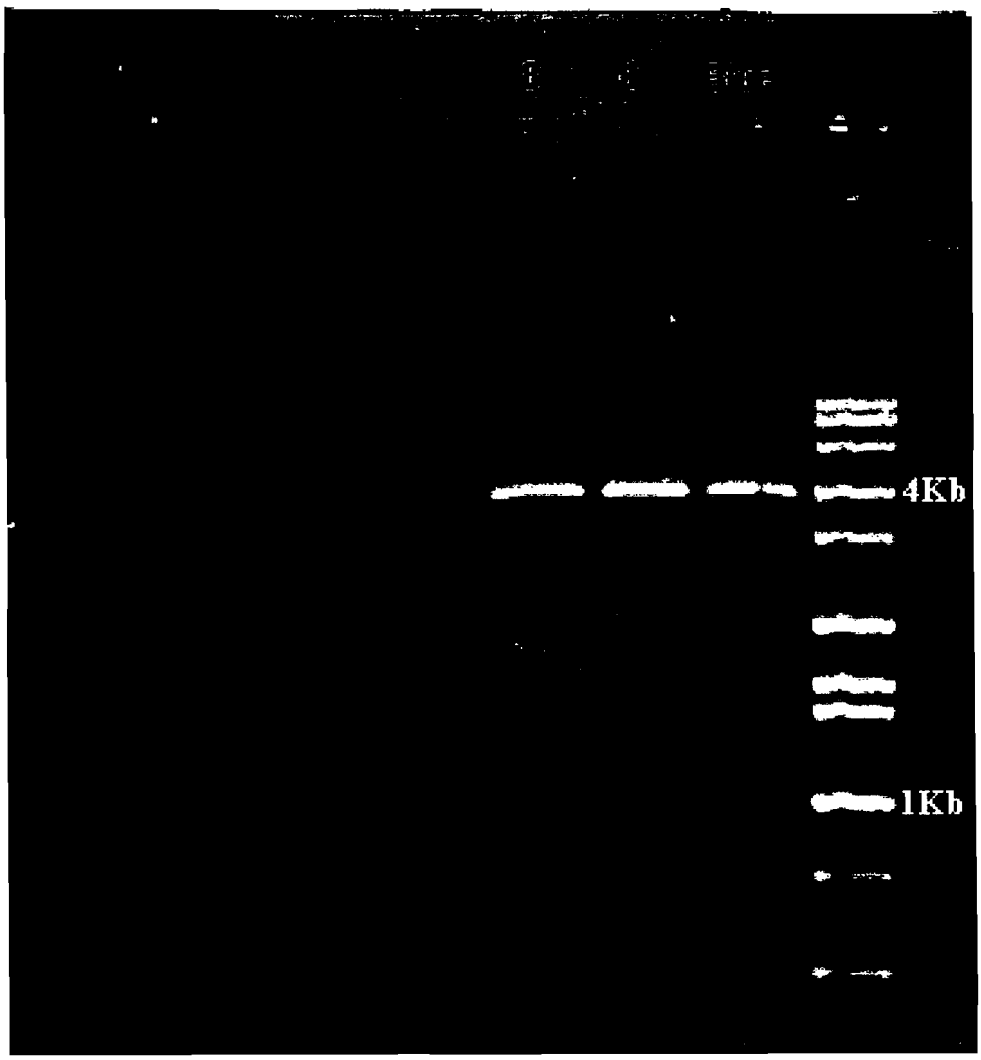


FIGURE 5. Restriction digestion of recombinant pT7Blue-3:secD clone isolated from colony 1. Lanes 1 through 4 are empty. Lane 5 contains the recombinant plasmid digested with *Kpn*I. Lane 6 contains the recombinant plasmid digested with *BamH*I. Lane 7 is empty.

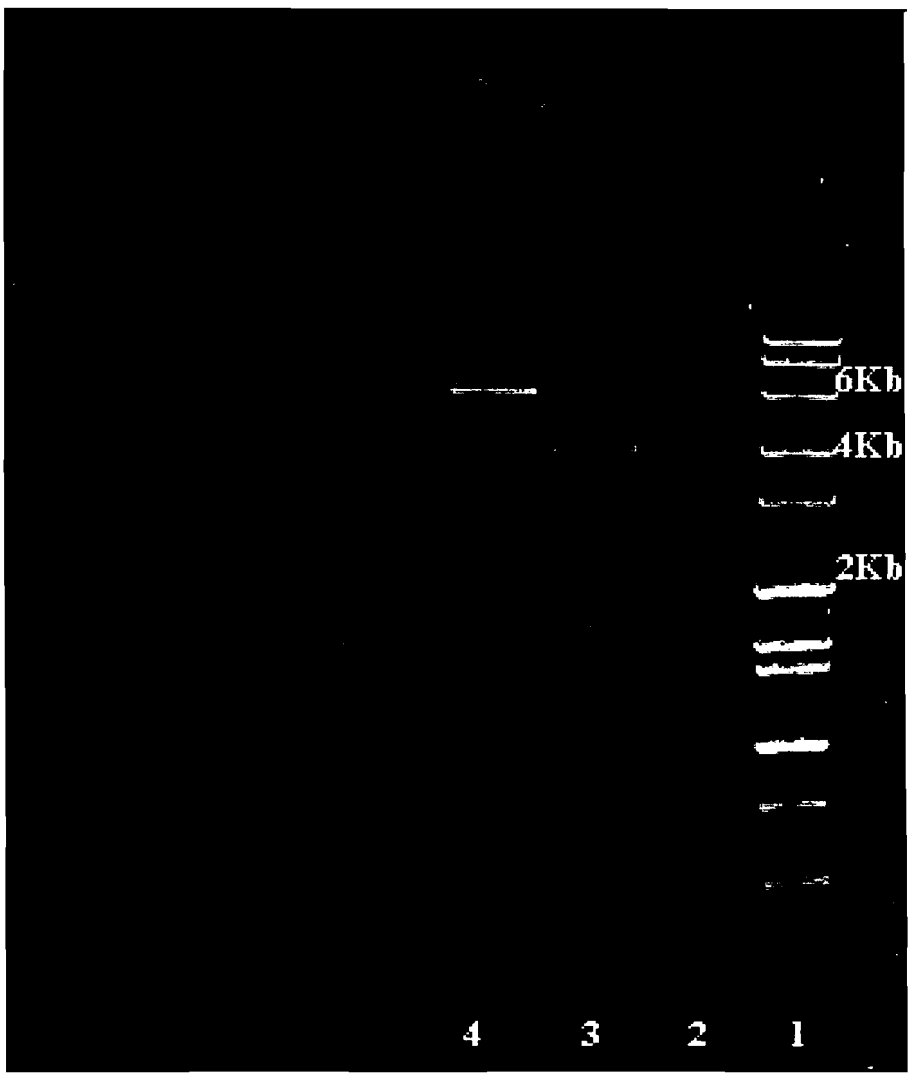

6Kb
4Kb
2Kb
4
3
2
1

FIGURE 6. Nucleotide sequence of the *Salmonella choleraesuis secD* gene. Only one strand is shown. The GenBank accession number for this sequence is: AF100611.

FIGURE 6. *Salmonella choleraesuis secD*

```
   1 GTGTTAAACCGTTATCCTTTGTGGAAGTACATCATGCTGGTCGTCGTGAT
  51 TATCGTCGGCCTGCTTTACGCACTTCCCAACCTGTATGGTGAGGATCCGG
 101 CTGTTCAAATCACTGGCGTGCGCGGCGTCGCCGCCAGTGAGCAAACGCTG
 151 ATCCAGGTCCAGAAAACGTTACAAGAAGAAAAAATTCCCGCTAAGTCTGT
 201 GGCACTGGAAGAGGGCGCTATTCTTGCGCGCTTCGACACCACCGATACCC
 251 AGCTACGCGCCCGTGAAGCGCTGATGAGCGTGCTGGGGGATAAATATGTC
 301 GTGGCGCTTAACCTTGCGCCGGCTACGCCGCGCTGGTTGGCGGCTATTCA
 351 CGCCGATCCGATGAAGCTCGGTCTTGATCTGCGCGGCGGCGTTCACTTCC
 401 TGATGGAAGTGGATATGGATACCGCGTTAGGTAAACTCCAGGAACAAAAT
 451 ATCGATAGCCTGCGCAGCGATCTGCGTGAAAAAGGCATTCCTTATACCAC
 501 CGTTCGTAAAGAAAATAATTATGGGTTGAGCATCACGTTCCGTGACAGCA
 551 AAGCGCGTGATGAGGCGATCGCCTATCTGACGCCGCGTCATCGTGACCTG
 601 GTGATTTCCAGCCAGAGCGGCAATCAACTTCGCGCGGTAATGACCGATGC
 651 GCGCCTGAGCGAAGCGCGTGAATACGCCGTTCAGCAGAACATTAATATTC
 701 TGCGTAACCGTGTTAACCAACTGGGCGTCGCGGAGCCTGTCGTACAACGC
 751 CAGGGCGCAGACCGTATTGTGGTTGAGCTGCCGGGTATCCAGGATACCGC
 801 TCGCGCGAAAGAAATTCTCGGCGCGACCGCGACGCTGGAATTCCGTCTGG
 851 TGAATACCAACGTTGACCAGGCCGCCGCTGCTGCCGGGCGCGTACCGGGT
 901 GATTCCGAAGTGAAACAGACCCGCGAAGGTCAGCCGGTGGTATTGTACAA
 951 ACGCGTGATCCTCACCGGCGACCATATCACCGACTCGACGTCCAGCCAGG
1001 ATGAATATAACCAACCGCAGGTTAACATCTCGCTGGATAGCGCAGGCGGC
1051 AATATCATGTCTAACTTCACTAAGGACAATATCGGCAAACCGATGGCGAC
1101 CCTGTTTGTGGAGTATAAAGACAGCGGTAAGAAAGACGCTAACGGTCGCG
1151 CCGTACTGGTGAAACAGGAAGAAGTGATTAACATTGCCAACATCCAGTCC
1201 CGGCTGGGTAACAGCTTCCGTATTACCGGTATCAGCAACCCGAACGAAGC
1251 GCGTCAGCTCTCGTTGTTGCTGCGTGCGGGCGCCCTGATTGCGCCGATTC
1301 AGATCGTTGAAGAGCGCACCATTGGCCCGACTCTGGGTATGCAGAACATC
1351 AAGCAGGGCCTGGAAGCCTGTCTGGCCGGTCTGGTGGTCTCCATCCTGTT
1401 CATGATCTTCTTCTATAAGAAGTTTGGTCTGATTGCGACCAGCGCGTTAG
1451 TGGCTAACCTGGTGTTGATTGTCGGTATCATGTCGCTGTTGCCGGGGGCA
1501 ACGCTCAGTATGCCGGGGATTGCGGGTATCGTGTTAACCCTTGCGGTGGC
1551 GGTCGACGCCAACGTTCTGATTAACGAACGTATCAAAGAAGAGTTGAGTA
1601 ATGGCCGCACGGTACAGCAGGCGATAAACGAAGGCTATGCCGGCGCGTTC
1651 AGTTCCATCTTCGATGCGAACATCACGACGCTGATTAAAGTCATCATTCT
1701 GTATGCTGTCGGTACCGGGGCAATTAAAGGGTTCGCGATTACGACCGGTA
1751 TCGGTGTGGCGACGTCGATGTTTACCGCGATTATCGGTACGCGTGCTATC
1801 GTAAACCTGCTATATGGCGGCAAGCGCGTCAAGAAGGTGTCAATCTGA
```

FIGURE 7. Nucleotide sequence of the *Enterobacter aerogenes secD* gene. Only one strand is shown. The GenBank accession nuber for this sequence is: AF163861.

FIGURE 7. *Enterobacter aerogenes secD*

```
   1 GTGTTAAACCGTTATCCTTTGTGGAAGTACGTTATGCTGGTCGTCGTGAT
  51 TATCGTCGGCCTGATTTACGCGCTTCCCAACCTGTATGGTGAGGATCCGG
 101 CTGTTCAAATCACTGGCGCGCGCGGCGTCGCCGCCAGTGAGCAAACGCTG
 151 ATCCAGGTCCAGAAAACTTTACAAGAAGAAAAAATCACCGCGAAGTCTGT
 201 GGCACTGGAAGAGGGCGCAATTCTTGCGCGCTTCGACACCACCGACACTC
 251 AGCTGCGAGCGCGTGAAGCGCTGGTTAACCTGCTGGGTGACAAATACGTC
 301 GTGGCGCTAAACCTTGCTCCGGCGACCCCGCGCTGGCTGGCGGCGATGTA
 351 TGCCGAGCCGATGAAGCTGGGTCTTGACCTACGCGGCGGCGTGCACTTCC
 401 TGATGGAAGTGGACATGGACACCGCATTAGGCAAACTGCAGGAACAAAAC
 451 ATCGATAGCCTGCGCAGCGAGCTGCGTGACAAAGGCATTCCTTACTCCAC
 501 CGTGCGTAAGGAAGACAACTACGGTCTGAGCATCGTCTTTCGCGACAGCG
 551 CGGCGCGCGATCAGGCTATCTCTTATCTCAGCCCGCGCCATCGCGATCTG
 601 GTTATTTCAAGCCAGGGTACCAATGCGCTGAAAGCGGTGATGACCGATGA
 651 ACGTCTGAAAGAAGCCCGTGAATACGCCGTTCAGCAGAACATCAACATTC
 701 TGCGTAACCGTGTAAACCAGCTCGGCGTGGCCGAACCGCTGGTGCAGCGT
 751 CAGGGTTCCGACCGTATCGTGGTTGAACTGCCGGGTATCCAGGATACCGC
 801 GCGTGCGAAGGAAATCCTTGGCGCGACCGCGACGCTGGAATTCCGTCTGG
 851 TGAACACCAACGTCGATCAGTCCGCAGCCGCTTCTGGTCGCGTACCGGGC
 901 GATTCGGAAGTGAAAGATACCCGTGAAGGCCGTCCGGTGGTGCTGTATAA
 951 GCGCGTGATCCTGACCGGCGACCATATCACCGACTCCACTTCCAGCATGG
1001 ACGAGTACAACCAGCCGCAGGTTAACATCTCGCTCGATAGCGCGGGTGGC
1051 AACATCATGTCTAACTTCACCAAGGACAATATCGGTAAACCGATGGCGAC
1101 CCTGTTCGTGGAGTACAAAGACAGCGGTAAGAAAGATGCCAACGGTCGCG
1151 CTATCCTTGCGAAAGAGGAAGAGGTGATTAACATCGCCAATATCCAGTCT
1201 CGCCTTGGCAATAGCTTCCGTATTACCGGTATCAGCAACCCGACCGAAGC
1251 GCGTCAGCTTTCTCTGCTGCTGCGTGCGGGCGCCCTGATTGCGCCGATTC
1301 AGATCGTTGAAGAGCGCACTATCGGCCCGACTCTGGGTATGCAGAACATC
1351 AAGCAGGGCCTGGAAGCGTGTCTGGCCGGTCTGGTGGTCTCTATCCTGTT
1401 CATGATCTTCTTCTATAAGAAGTTCGGTCTGATTGCGACCTCCGCGCTGA
1451 TTGCTAACCTGGTGCTGATTGTCGGTATTATGTCCCTGATTCCTGGGGCG
1501 ACGCTGACCATGCCGGGGATCGCGGGTATCGTCTTAACCCTTGCGGTGGC
1551 GGTCGATGCTAACGTACTGATTAACGAGCGTATCAAAGAAGAGCTGAGCA
1601 ACGGACGTACCGTTCAGCAGGCGATTGACGAAGGTTATAAGGGCGCATTC
1651 AGCTCCATCTTCGACGCCAACGTGACGACGCTAATTAAGGTTATCATCCT
1701 GTACGCGGTCGGTACCGGTGCCATTAAAGGGTTTGCGATTACTACCGGTA
1751 TCGGTATCGCGACCTCAATGTTTACCGCTATCGTCGGCACCCGTGCCATC
1801 GTGAACCTGCTGTACGGCGGCAAGCGCGTCAAGAAGCTGTCAATCTGA
```

FIGURE 8. Nucleotide sequence of the *Citrobacter freundii secD* gene. Only one strand is shown. The GenBank accession number for this sequence is: AF179925.

FIGURE 8. *Citrobacter freundii secD*

```
   1 GTGTTAAACCGTTATCCTTTGTGGAAGTATGTCATGCTGATCGTCGTTAT
  51 TGTCGTCGGTCTGCTCTATGCGCTTCCCAACCTGTATGGTGAGGATCCGG
 101 CTGTTCAAATCACTGGCGTGCGCGGTGCCGCCGCCAGTGAGCAAACGCTG
 151 ATCCAGGTCCAGAAAACGTTACAAGAAGATAAAATCACACCTAAGTCTGT
 201 GGCTCTGGAAGAGGGTGCTATTCTCGCGCGCTTCGACTCCACTGATACCC
 251 AACTGCGCGCACGTGAAGTATTAATGGGCGTACTGGGCGATCAGTACGTC
 301 GTGGCGCTTAACCTTGCTCCGGCAACCCCGCGCTGGCTGGCTGCCATTCA
 351 TGCAGAACCGATGAAACTCGGTCTTGACCTGCGTGGCGGCGTTCACTTCC
 401 TGATGGAAGTGGATATGGATACCGCGCTGAGTAAGTTGCAGGAACAAAAT
 451 ATCGATAGCCTGCGCAGCGATCTGCGTGAAAAAGGCATTCCTTATACTAC
 501 CGTTCGCAAAGAAGATAACTACGGTCTGAGCATCACTTTCCGTGATAGCA
 551 GCGCCCGTGATGAAGCTATCTCTTACTTGAGCAAACGTCACCAGGATCTG
 601 GTTATTTCCAGCCAGGGCAGCAATGCGCTGCGCGCTGTAATGACCGATGC
 651 GCGCCTGAGTGAAGCCCGTGAGTACGCGGTACAGCAGAACATCAACATTC
 701 TGCGTAATCGTGTTAACCAACTGGGTGTTGCTGAGCCGGTGGTACAACGT
 751 CAAGGTGCTGACCGCATCGTGGTCGAACTGCCGGGTATTCAGGATACCGC
 801 GCGTGCGAAAGAAATTCTGGGCGCAACCGCGACGCTGGAGTTCCGTCTGG
 851 TAAACACCAACGTTGACCAGTCCGCTGCAGCATCGGGTCGTGTTCCGGGT
 901 GATTCGGAAGTGAAACAGTCTCGTGAAGGTCAGCCTGTTGTGCTGTACAA
 951 GCGCGTGATCCTGACCGGTGACCATATCACCGACTCCACTTCCAGCCAGG
1001 ACGAATACAATCAGCCGCAGGTTAACATCTCGCTGGATAGCGCGGGTGGT
1051 AACATCATGTCCAACTTCACCAAGGACAACATCGGTAAGCCGATGGCGAC
1101 CCTGTTCGTGGAGTACAAAGACAGCGGTAAGAAAGATGCTAACGGTCGTG
1151 CAGTACTGGTGAAACAGGAAGAGGTGATTAACATTGCCAATATCCAGTCT
1201 CGCCTGGGTAACAGCTTCCGTATCACCGGTATCAACAACCCGAACGAAGC
1251 GCGTCAGCTTTCTCTGCTGTTACGTGCCGGTGCGTTGATTGCGCCGATTC
1301 AGATTGTGGAAGAACGTACCATCGGCCCAACCCTGGGTATGCAGAACATT
1351 AAACAGGGTCTGGAAGCGTGTCTGGCCGGTCTGGTGGTGTCTATCCTGTT
1401 CATGATCTTCTTCTATAAGAAGTTCGGTCTGATTGCGACCAGTGCGCTGA
1451 TTGCTAACCTGGTATTGATTGTCGGTATTATGTCGTTGCTGCCGGGCGCT
1501 ACGCTAAGTATGCCGGGGATTGCGGGGATTGTCTTAACCCTTGCCGTCGC
1551 GGTCGATGCGAACGTACTGATCAACGAACGTATCAAAGAAGAGTTGAGTA
1601 ACGGGCGTTCGGTGCAGCAAGCCATCAACGAAGGTTATGCGGGCGCATTT
1651 AGCTCCATCTTTGATGCGAACATTACGACGCTGATCAAGGTCATCATCCT
1701 GTACGCAGTTGGCACCGGGGCAATTAAAGGGTTCGCGATTACTACCGGTA
1751 TTGGTGTGGCGACGTCGATGTTTACCGCGATTGTCGGTACACGTGCCATC
1801 GTAAACCTGTTGTATGGCGGCAAGCGCGTCAAGAAGCTGTCAATCTGA
```

FIGURE 9. Nucleotide sequence of the *Klebsiella pneumoniae secD* gene. Only one strand is shown. A GenBank accession number for this sequence has been applied for.

FIGURE 9. *Klebsiella pneumoniae secD*

```
   1 GTGTTAAACCGTTATCCTTTGTGGAAGTACGTCATGCTGGTCGTCGTCAT
  51 TGTCGTCGGCCTGATCTATGCGCTTCCCAACCTGTATGGTGAGGATCCGG
 101 CTGTTCAAATCACTGGCGCGCGCGGCGTCGCCGCCAGTGAGCAAACGCTG
 151 ATCCAGGTCCAGAAAACTTTACAAGAAGAAAAAATAACCGCGAAGTCTGT
 201 GGCACTGGAAGAGGGCGCTATTCTTGCGCGCTTCGACACCACTGATACCC
 251 AGCTGCGCGCGCGCGAAGCGCTGCTTAACGTGCTGGGTGACAAATACGTC
 301 GTGGCGCTGAACCTTGCTCCGGCAACCCCACGCTGGCTGGCGGCGCTCTA
 351 TGCCGAGCCGATGAAGCTGGGTCTTGATCTGCGCGGCGGCGTGCACTTCC
 401 TGATGGAAGTCGACATGGACACCGCGTTAGGCAAACTGCAGGAACAGAAC
 451 ATCGATAGCCTGCGCAGCGAGCTGCGCGACAAAGGCATTCCTTACGCGAC
 501 CGTGCGTAAAGAAGACAACTACGGTCTGAGCATCGTCTTCCGCGACAGCG
 551 CGGCGCGCGACCAGGCTATCTCTTATCTCAGCCCTCGCCATCGCGATCTG
 601 GTTATTTCCTCTCAGGGTGACAACAGCCTGAAAGCGGTGATGACCGATGA
 651 GCGCCTGAAAGAAGCTCGTGAGTACGCGGTTCAGCAGAACATTAATATCC
 701 TGCGTAACCGTGTGAACCAGCTGGGCGTCGCCGAGCCGCTGGTTCAGCGT
 751 CAGGGCGCTGACCGCATCGTGGTTGAGCTGCCGGGTATCCAGGATACCGC
 801 GCGTGCGAAGGAAATCCTTGGCGCGACCGCGACCCTCGAGTTCCGTCTGG
 851 TCAATAATAACGTGGACCAGTCCGCCGCCGCCTCTGGCCGCGTGCCGGGC
 901 GACTCGGAAGTGAAGCAGACGCGTGAAGGCCAGCCGGTAGTGCTGTACAA
 951 GCGGGTGATCCTGACCGGGGACCATATCACCGACTCGACCTCCAGCATGG
1001 ATGAGTACAACCAGCCGCAGGTCAACATCTCGCTGGACAGCGCGGGCGGT
1051 AACATCATGTCTAACTTCACTAAGGACAACATCGGCAAGCCGATGGCGAC
1101 CCTGTTCGTGGAGTATAAAGACAGCGGTAAGAAAGACGCTAACGGTCGCG
1151 CTATCCTGGCGAAAGAGGAAGAGGTGATTAACATCGCCAACATCCAGTCG
1201 CGTCTGGGTAACAGCTTCCGTATCACCGGTATCAGCAATCCGAACGAAGC
1251 GCGTCAGCTGTCGCTGCTGCTGCGTGCCGGCGCGCTGATTGCGCCGATTC
1301 AGATCGTCGAAGAGCGGACTATCGGCCCAACCCTGGGAATGCAGAACATC
1351 AAGCAGGGTCTGGAAGCCTGTCTGGCCGGTCTGGTGGTCTCTATCCTGTT
1401 CATGATCCTCTTCTATAAGAAGTTCGGCCTGATTGCGACCTCGGCGCTGA
1451 TTGCCAACCTGATACTGATTGTCGGCATTATGTCCCTGATCCCGGGGGCG
1501 ACGCTGACCATGCCGGGCATCGCGGGTATCGTGTTAACCCTTGCGGTGGC
1551 GGTCGATGCTAACGTTCTGATCAACGAGCGTATCAAAGAAGAGCTGAGCA
1601 ACGGGCGTACCGTTCAGCAGGCGATTGACGAAGGTTATCGCGGCGCGTTC
1651 AGCTCGATCTTCGACGCCAACGTGACGACGCTTATCAAAGTGATCATCCT
1701 GTACGCGGTCGGTACCGGTGCCATTAAAGGGTTTGCGATTACCACCGGTA
1751 TCGGTATCGCGACCTCAATGTTTACCGCTATCGTCGGCACCCGTGCCATC
1801 GTGAACCTGCTGTACGGCGGCAAGCGCGTCAAGAAGCTGTCAATC
```

TABLE 3. Comparisons of nucleotide sequences of the *Enterobacteriaceae secD* gene. Each comparison contains the number of identical nucleotides divided by the total number of nucleotides giving the percent identity. The average gene identity of these homologous genes was 87%.

Table 3. Gene similarity of *secD* in the Enterobacteriaceae sequenced

	E. coli	*S. choleraesuis*	*E. aerogenes*	*C. fruendii*	*K. pneumoniae*
E. coli	-----	1607/1848= 0.87	1565/1848= 0.85	1629/1848= 0.88	1563/1845= 0.85
S. choleraesuis		-----	1603/1848= 0.87	1612/1848= 0.87	1592/1845= 0.86
E. aerogenes			-----	1589/1848= 0.86	1702/1848= 0.92
C. fruendii				-----	1576/1845= 0.85
K. pneumoniae					-----

cloning and sequencing of additional *secD* genes from other members of the family.

V. Sequence Comparison of *secD* (From Other *Enterobacteriaceae*)

The *secD* gene sequences from other *Enterobacteriaceae* were compared to the *E. coli secD* gene sequence to assess homology. Sequences were also compared to all the other gene sequences for homology. The sequence comparisons were performed using the BLAST software "blastn". The percent identity for each pair was calculated by dividing the number of identical residues by the total number of residues. The average identity was approximately 87%, the results of each are shown in table 3.

VI. Translated Amino Acid Sequences From the *Enterobacteriaceae*

The translation of each nucleotide sequence is shown in figures 10 to 13. The results of "blastp" comparison of these translated amino acid sequences from each sequence are shown in table 4 which shows an overall average 93% residue identity and 97% similar residues. Similar residues are amino acids with similar side groups, *i.e.* a polar amino acid replaces a polar amino acid. The protein shows similarity to many GTPases and kinases using the protein analysis software PredictProtein (analysis not shown). Potential transmembrane regions were predicted using a variety of TM region prediction software programs and the results are shown in table 5. The molecular weights shown in the discussion (table 5) for *E. aerogenes*, *S. choleraesuis*, and *C. freundii* were calculated with the compute Ip/Mw tool from the expert protein analysis system (expasy) proteomics server from the Swiss Institute of BioInformatics (SIB). No MW was calculated for *K. pneumoniae* due to the missing termination codon. TM analysis (TMHMM, TopPred2, and HMMTOP) also predicts the N- and C-termini orientation of each protein as cytoplasmic.

FIGURE 10. The amino acid sequence of *Salmonella choleraesuis* SecD. The accession number is AF100611.

FIGURE 10. *Salmonella choleraesuis* SecD

```
  1 MLNRYPLWKYIMLVVVIIVGLLYALPNLYGEDPAVQITGVRGVAASEQTL
 51 IQVQKTLQEEKIPAKSVALEEGAILARFDTTDTQLRAREALMSVLGDKYV
101 VALNLAPATPRWLAAIHADPMKLGLDLRGGVHFLMEVDMDTALGKLQEQN
151 IDSLRSDLREKGIPYTTVRKENNYGLSITFRDSKARDEAIAYLTPRHRDL
201 VISSQSGNQLRAVMTDARLSEAREYAVQQNINILRNRVNQLGVAEPVVQR
251 QGADRIVVELPGIQDTARAKEILGATATLEFRLVNTNVDQAAAAAGRVPG
301 DSEVKQTREGQPVVLYKRVILTGDHITDSTSSQDEYNQPQVNISLDSAGG
351 NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAVLVKQEEVINIANIQS
401 RLGNSFRITGISNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
451 KQGLEACLAGLVVSILFMIFFYKKFGLIATSALVANLVLIVGIMSLLPGA
501 TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAINEGYAGAF
551 SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIIGTRAI
601 VNLLYGGKRVKKLSI
```

FIGURE 11. The amino acid sequence of *Enterobacter aerogenes* SecD. The accession number is AF163861.

FIGURE 11. *Enterobacter aerogenes* SecD

```
  1 MLNRYPLWKYVMLVVVIIVGLIYALPNLYGEDPAVQITGARGVAASEQTL
 51 IQVQKTLQEEKITAKSVALEEGAILARFDTTDTQLRAREALVNLLGDKYV
101 VALNLAPATPRWLAAMYAEPMKLGLDLRGGVHFLMEVDMDTALGKLQEQN
151 IDSLRSELRDKGIPYSTVRKEDNYGLSIVFRDSAARDQAISYLSPRHRDL
201 VISSQGTNALKAVMTDERLKEAREYAVQQNINILRNRVNQLGVAEPLVQR
251 QGSDRIVVELPGIQDTARAKEILGATATLEFRLVNTNVDQSAAASGRVPG
301 DSEVKDTREGRPVVLYKRVILTGDHITDSTSSMDEYNQPQVNISLDSAGG
351 NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAILAKEEEVINIANIQS
401 RLGNSFRITGISNPTEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
451 KQGLEACLAGLVVSILFMIFFYKKFGLIATSALIANLVLIVGIMSLIPGA
501 TLTMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYKGAF
551 SSIFDANVTTLIKVIILYAVGTGAIKGFAITTGIGIATSMFTAIVGTRAI
601 VNLLYGGKRVKKLSI
```

FIGURE 12. The amino acid sequence of *Citrobacter freundii* SecD. The accession number is AF179925.

FIGURE 12. *Citrobacter freundii* SecD

```
  1 MLNRYPLWKYVMLIVVIVVGLLYALPNLYGEDPAVQITGVRGAAASEQTL
 51 IQVQKTLQEDKITPKSVALEEGAILARFDSTDTQLRAREVLMGVLGDQYV
101 VALNLAPATPRWLAAIHAEPMKLGLDLRGGVHFLMEVDMDTALSKLQEQN
151 IDSLRSDLREKGIPYTTVRKEDNYGLSITFRDSSARDEAISYLSKRHQDL
201 VISSQGSNALRAVMTDARLSEAREYAVQQNINILRNRVNQLGVAEPVVQR
251 QGADRIVVELPGIQDTARAKEILGATATLEFRLVNTNVDQSAAASGRVPG
301 DSEVKQSREGQPVVLYKRVILTGDHITDSTSSQDEYNQPQVNISLDSAGG
351 NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAVLVKQEEVINIANIQS
401 RLGNSFRITGINNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
451 KQGLEACLAGLVVSILFMIFFYKKFGLIATSALIANLVLIVGIMSLLPGA
501 TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRSVQQAINEGYAGAF
551 SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIVGTRAI
601 VNLLYGGKRVKKLSI
```

FIGURE 13. The amino acid sequence of *Klebsiella pneumoniae* SecD translated from a nucleotide sequence missing the termination codon.

FIGURE 13. *Klebsiella pneumoniae* SecD

```
  1 MLNRYPLWKYVMLVVVIVVGLIYALPNLYGEDPAVQITGARGVAASEQTL
 51 IQVQKTLQEEKITAKSVALEEGAILARFDTTDTQLRAREALLNVLGDKYV
101 VALNLAPATPRWLAALYAEPMKLGLDLRGGVHFLMEVDMDTALGKLQEQN
151 IDSLRSELRDKGIPYATVRKEDNYGLSIVFRDSAARDQAISYLSPRHRDL
201 VISSQGDNSLKAVMTDERLKEAREYAVQQNINILRNRVNQLGVAEPLVQR
251 QGADRIVVELPGIQDTARAKEILGATATLEFRLVNNNVDQSAAASGRVPG
301 DSEVKQTREGQPVVLYKRVILTGDHITDSTSSMDEYNQPQVNISLDSAGG
351 NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAILAKEEEVINIANIQS
401 RLGNSFRITGISNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
451 KQGLEACLAGLVVSILFMILFYKKFGLIATSALIANLILIVGIMSLIPGA
501 TLTMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYRGAF
551 SSIFDANVTTLIKVIILYAVGTGAIKGFAITTGIGIATSMFTAIVGTRAI
601 VNLLYGGKRVKKLSI
```

TABLE 4. The translated amino acid sequences of *Enterobacteriaceae* species SecD sequenced. In each comparison upper line contains the number of identical residues divided by the total giving the percent Identity. The lower line contains the number of similar residues divided by the total residues giving the percent similarity. The overall average of identical residues was 93% and the average of similar residues was 97%.

Table 4. Translated amino acid similarity of SecD in the Enterobacteriaceae sequenced

	E. coli	*S. choleraesuis*	*E. aerogenes*	*C. fruendii*	*K. pneumoniae*
E. coli	-----	587/615= 0.95 Id 606/615= 0.98 Sim	566/615= 0.92 Id 602/615= 0.97 Sim	589/615= 0.96 Id 605/615= 0.98 Sim	573/615= 0.93 Id 602/615= 0.97 Sim
S. choleraesuis		-----	571/615= 0.93 Id 601/615= 0.97 Sim	586/615= 0.95 Id 604/615= 0.98 Sim	572/615= 0.93 Id 600/615= 0.97 Sim
E. aerogenes			-----	569/615= 0.92 Id 600/615= 0.97 Sim	600/615= 0.97 Id 610/615= 0.99 Sim
C. fruendii				-----	571/615= 0.93 Id 598/615= 0.97 Sim
K. pneumoniae					-----

TABLE 5. The six potential transmembrane regions, molecular weight, and number of amino acids are shown for *S. choleraesuis*, *E. aerogenes*, *C. freundii*, and *K. pneumoniae* SecD. These are compared to the *E. coli*, *Mycobacterium tuberculosis*, and *M. leprae* SecD. Note that each contains six transmembrane regions and each contains some similarity to the others with the exception of *M. leprae*. The molecular weight shown for *E. aerogenes*, *S. choleraesuis*, and *C. freundii* was calculated with the compute Ip/Mw tool from ExPASy. No MW was calculated for *K. pneumoniae* due to the missing termination codon.

Table 5. Potential transmembrane regions of SecD

E. coli	*S. choleraesuis*	*E. aerogenes*	*C. fruendii*	*K. pneumoniae*	*Haemophilus Influenzae*	*M. tuberculosis*	*M. leprae*
10 to 30	10 to 30	10 to 30	10 to 30	10 to 30	11 to 31	13 to 33	225 to 245
452 to 472	455 to 472	452 to 472	454 to 472	457 to 474	453 to 473	385 to 405	296 to 316
477 to 497	477 to 495	476 to 497	477 to 495	479 to 500	475 to 495	410 to 430	321 to 341
504 to 524	506 to 524	506 to 524	506 to 524	505 to 522	497 to 517	441 to 461	352 to 372
564 to 584	556 to 574	550 to 573	556 to 574	556 to 574	562 to 582	489 to 509	400 to 420
586 to 605	579 to 597	579 to 597	579 to 597	579 to 597	585 to 605	514 to 534	425 to 445
615 AA	615 AA	615 AA	615 AA	615 AA	616 AA	573 AA	701 AA
66,632 MW	66,632 MW	66,744 MW	66,573 MW	---------	66,986 MW	60,267 MW	75,281 MW

VII. Amino Acid Comparison of *S. choleraesuis* SecD to *E. coli* SecD

The translated *S. choleraesuis* SecD amino acid sequence was compared to the translated *E. coli* SecD amino acid sequence to assess homology of the sequences. The comparison is shown in figure 14. This sequence comparison was done using the BLAST software "compare two sequences" in "blastp" mode from NIH on the internet (Altschul et al., 1997) and is shown in table 4. The percent identity of the sequence was approximately 95% and the similarity was approximately 98%.

VIII. Multiple Sequence Alignment of All Translated Amino Acid Sequences

The amino acid sequences of all *secD* genes were compared to the translated *E. coli* sequence, as shown in figure 15. The figure shows a very conserved consensus amino acid sequence for the SecD protein and where substitutions occur they are by similar amino acids, they occur at the same position in two or more species, or they are at a very specific amino acid position (see position 93, 184, or 209).

IX. Expression Vector Analysis

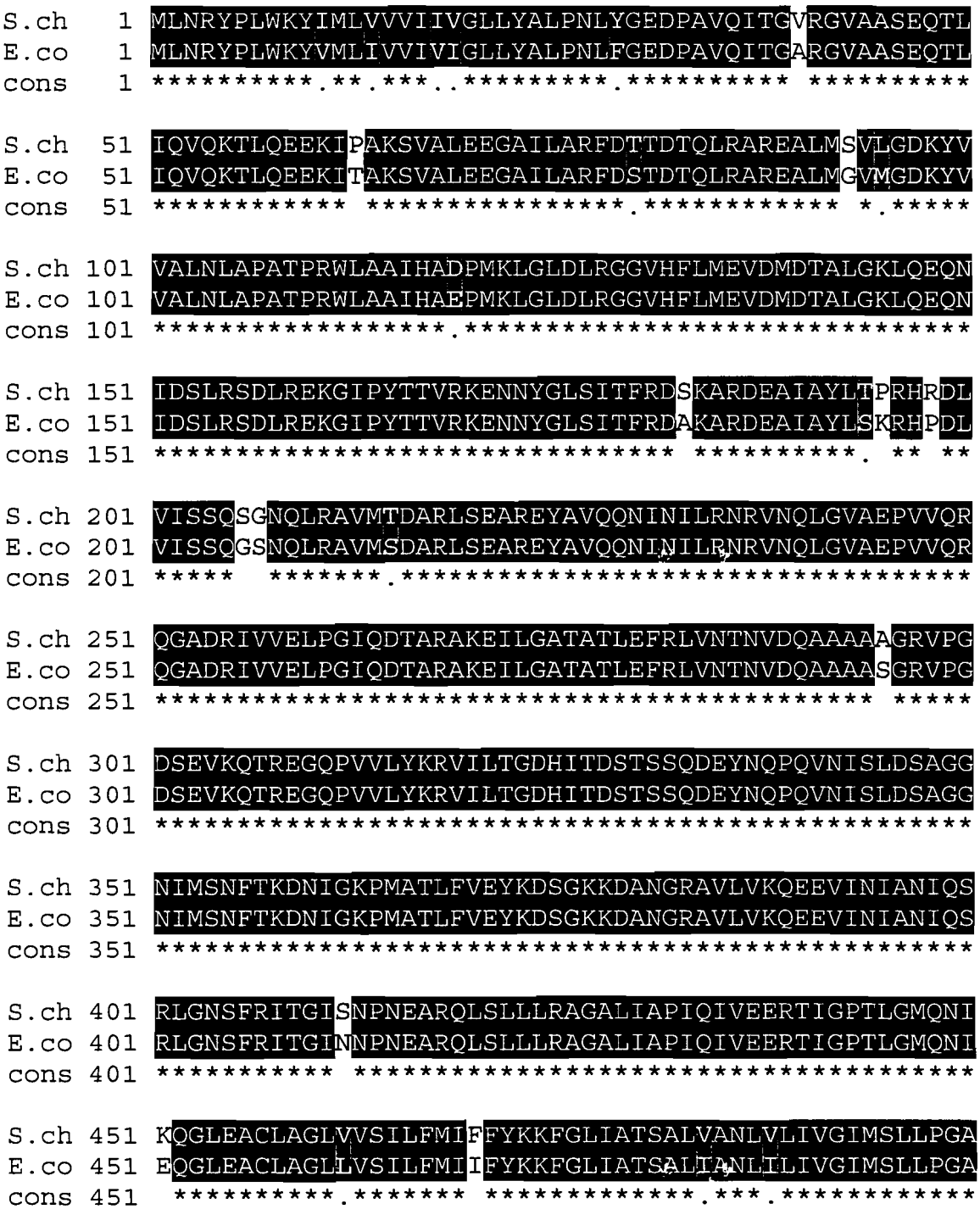
The agarose gel in figure 16 shows the expression vector restriction digestion. This confirmed proper insertion of the *S. choleraesuis secD* gene into the pQE30 expression vector.

X. SecD mutant complementation

The recombinant expression plasmid pQE30:12-5 containing the *Sac*I-*secD* *S. choleraesuis* gene was used to complement the *E. coli* K-12 SecD cold sensitive (Cs) mutant KJ173. The protein expressed by the recombinant plasmid complemented the cold sensitivity of the mutant in induced cells as evidenced by the growth at 23°C (table 6), indicating that the protein was analogous to SecD in *E. coli*.

FIGURE 14. Comparison of the translated amino acid sequence of *S. choleraesuis* SecD to *E. coli* SecD. Conserved amino acids are background shaded in black, with similar amino acid substitutions background shaded in gray. Dissimilar amino acid substitutions are denoted by a white background. Identity is denoted with a (*) on the consensus line, similar residues *i.e.* isoleucine substitution for valine is denoted with a (.).

Figure 14. Multiple sequence alignment



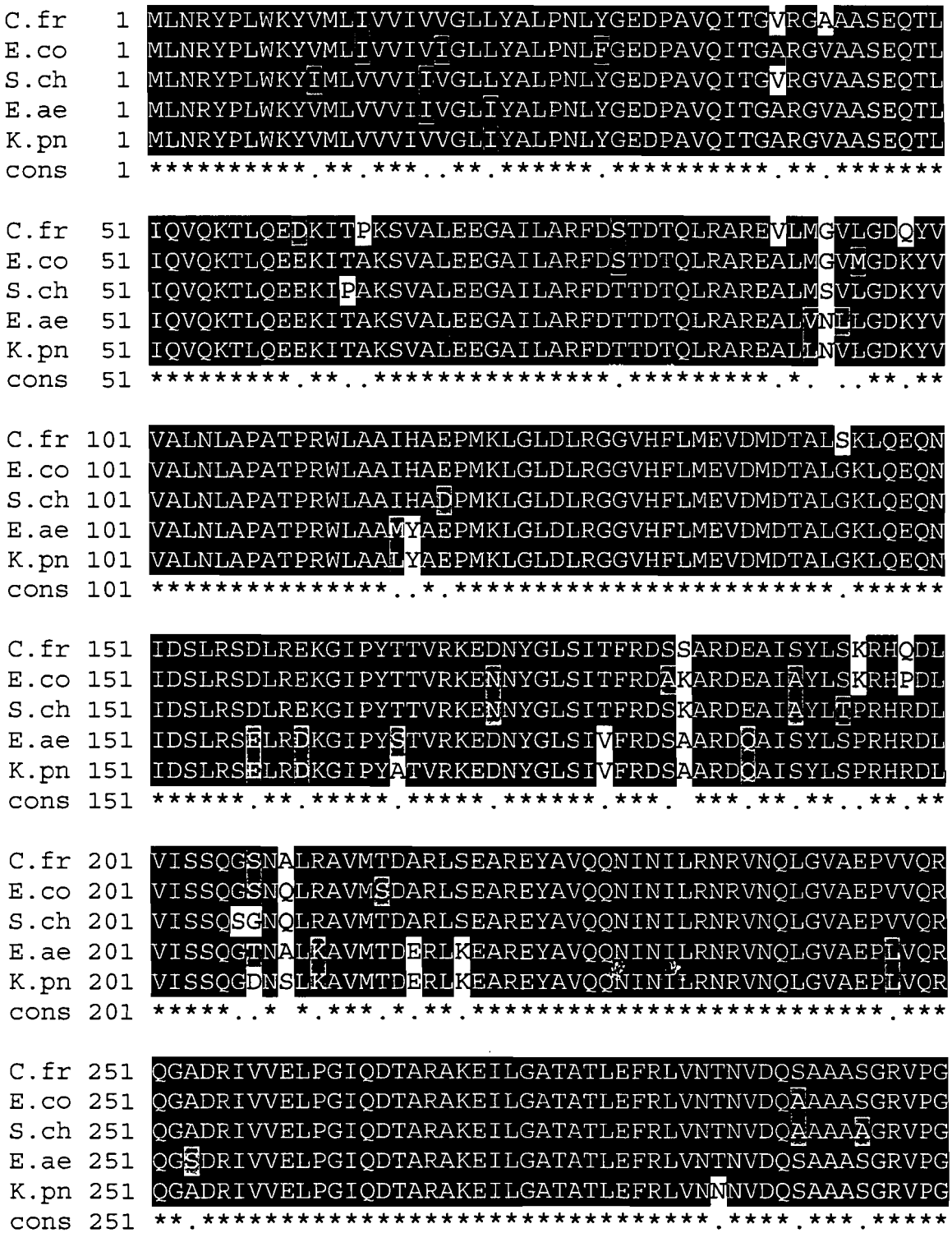
```
S.ch 501  TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAINEGYAGAF
E.co 501  TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYRGAF
cons 501  ****************************************** *** ***

S.ch 551  SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIIGTRAI
E.co 551  SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIVGTRAI
cons 551  *********************************************.*****

S.ch 601  VNLLYGGKRVKKLSI
E.co 601  VNLLYGGKRVKKLSI
cons 601  ***************
```

FIGURE 15. Multiple sequence alignment comparison of the translated SecD amino acid sequences of *S. choleraesuis, E. aerogenes, C. freundii, and K. pneumoniae* to *E. coli*. Conserved amino acids are background shaded in black, with similar amino acid substitutions background shaded in gray. Dissimilar amino acid substitutions are denoted by a white background. Identity is denoted with a (*) on the consensus line, similar resides *i.e.* isoleucine substitution for valine is denoted with (.).

FIGURE 15. Multiple Sequence Alignment



```
C.fr 301  DSEVKQSREGQPVVLYKRVILTGDHITDSTSSQDEYNQPQVNISLDSAGG
E.co 301  DSEVKQTREGQPVVLYKRVILTGDHITDSTSSQDEYNQPQVNISLDSAGG
S.ch 301  DSEVKQTREGQPVVLYKRVILTGDHITDSTSSQDEYNQPQVNISLDSAGG
E.ae 301  DSEVKDTREGRPVVLYKRVILTGDHITDSTSSMDEYNQPQVNISLDSAGG
K.pn 301  DSEVKQTREGQPVVLYKRVILTGDHITDSTSSMDEYNQPQVNISLDSAGG
cons 301  *****..***.**********************.*****************

C.fr 351  NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAVLVKQEEVINIANIQS
E.co 351  NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAVLVKQEEVINIANIQS
S.ch 351  NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAVLVKQEEVINIANIQS
E.ae 351  NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAILAKEEEVINIANIQS
K.pn 351  NIMSNFTKDNIGKPMATLFVEYKDSGKKDANGRAILAKEEEVINIANIQS
cons 351  ************************************.*.*.***********

C.fr 401  RLGNSFRITGINNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
E.co 401  RLGNSFRITGINNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
S.ch 401  RLGNSFRITGISNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
E.ae 401  RLGNSFRITGISNPTEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
K.pn 401  RLGNSFRITGISNPNEARQLSLLLRAGALIAPIQIVEERTIGPTLGMQNI
cons 401  ***********.**.************************************

C.fr 451  KQGLEACLAGLVVSILFMIFFYKKFGLIATSALIANLVLIVGIMSLLPGA
E.co 451  EQGLEACLAGLLVSILFMIIFYKKFGLIATSALIANLILIVGIMSLLPGA
S.ch 451  KQGLEACLAGLVVSILFMIFFYKKFGLIATSALVANLVLIVGIMSLLPGA
E.ae 451  KQGLEACLAGLVVSILFMIFFYKKFGLIATSALIANLVLIVGIMSLIPGA
K.pn 451  KQGLEACLAGLVVSILFMILFYKKFGLIATSALIANLILIVGIMSLIPGA
cons 451  .**********.*******.*************.***.********.***

C.fr 501  TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRSVQQAINEGYAGAF
E.co 501  TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYRGAF
S.ch 501  TLSMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAINEGYAGAF
E.ae 501  TLTMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYKGAF
K.pn 501  TLTMPGIAGIVLTLAVAVDANVLINERIKEELSNGRTVQQAIDEGYRGAF
cons 501  **.*********************************.*****.***.***

C.fr 551  SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIVGTRAI
E.co 551  SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIVGTRAI
S.ch 551  SSIFDANITTLIKVIILYAVGTGAIKGFAITTGIGVATSMFTAIIGTRAI
E.ae 551  SSIFDANVTTLIKVIILYAVGTGAIKGFAITTGIGIATSMFTAIVGTRAI
K.pn 551  SSIFDANVTTLIKVIILYAVGTGAIKGFAITTGIGIATSMFTAIVGTRAI
cons 551  *******.*****************************.********.*****

C.fr 601  VNLLYGGKRVKKLSI
E.co 601  VNLLYGGKRVKKLSI
S.ch 601  VNLLYGGKRVKKLSI
E.ae 601  VNLLYGGKRVKKLSI
K.pn 601  VNLLYGGKRVKKLSI
cons 601  ***************
```

FIGURE 16. Restriction digestion of recombinant expression pQE30:secD clones containing *Salmonella choleraesuis secD*. Lane 1 is colony 12-5 plasmid digested with *Kpn*I. Lane 2 is 12-5 plasmid digested with *BamH*I. Lane 3 is plasmid 3-3 digested with *Kpn*I. Lane 4 is 3-3 plasmid digested with *BamH*I. Lane 5 contains molecular wt. markers. Lane 6 is 3-4x plasmid digested with *Kpn*I. Lane 7 is 3-4x plasmid digested with *BamH*I. Plasmids in lane 4 show the proper restriction pattern.

Figure 16. Expression vector restriction digestion

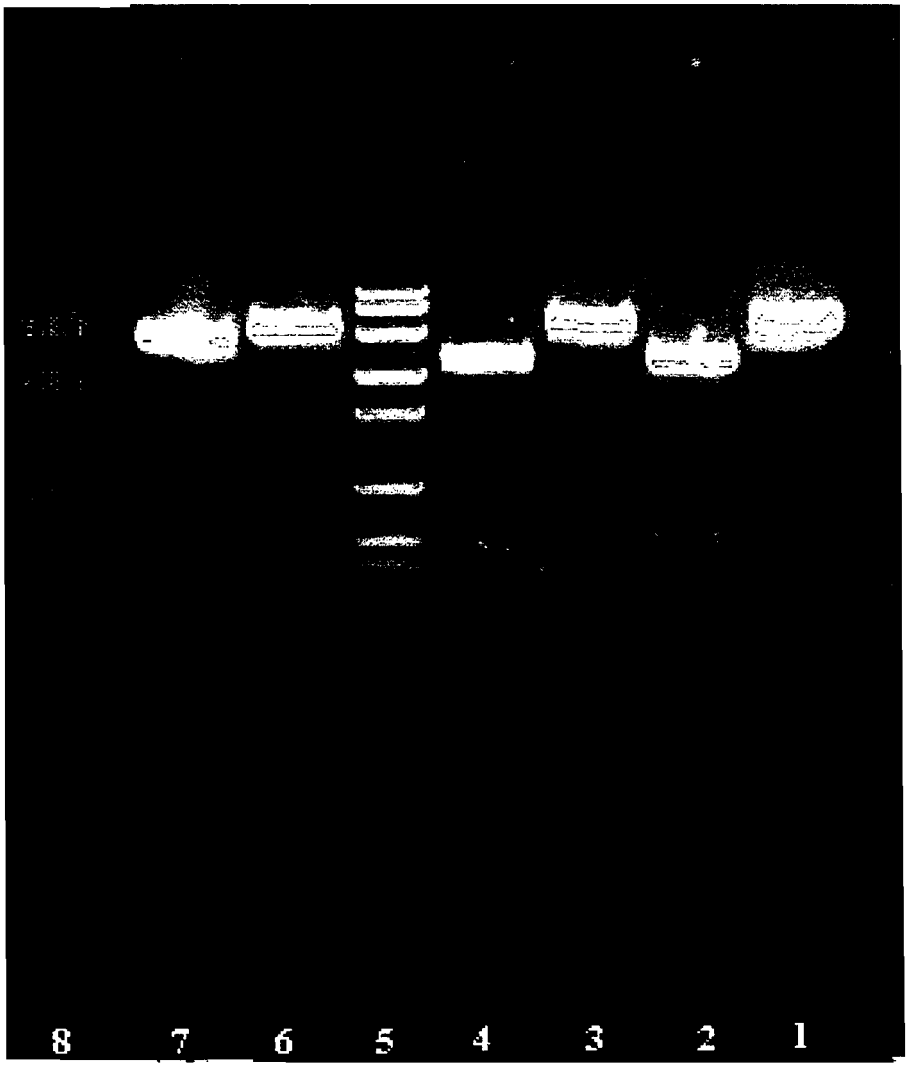


TABLE 6. Complementation of cold sensitive secD mutants. The 12-5 plasmids contain the SecD protein, the pQE16 plasmids contain an unrelated protein used as a protein control. Column one lists the recombinant plasmid tube tested. Column two shows the experimental reading for each tube minus the average absorbance of the negative control (pQE31) value of 0.0223. Column three is the result divided by the negative control average. Column four shows the increase in growth over the negative control, i.e. 12-5 #1 shows a 3.17 fold increase in growth over the negative control.

Table 6. Complementation of SecD Cs mutants

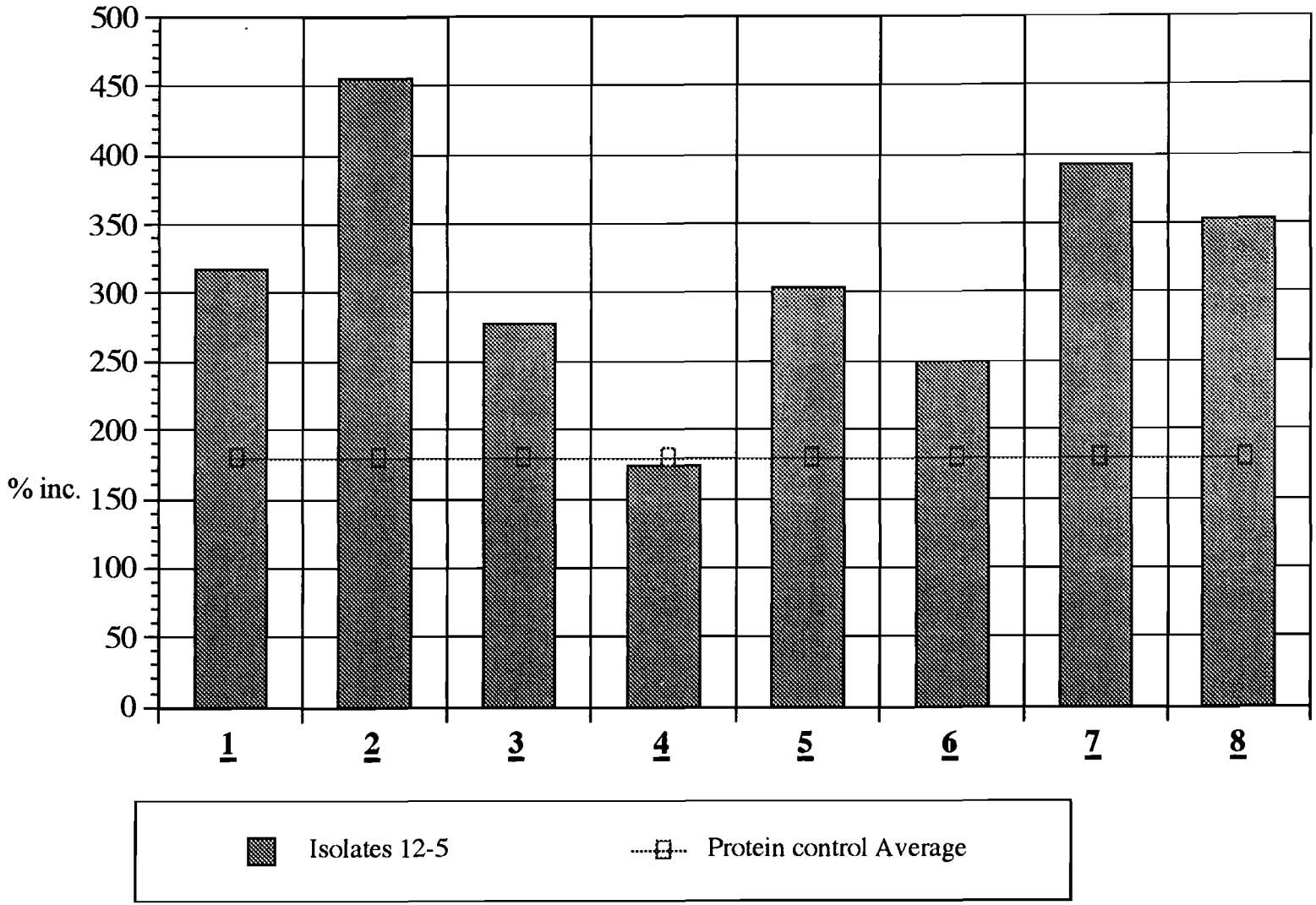
SecD Protein	experimental reading - neg. control average =	X / neg. cont. ave.	increase in growth over negative control
12-5 1	0.093 - 0.0223	0.0707/0.0223	3.17
12-5 2	0.124 - 0.0223	0.1017/0.0223	4.56
12-5 3	0.084 - 0.0223	0.0617/0.0223	2.77
12-5 4	0.061 - 0.0223	0.0387/0.0223	1.74
12-5 5	0.090 - 0.0223	0.0677/0.0223	3.04
12-5 6	0.078 - 0.0223	0.0557/0.0223	2.5
12-5 7	0.110 - 0.0223	0.0877/0.0223	3.93
12-5 8	0.101 - 0.0223	0.0787/0.0223	3.53
			3.16 ave. increase
Unrelated Protein			
pQE16 1	0.080 - 0.0223	0.0577/0.0223	2.59
pQE16 2	0.054 - 0.0223	0.0317/0.0223	1.42
pQE16 3	0.074 - 0.0223	0.0517/0.0223	2.32
pQE16 4	0.060 - 0.0223	0.0377/0.0223	1.69
pQE16 5	0.043 - 0.0223	0.0207/0.0223	0.93
			1.79 ave. increase

Complementation was demonstrated by the rapid growth of transformed *E. coli* Cs KJ173 mutants, compared to the pQE31 negative control and the pQE16, a non SecD protein producing control. The average increase in growth of the pQE30:12-5 over the pQE31 only negative control was 316%. The average increase in growth of the recombinant pQE16 protein producing control over the pQE31 negative control was 179% as shown in figure 17.

A one-way (single factor) analysis of variance (ANOVA) was performed on the data (Sigma Plot software) to test the affect of plasmid transformation on the mutants. The groups mean square (0.00669) divided by the error mean square (0.00279) gave an F value of 23.991 $P < 0.001$ indicating that not all of the population means were equal. A Tukey test was performed which showed the data consisted of three groups. The pQE30:12-5 complemented mutants differed from both the pQE31 and the pQE16 control groups. It also showed the pQE16 group which produced a protein unrelated to protein secretion differed from the pQE31 only group. These tests show mutants complemented with the pQE30:12-5 construct containing the *S. choleraesuis secD* gene had a statistically significantly greater amount of growth than the mutants transformed with the pQE31 or pQE16 plasmids.

FIGURE 17. Complementation of the cold sensitive secD mutants. The numbers along the x-axis refer to the recombinant clone tubes in table 5. The shaded bars represent the reading for each experimental tube. The horizontal line is the average reading for the protein control pQE16. The data was coded by subtracting the average absorbance of the negative control pQE31 value (0.0223) from each reading.

FIGURE 17. SecD mutant complementation



Discussion

DNA from the *Enterobacteriaceae,* a common family of human and veterinary opportunistic pathogens, was isolated and molecular biology techniques used to evaluate the presence of the *secD* gene. The *secD* gene was shown to be present in several members of the family. The genes from the various members were very highly conserved, with an average nucleotide identity of 87% (table 3) and an average amino acid homology identity of 93% with 97% similar residues (table 4).

Primers and PCR

Primers designed from the *E. coli secD* gene amplified the *secD* gene from several members of the *Enterobacteriaceae* family due to the homology of the N- and C- terminal sequences, which were identical in all of the species sequenced. This allowed N- and C- terminal sequencing of the *secD* gene, but the middle portion of the gene required nested primers for sequencing. Nested primers were designed from what was thought to be an apparently conserved region, chosen due to the identical base sequence in both the *E. coli* and *S. choleraesuis* sequences. The nested primers were used to sequence the *S. choleraesuis, E. aerogenes, K. pneumoniae,* and *C. freundii* internal sequences, even though some of the sequences contained up to three mis- matched bases per primer.

PCR was unable to amplify *secD* from *Proteus ssp.*. The annealing temperature was reduced and gradient PCR performed in an attempt to amplify the gene from *Serratia marcescens* and *Yersinia enterocolitica.* The *E. coli secD* primers amplified several bands, but a southern blot was needed before the proper band could be located and cloned, so these were not sequenced. PCR was performed (data not shown) on twenty four members of *Salmonella ssp.* to show *secD* is ubiquitous in the genus, and this trend is predicted to be true for the remaining genera of the family.

Gene Sequence Comparison

The *S. choleraesuis* and *E. coli secD* gene sequences were compared to determine the homology of the genes using BLAST from NCBI (Altschul et al., 1997). The sequence homology to *E. coli* was 87% for *S. choleraesuis secD*, which motivated the sequencing of other members. The average homology of the members sequenced was 87%. The greatest gene identity occurred between *E. aerogenes* and *K. pneumoniae* at 92% and the lowest occurred between *E. coli* and *K. pneumoniae* at 85% (table 3). This shows that the *secD* gene is very conserved in the *Enterobacteriaceae*.

Amino Acid Sequence

The translated amino acid sequences were compared and they shared an average of 93% identical and 97% similar residues (table 4). This shows that the primary amino acid sequence is more conserved than the nucleotide sequence. The similarity varied among the sequences with the greatest between *E. aerogenes* and *K. pneumoniae* and the least between *E. aerogenes* and *E. coli* or *C. freundii* (table 4). Figure 16 shows a very highly conserved SecD present in the family. This alignment of the translated amino acid sequences allows the preliminary prediction that the proteins may be analogous and that they may have a similar function in the membrane. Where amino acid substitutions occur, they tend to be position specific (clustered in a general area or more than 1 substitution at a specific site) and tend to be substituted by similar amino acids. There are five amino acid positions which show an unusual amount of substitution; four of which occur in the periplasmic domain at positions 93, 184, 207, and 209, and one occurs at position 547.

Proteins that have more than 20% identity in their primary sequence are considered homologous, meaning that the proteins share a common ancestor (www.cryst.bbk.ac.uk/PPS2/course/section2/implications.html). But in general,

most common sequence search engines use a cut off of about 30% identity in homology searches (Branden and Tooze, 1991).

The importance of showing sequence homology of the proteins is that it is a very good indicator of the possible structure and function of the sequences. Comparing the new sequences to a known sequence with a known function may allow prediction of its possible function and its possible structure in the membrane.

Conserved Transmembrane Domains

There are six potential transmembrane domains in *E. coli* SecD described by Beckwith et al. (1994). By examining the *S. choleraesuis* sequence with an Eisenburg hydropathy plot and several transmembrane (TM) region prediction programs (TMPREDD, TMHMM, and SOSUI), it is possible to predict transmembrane domains from the amino acid sequence (table 5). This analysis of the *S. choleraesuis* SecD also shows six apparent transmembrane regions, which appear to be conserved. The large periplasmic domain also appears to be conserved in *S. choleraesuis* SecD. These apparent secondary structures in *E. coli* SecD were also conserved throughout the genera studied. Positions of transmembrane segments and the position of the periplasmic domain remaining conserved throughout, as shown in the individual hydropathy plots (table 5 and figures 18 to 22).

Hydropathy plots are used to predict the position of transmembrane α-helices. A hydrophobicity scale, where each amino acid is given a value to represent its hydrophobicity, is used to calculate a hydropathic index for each position in the amino acid sequence (Branden and Tooze, 1991). The hydropathic index shows the mean value of hydrophobicity for approximately 21 amino acids about each amino acid position of the sequence. The hydropathy index increases in value with increasing hydrophobicity and this hydropathy plot

Figure 18. Hydropathy plot of the *E. coli* SecD protein. An Eisenburg hydropathy plot was performed on the amino acid sequence from GenBank. Accession number X56175.

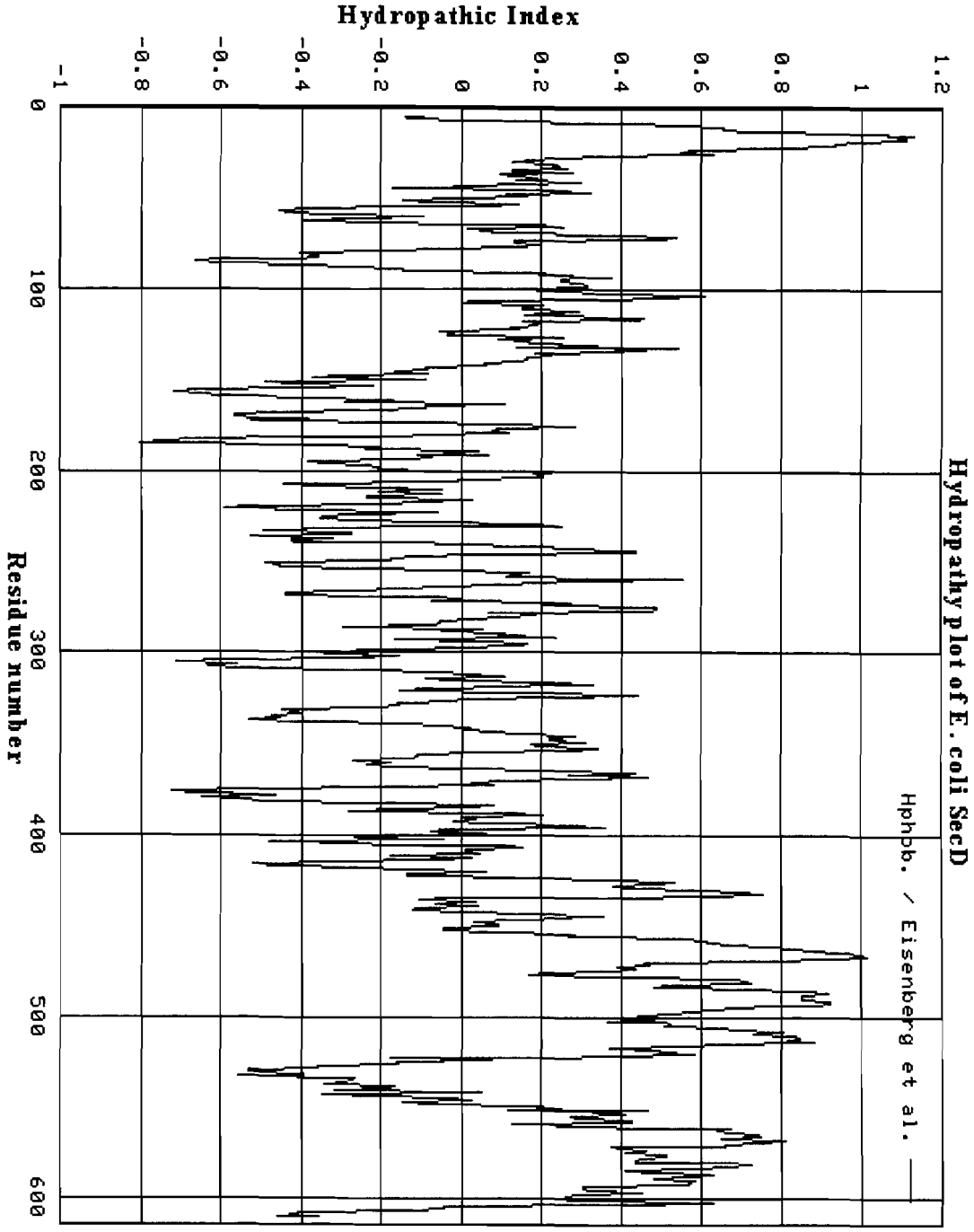

Hydropathy plot of E. coli SecD
Hydropathic Index
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
100
200
300
400
500
600
Residue number
Hphob. / Eisenberg et al.

Figure 19. Hydropathy plot of the *S. choleraesuis* SecD protein. An Eisenburg hydropathy plot was performed on the amino acid sequence from GenBank.

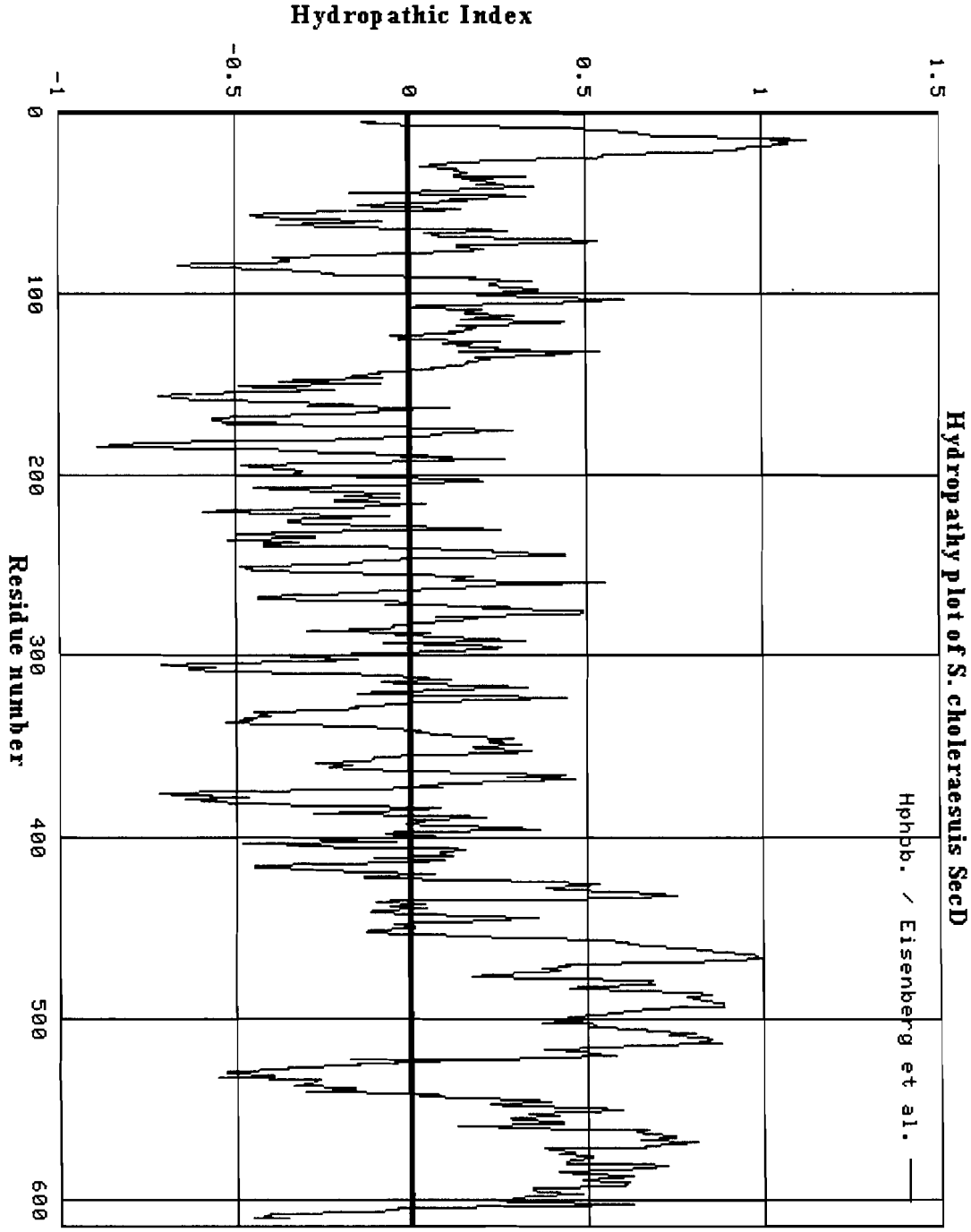

Hydropathic Index
-1
-0.5
0
0.5
1
1.5
0
100
200
300
400
500
600
Residue number
Hphob. / Eisenberg et al.
Hydropathy plot of S. choleraesuis SecD

Figure 20. Hydropathy plot of the *C. freundii* SecD protein. An Eisenburg hydropathy plot was performed on the amino acid sequence from GenBank.

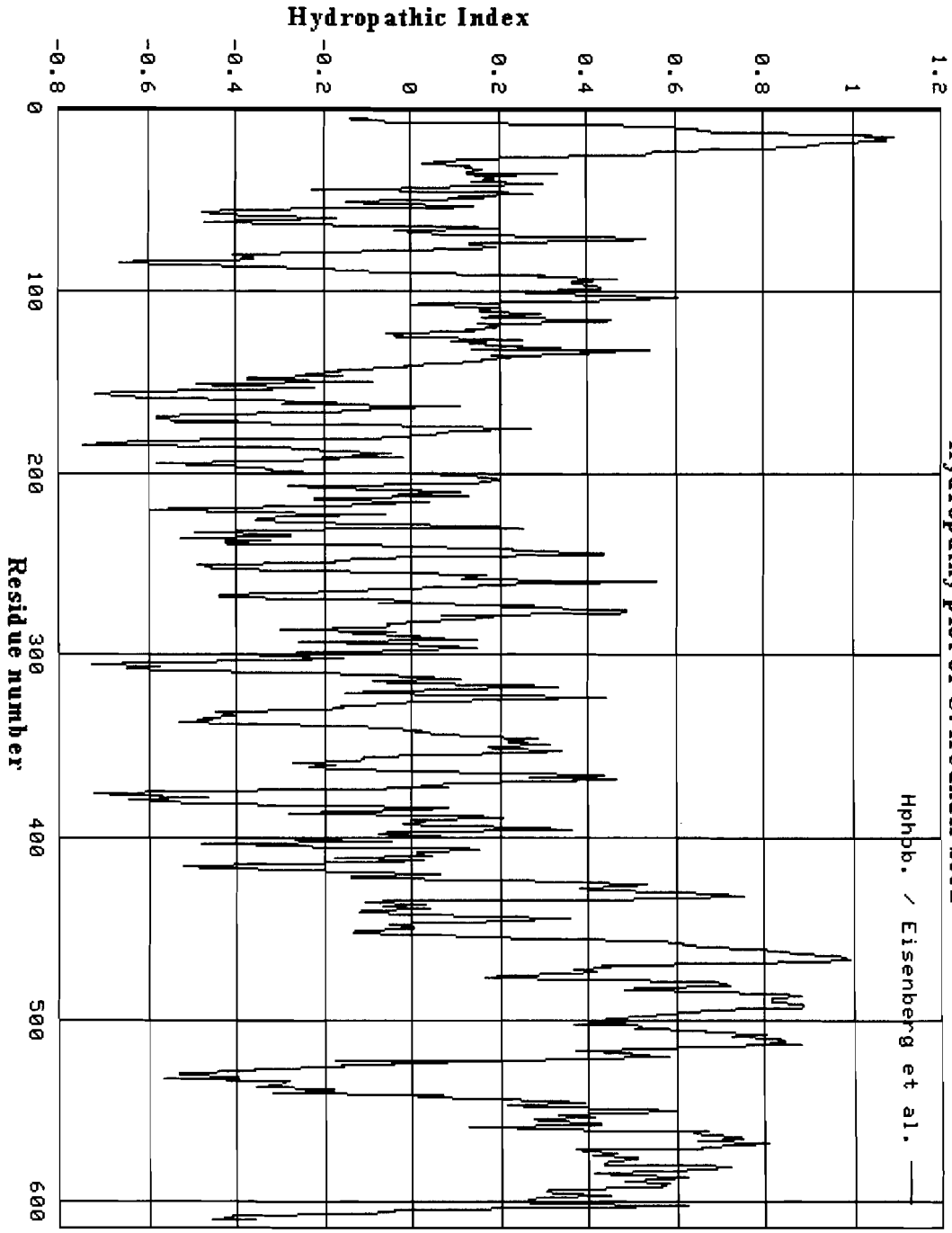

Hydropathy plot of C. freundii SecD
Hydropathic Index
Residue number
Hphbb. / Eisenberg et al.
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
100
200
300
400
500
600

Figure 21. Hydropathy plot of the *E. aerogenes* SecD protein. An Eisenburg hydropathy plot was performed on the amino acid sequence from GenBank.

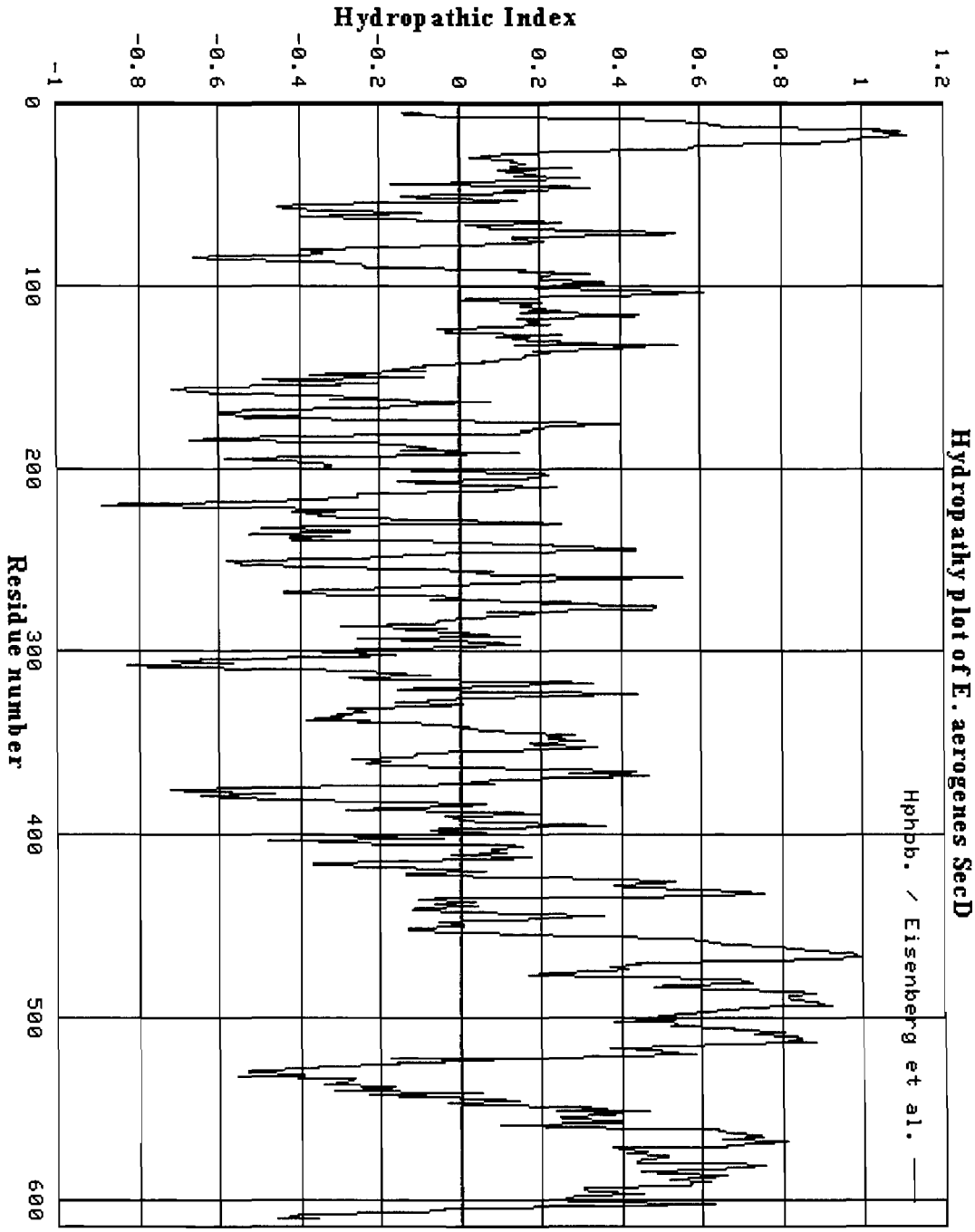
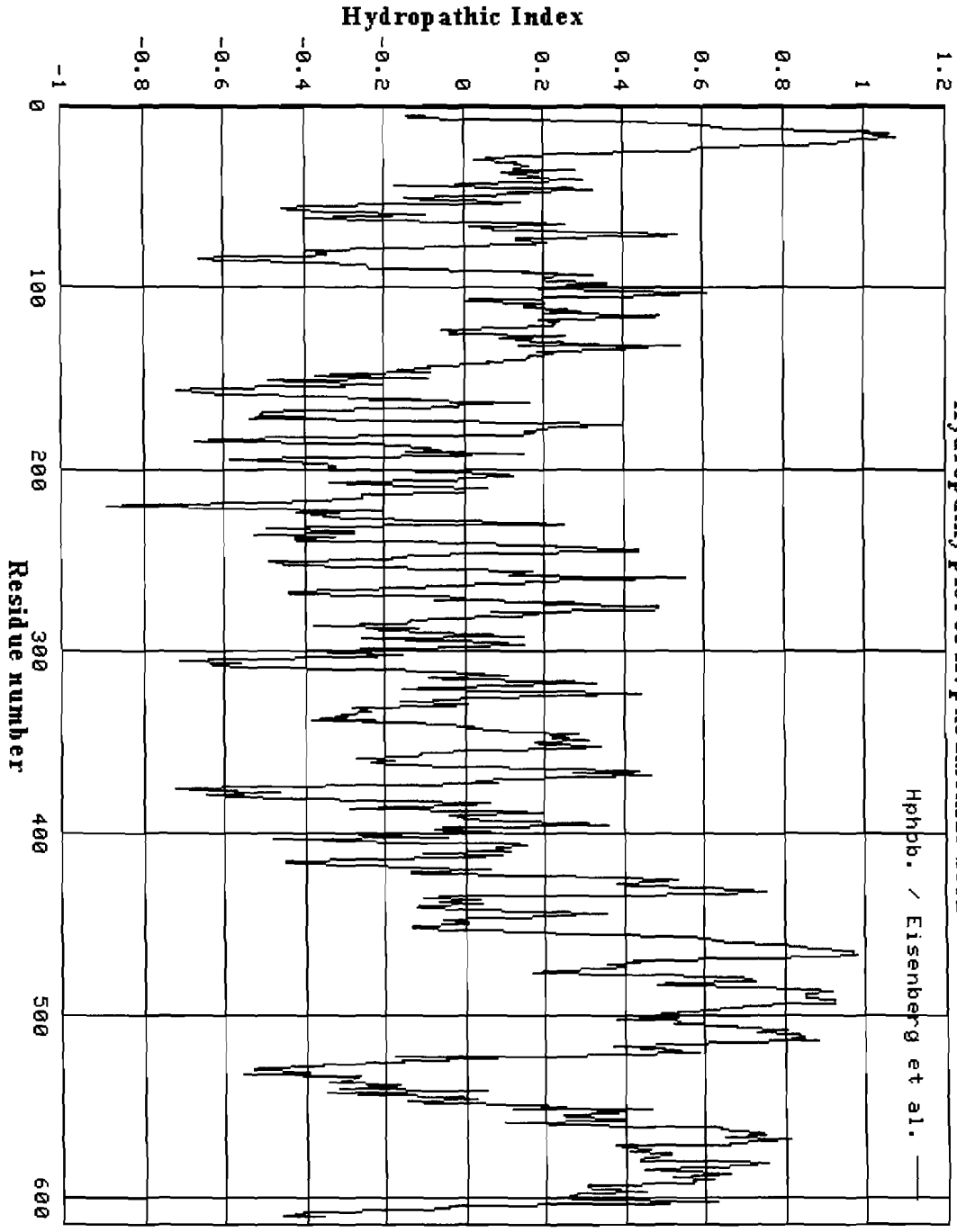

Hydropathy plot of E. aerogenes SecD
Hydropathic Index
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
Residue number
0
100
200
300
400
500
600
Hphob. / Eisenberg et al.

Figure 22. Hydropathy plot of the *K. pneumoniae* SecD protein. An Eisenburg hydropathy plot was performed on the translated amino acid sequence.


Hydropathy plot of K. pneumoniae SecD
Hydropathic Index
Residue number
Hphpb. / Eisenberg et al.
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
100
200
300
400
500
600

is used to identify the position of possible transmembrane helices by residue number (Branden and Tooze, 1991). Additionally, the number of putative transmembrane regions and size of the protein is similar to other gram negative and acid fast bacteria shown to contain SecD, namely *Haemophilus influenzae*, *Mycobacterium leprae*, and *M. tuberculosis* are shown (table 5).

Protein Production

The initial attempts to overproduce the *S. choleraesuis* SecD protein used the cloned gene which included the leader sequence. These attempts did not produce the desired results. So a truncated version of the gene was cloned into the expression vector. This truncated gene was missing the N-terminal portion of the gene, including the leader sequence. It was constructed by cleaving the *secD* gene at its *BamHI* site (position 93). The protein band produced by the recombinant vector is the same size, at approximately 54 kDa in 10% SDS-PAGE, as is the *E. coli* SecD protein. Attempts to confirm this as the SecD protein through 6-His tag isolation were inconclusive, so another strategy was used as a confirmatory methodology (below).

Mutant Complementation

The *S. choleraesuis* SecD amino acid sequence was compared to the *E. coli* SecD sequence in part to assess the possible success of complementing the *E. coli* cold sensitive (Cs) mutant with the *S. choleraesuis secD* gene. Mutant complementation was performed using the full *S. choleraesuis secD* gene in a pQE30 expression vector construct. Several clones were tested for the ability to complement the Cs secD mutants, data for the pQE30:12-5 construct is shown in table 6. Complementation of the *E. coli secD* Cs mutant KJ173 with *S. choleraesuis* SecD show that these two proteins are analogous. This is demonstrated by the clones ability, when transformed into KJ173, to overcome its deletion mutation. This complementation when taken *in toto* with the gene

and amino acid sequence similarity shows that the *S. choleraesuis secD* gene and the *E. coli secD* gene code for analogous proteins.

Summary

The question posed for this research was, does SecD occur in other members of the *Enterobacteriaceae* that is homologous and analogous to that of *E. coli*? To answer this question several pathogenic members of this family were obtained, chromosomal DNA was isolated from each, and PCR was performed using primers designed from the *E. coli secD* gene. PCR products of the appropriate size were agarose gel extracted, purified, and ligated into the plasmid cloning vector pT7Blue-3. Several colony purified recombinant clones from each species were restriction mapped to confirm proper insert. Clones containing the proper insert were sequenced. The *secD* gene sequence for each species was constructed using 3- 4 sequencing runs to determine each nucleotide position. Each sequence was compared to the *E. coli secD* gene and each other for DNA homology. Comparisons revealed that the sequences were homologous, with an average identity of 87%. Translated amino acid sequences compared using BLASTp and CLUSTALw show homology, with the average amino acid identity 93% and 97% similar residues. Analysis of the secondary structure shows strong transmembrane similarities (6 regions) among all SecD. The relative positions of the putative *E. coli* SecD transmembrane regions and large periplasmic domain were found to be conserved in each sequence.

PCR performed on twenty four members of *Salmonella* to show *secD* is ubiquitous in the genus and this is predicted to hold true for the remaining genera of the family. Complementation studies performed using a cold sensitive *secD* mutant strain were successful, complemented by the protein expressed from a recombinant expression vector containing the *Salmonella choleraesuis secD* gene. This construct was also used to produce the *S. choleraesuis* SecD protein for comparison to *E. coli* SecD on SDS-PAGE. Protein expression produced a

54kDa product that migrated to the same position as *E. coli* SecD. Conclusions drawn are the *E. coli* and *S. choleraesuis secD* genes are analogous and the other sequenced *secD* genes are very conserved in the family, with all showing homology to the *E. coli secD* gene.

References

Akiyama, Y. and Ito, K. (1987). Topology analysis of the SecY protein, an integral membrane protein involved in protein export in *Escherichia coli*. The EMBO Journal *6(11)*,3465-3470.

Altschul, S.F., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research *25*,3389-3402.

Arkowitz, R., and Bassilana, M. (1994). Protein translocation in *Escherichia coli*. Biochimica et Biophysica Acta *1197*,311-343.

Arkowitz, R., and Wickner, W. (1994). SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation. The EMBO Journal *13(4)*,954-963.

Brandez, C. and Tooze, J. (1991). Introduction to protein structure. (New York, New York: Garland Publishing, Inc.) pp. 1-302.

Brundage, L., Hendrick, J., Schiebel, E., Driessen, A., and Wickner, W. (1990). The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. Cell *62*,649-657.

Collier, D., Bankaitis, V., Weiss, J., and Bassford, P. , Jr. (1988). The antifolding activity of Sec B promotes the export of the E. coli maltose-binding protein. Cell *53*,273-283.

Dalbey, R., and Robinson, C. (1999). Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. TiBS *24(1)*, 1-13.

Danese, P., and Silhavy, T. (1998). Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. Annual Review of Genetics *32*, 59-94.

Duong, F., and Wickner, W. (1997A). The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. The EMBO Journal *16(16)*,4871-4879.

Duong, F., and Wickner, W. (1997B). Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. The EMBO Journal *16(10)*,2756-2768.

Economou, A., Pogliano, J., Beckwith, J., Oliver, D., and Wickner W. (1995). SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. Cell *83*,1171-1181.

Gardel, C., Benson, S., Hunt, J., Michaelis, S., and Beckwith, J. (1987). *secD*, a new gene involved in protein export in *Escherichia coli.* Journal of Bacteriology *169(3)*,1286-1290.

Gardel, C., Johnson, K., Jacq, A., and Beckwith, J. (1990). The *secD* locus of *E. coli* codes for two membrane proteins required for protein export. The EMBO Journal *9(10)*,3209-3216.

Hanada, M., Nishiyama, K-i, Mizushima, S., and Tokuda, H. (1994). Reconstitution of an efficient protein translocation machinery compromising SecA and the three membrane proteins, SecY, SecE, and SecG (p12). The Journal of Biological Chemistry *269(38)*,23625-23631.

Hanada, M., Nishiyama, K-i., and Tokuda, H. (1996). SecG plays a critical role in protein translocation in the absence of the proton motive force as well as at low temperature. FEBS Letters *381*,25-28.

Hartl, F.-U., Lecker, S., Schiebel, E., Hendrick, J., and Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. Cell *63*, 269-279.

Hayashi, S., and Wu, H. (1990). Lipoproteins in bacteria. Journal of Bioenergetics and Biomembranes *22(3)*, 451-460.

Ito, K. (1984). Identification of the *secY (prlA)* gene product involved in protein export in *Escherichia coli.* Molecular & General Genetics *197(2)*,204-208.

Ito, K. (1992). SecY and integral membrane components of the *Escherichia coli* protein translocation system. Molecular Microbiology *6(17)*,2423-2428.

Jungnickel, B., Rapoport, T., and Harmann, E. (1994). Protein translocation: common themes from bacteria to man. *FEBS Letters 346*: 73-77.

Kawasaki, S., Mizushima, S., and Tokuda, H. (1993). Membrane vesicles containing overproduced SecY and SecE exhibit high translocation ATPase activity and countermovement of protons in a SecA- and presecretory protein-dependent manner. The Journal of Biological Chemistry *268(11)*,8193-8198.

Kim, Y., Rajapandi, T., and Oliver, D. (1994). SecA protein is exposed to the periplasmic surface of the *E. coli* inner membrane in its active state. Cell *78*, 845-853.

Kosic, N., Sugai, M., Fan, C., and Wu, H. (1993). Processing of lipid-modified prolipoprotein requires energy and *sec* gene products in vivo. Journal of Bacteriology *175(19)*, 6113-6117.

Martoglio, B., and Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. Trends in Cell Biology *8(10)*, 1-12.

Matsuyama, S., Fujita, Y., Sagara, K., and Mizushima, S. (1992). Overproduction, purification and characterization of SecD and SecF, integral membrane components of the protein translocation machinery of *Escherichia coli*. Biochimica et Biophysica Acta *1122*,77-84.

Meyer, T., Menetret, J-F., Breitling, R., Miller, K., Akey, C., and Rapoport, T. (1998). The bacterial Sec Y/E translocation complex forms channel-like structures similar to those of the eukaryotic sec61p complex. Journal of Molecular Biology *285*,1789-1800.

Nishiyama, K., Hanada, M., and Tokuda, H. (1994). Disruption of the gene encoding p12 (Sec G) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature. The EMBO Journal *13(14)*,3272-3277.

Pogliano, J., and Beckwith, J. (1993). The cs sec mutants of *Escherichia coli* reflect the cold sensitivity of protein export itself. Genetics *133*,763-773.

Pogliano, J., and Beckwith, J. (1994A). SecD and SecF facilitate protein export in *Escherichia coli.* The EMBO Journal *13(3)*,554-561.

Pogliano, J., and Beckwith, J. (1994B). Genetic and molecular characterization of the *Escherichia coli* secD operon and its products. Journal of Bacteriology, *February 176(3)*, 804-814.

Randall, L., Hardy, S. (1995). High selectivity with low specificity: how SecB has solved the paradox of chaperone binding. Trends in Biochemical Sciences *20*, 65-69.

Randall, L., Hardy, S., Topping, T., Smith, V., Bruce, J., and Smith, R. (1998). The interaction between the chaperone SecB and its ligands: Evidence for multiple subsites for binding. Protein Science *7*, 2384-2390.

Riggs, P., Derman, A., and Beckwith, J. (1988). A mutation affecting the regulation of a *secA-lacZ* fusion defines a new *sec* gene. Genetics *118*,571-579.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual. (Cold Spring Harbor Laboratory Press, John Wiley and Sons, Inc. U.S.A.)

Schiebel, E., Driessen, A., Hartl, F-U., and Wickner, W. (1991). $\Delta\mu_H^+$ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell *64*, 927-939.

Settles, A., and Martienssen, R. (1998). Old and new pathways of protein export in choroplasts and bacteria. Trends in Cell Biology *8 (12)*, 1-15.

Stephens, C., and Shapiro, L. (1997). Bacterial protein secretion -A target for new antibiotics? Chemistry & Biology *4(9)*, 1-9.

Sugai, M. and Wu, H. (1992). Export of the outer membrane lipoprotein is defective in *secD, secE*, and *secF* mutants of *Escherichia coli.* Journal of Bacteriology *174(8)*,2511-2516.

Taura, T., Akiyama, Y., and Ito, K. (1994). Genetic analysis of SecY: additional export-defective mutations and factors affecting their phenotypes. Molecular and General Genetics *243*,261-269.

Taura, T., Baba, T., Akiyama, Y., and Ito, K. (1993). Determinants of the quantity of the stable SecY complex in the *Escherichia coli* cell. Journal of Bacteriology *175(24)*,7771-7775.

Tokuda, H. (1994). Biochemical characterization of the presecretory protein translocation machinery of *Escherichia coli*. FEBS Letters *346*, 65-68.

Tusnady, G. and Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. Journal of Molecular Biology *283*,489-506.

Wickner, W.,and Leonard, M. (1996). Minireview: *Escherichia coli* preprotein translocase. The Journal of Biological Chemistry *271(47)*, 1-9.

Yang, Y-B., Yu, N., and Tai, P. (1995). SecE-depleted membranes of *Escherichia coli* are active. Journal of Biological Chemistry *272(21)*, 1-13.

Signature of Graduate Student

Signature of Major Advisor

I, Frank C. Kuhn, hereby submit this thesis to Emporia State University as partial fulfillment of the requirements for an advanced degree. I agree that the Library of the University may make it available to use in accordance with its regulations governing materials of this type. I further agree that quoting, photocopying, or other reproduction of this document is allowed for private study, scholarship (including teaching) and research purposes of a nonprofit nature. No copying which involves potential financial gain will be allowed without written permission of the author.

Signature of Author

12-9-99

Date

Cloning and functional homology of SecD within the *Enterobacteriaceae*

Title of Thesis

Signature of Graduate Office Staff

December 10, 1999

Date Received