

BNL-66447

RECEIVED
JUN 09 1999
OSTI

DNA Uptake by Transformable Bacteria

Sanford A. Lacks

Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

Correspondence: E-mail, lacks@bnl.gov; Tel. (516) 344-3369; Fax (516) 344-3407.

Historical introduction

The first dim realization that hereditary material was transferred from male to female animals probably arose among early pastoralists some ten thousand years ago, if not before. However, it was not until 1944 that the chemical substance of heredity, DNA, was directly shown to effect genetic change by demonstration of its role in bacterial transformation (Avery, McCarty & McLeod, 1944). The work of Avery and colleagues indicated that cells of *Streptococcus pneumoniae* could take up naked DNA. In the next decade, conjugative transfer of genetic markers, requiring cellular contact, was observed in bacteria (Lederberg & Tatum, 1946), and the introduction of viral DNA was shown to accompany bacteriophage infection (Hershey & Chase, 1952).

By the time that the double-stranded, helical structure of DNA was proposed in 1953, a number of animal and plant viruses had been purified and analyzed and found to contain either DNA or RNA. It was realized that all viruses contain nucleic acid as their genetic material, and that the nucleic acid must be introduced into the cell to infect it. A natural ability to be transformed genetically by free DNA was discovered in additional bacterial species including *Haemophilus influenzae* and *Bacillus subtilis*. During the 1950's extrachromosomal DNA elements called plasmids were discovered in bacteria, and many plasmids were found to carry the ability to be transferred by conjugation. Thus, mechanisms for DNA transfer appeared to be widespread.

Beginning in the late 1950's, the use of isotopically labeled material showed that DNA was physically introduced into recipient cells during transformation (Fox, 1957; Lerman & Tolmach, 1957). It was shown that transforming DNA is converted

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

**Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.**

to single-stranded segments during entry (Lacks, 1962) and that these lengthy segments are incorporated into the chromosomal DNA of the recipient cell (Fox & Allen, 1964; Notani & Goodgal, 1966). This molecular processing of transforming DNA and the realization that the development of competence for DNA uptake was a transient property of the cell, itself under elaborate control (Tomasz & Hotchkiss, 1964), that required synthesis of a dozen proteins (Morrison & Baker, 1979) pointed to the complexity of the natural transformation systems. Meanwhile, conjugation was also shown to result in transfer of single-stranded DNA (Cohen *et al.*, 1968), and this process, too, required the function of at least a dozen genes (see review by Wilkins, 1995).

In the 1970's various methods for artificially introducing DNA into cells began to be elaborated. These artificial mechanisms will be discussed here only briefly. Also, the mechanisms of viral transfer and conjugation will not be presented in much detail. The emphasis of this review will be on DNA uptake by the natural mechanisms of bacterial transformation. Research up to the mid-1970's on this topic was previously reviewed (Lacks, 1977b). Several fine recent reviews cover DNA uptake more broadly or from different perspectives (Dreiseikelman, 1994; Dubnau, 1991, 1997; Sabelnikov, 1994; Lorenz & Wackernagel, 1994; Lunsford, 1998; Palmen, Driessens & Hellingwerf, 1994). In the case of transformation, discovery of the mechanisms of competence regulation during the past decade has facilitated the identification of genes involved in DNA uptake. Therefore, most of the components of the uptake systems are now known, although their arrangements and precise functions are not. No case of DNA uptake is yet understood in detail, but it should be instructive at this time to assess the state of our knowledge.

Overview of DNA uptake mechanisms

The various processes of DNA uptake by cells can be categorized as: viral DNA entry, conjugation, or transformation. Within each category, a variety of mechanisms have been found. However, considerable similarities occur among the different mechanisms of conjugation and, especially, transformation. All of these natural mechanisms of DNA transfer are quite elaborate and involve multiple protein components, as the case may be, of the virus, the donor cell, and the recipient cell. The mechanisms of viral infection and conjugation will be discussed mainly with respect to their relevance to transformation.

Infection by viruses

Introduction of DNA in viral infection always involves interaction of viral proteins with receptors on the cell surface. Some viruses have elaborate structures and appendages composed of multiple proteins. For example, *Escherichia coli* T4 phage contains a tail-like appendage implicated in DNA uptake. As illustrated in Figure 1A, tail fibers bind to surface lipopolysaccharide (Cerritelli *et al.*, 1996), after which the tail core penetrates the outer membrane, fusing it with the inner cell membrane of the Gram-negative bacterium, to allow passage of phage DNA into the cell (Tarahovsky *et al.*, 1991). The phage protein gp2 may serve as a "pilot" that attaches to an end of incoming DNA to protect it from cellular exonucleases (Lipinska *et al.*, 1989) and perhaps to assist its entry. Such pilot proteins have been proposed in various DNA uptake systems but with little direct evidence, possibly because a single protein molecule per viral genome is difficult to detect.

A simpler virus, such as the filamentous phage fd, contains only a few coat

proteins. This virus infects "male" cells of *E. coli*, that is cells which contain the sex-factor plasmid F and which, therefore, extrude fibrils, called pili, from their surface. A minor coat protein, gp3, located at one end of the filamentous virus, binds to a pilus (Gray, Brown & Marvin, 1981), and the virus is drawn to the cell surface as the pilus retracts. Oligomers of this same gp3 protein then form a transmembrane channel (Glaser-Wuttke, Keppner & Rasched, 1989) through which the single-stranded fd DNA enters the cell. In addition to the pilin proteins forming the pilus, entry may require other host proteins such as TolA, mutants of which cannot be infected by fd. (Sun & Webster, 1986). TolA is a membrane protein, and it may contribute to the entry channel (Levengood & Webster, 1989).

Viral transfer mechanisms can introduce either single- or double-stranded DNA. The other natural mechanisms of DNA transfer, however, introduce only single-stranded DNA. This may reflect their function, which is to transfer genetic information from one cell to another, unlike the virus which acts as an independent entity and takes over the cell machinery.

Conjugation — cell to cell transfer

The key feature of conjugative DNA transfer is its requirement for close contact between cells, so that the DNA passes directly from donor cell to recipient without external exposure. Close contact is achieved by the action of donor cell pili, which first bind to the recipient cell surface. Upon retraction of the pili, the two cells are juxtaposed, and a channel through which single-stranded DNA can pass is formed (reviewed in Wilkins, 1995).

Conjugative transfer systems are often present on plasmids in both Gram-positive and Gram-negative bacteria. They allow transmission of the plasmid to other

cells independent of gene transfer mechanisms in the host. Typically, up to two dozen *tra* genes in the plasmid may encode components of the transfer system (Wilkins, 1995). In addition to the pilins and proteins that process and export them, these genes encode proteins that make up the transmembrane channel, and proteins that act on the plasmid DNA prior to its transfer, and even, by accompanying the DNA, after its transfer to the recipient cell (Rees & Wilkins, 1991).

Once the cells are in contact, a bridgelike channel must be formed between them for passage of the DNA. Many of the *tra* genes presumably encode proteins for this structure. Recipient proteins, such as OmpA, may play roles in cell contact and channel formation (Manoil & Rosenbusch, 1982; Sugawara & Nikaido, 1992). Several *tra* gene products process the donor DNA. In the case of F-factor, the protein TraI, a DNA-dependent ATPase with helicase activity, makes a single-strand break at the *oriT* site on the plasmid, binds covalently to the newly formed 5'-end, and unwinds the strand (Matson, Nelson & Morton, 1993). It is not clear whether TraI leads the donor strand into the recipient cell, as depicted in Fig. 1B, or whether it binds to the intercellular pore so that donor DNA loops into the cell (Wilkins, 1995). A single-stranded DNA binding protein is encoded by the F-factor; it may bind to the complementary strand, which is used as a template to resynthesize in the donor cell the strand corresponding to the one transferred.

Related to conjugative transfer is the interkingdom transmission of bacterial plasmid DNA from *Agrobacterium tumefaciens* to plant cells to form crown gall tumors (for reviews, see Kado, 1991; Citovsky & Zambryski, 1993). In this case, the Ti plasmid, which is also capable of conjugative transfer between bacteria, carries a

set of approximately two dozen *vir* genes that when activated, remove a single-strand segment of plasmid DNA, called T-DNA and transfer it through a conjugative bridge structure through the membrane of a plant cell, where it finds its way into the nucleus and becomes incorporated into the plant genome. The *virD1* and *virD2* products form an endonuclease that cuts at the T-DNA borders to release the strand segment. The VirD2 protein is bound to the 5-end of T-DNA, probably covalently, where it may act as a pilot protein to shepherd the strand into the host cell nucleus, protecting it from exonucleolytic digestion, and/or conserving the original bond energy for ligation to plant DNA (Dürrenberger *et al.*, 1989). Some of the *vir* genes are homologous to *tra* genes, for example *virB2* and *traA*, the pilin-encoding gene of the F-factor (Shirasu & Kado, 1993). Another gene, *virB11*, shares sequence similarity with *comGA*, which is required for transformation (Albano, Breitling & Dubnau, 1989).

Genetic transformation by free DNA

Natural bacterial transformation

Genetic transformation by purified DNA has now been demonstrated in over 40 species of bacteria (Lorenz & Wackernagel, 1994), but it is by no means a universal capability. As in conjugation, many specific proteins are required for transformation. However, since donor DNA is not accompanied by proteins, probably even under natural conditions, the proteins required for DNA uptake are produced in the recipient cell, rather than in the donor cell, as is the case for conjugation. It is not known whether under natural circumstances donor DNA is extruded from or leaks out of the cell or is released on lysis of cells, but extracellular

DNA can be detected in cultures (reviewed by Lorenz & Wackernagel, 1994).

The process of natural transformation allows large amounts of DNA--as much as 10% of their DNA content--to be taken up by cells (for a review, see Lacks, 1988). Chromosomal DNA markers routinely transform more than 1% of the recipient population; homogeneous, plasmid DNA containing a chromosomal marker can transform more than half of the cells. In most species all of the cells become competent to take up DNA; in *E. subtilis*, however, only ~10% of the population becomes competent. The DNA that is taken up by the recipient cell can be integrated with very high efficiency, with estimates ranging between 0.5 and 1.0.

Artificial transformation--electroporation

In contrast to natural transformation, artificial methods introduce relatively small amounts of DNA into cells. Two types of methods have been used. A cell-DNA mixture is subjected to abrupt shifts in divalent cation concentration and temperature in one method (Mandel & Higa, 1970) and to high-voltage electrical pulses in the other (Harlander, 1987). The former was more important historically, but the latter method, called electroporation, has become the method of choice because of its simplicity and general applicability. Although the precise mechanisms by which these agents act are not known; it is presumed that they form transient pores in the cell membrane(s), through which DNA can pass (for reviews see Chang *et al.*, 1992; Sabelnikov, 1994).

Despite the small quantities of DNA introduced by artificial transformation, it has proven extremely useful in working with the many species that are not naturally transformable. The DNA is taken up in the form presented to the cells without degradative processing, which is advantageous for establishing a plasmid. This

compensates, in part, for the relatively poor uptake.

Molecular fate of DNA in transformation

Initial attachment of double-stranded DNA

In all known systems of natural transformation, DNA must be double-stranded, which is its native state, to be efficiently taken up. Although, denatured (Miao & Guild, 1970) or single-stranded (Barany & Boeke, 1983) DNA can enter cells, its ability to do so is less than 1% of double-stranded DNA. Kinetic studies, in which the concentration of donor DNA is varied, suggest the existence on the surface of Gram-positive bacteria of ~50sites that can reversibly bind dsDNA. Very quickly, however, DNA becomes irreversibly bound, so that it can no longer be washed away, although its sensitivity to external agents, like DNase, shows it to be still outside the cell. With *S. pneumoniae*, donor DNA at this stage can be examined in mutants lacking the EndA nuclease required for entry (Lacks, Greenberg & Neuberger, 1974) or in wild-type cells treated with chelating agents (Seto & Tomasz, 1974), which block the nuclease action. With *B. subtilis*, the entry process is slower, and the external state of DNA can be investigated without inhibiting its entry. These Gram-positive species and others are able to take up DNA indiscriminantly with regard to its nucleotide sequence.

At least some Gram-negative species, typified by *H. influenzae*, require for binding a 9- to 11-nt sequence found specifically in their own DNA (Danner *et al.*, 1980). Cells of *Neisseria gonorrhoeae* also require a specific sequence, but one that is different from *H. influenzae* (Goodman & Scocca, 1988). In addition to the sequence requirement, DNA bound by these species becomes resistant to

extracellular agents. In the case of *H. influenzae* it has been shown that such bound DNA is protected by cell membranes in a vesicle (Kahn, Maul & Goodgal, 1982) called a transformasome. Within the transformasome, donor DNA remains double-stranded, but as in the case of Gram-positive cells, it is converted to single strands on entry into the cell proper (Barany, Kahn & Smith, 1983). According to one model (Sabelnikov, 1994), DNA binds to receptors on the outer membrane that recognize the specific sequence. This triggers curvature of the membrane, which envelopes the DNA to form a vesicle within the outer membrane. The vesicle, or transformasome, then contacts the inner membrane, fusing with it at the location of a competence-induced structure for DNA entry.

DNA degradation and entry of single strands

External strand breaks

It is likely that DNA binds initially to a cellular receptor in a reversible manner. However, in the transformation of *S. pneumoniae* and *B. subtilis*, a stage of irreversible binding in which the DNA is still external can be identified. Such binding is an active process with its own ionic and energy requirements. In *S. pneumoniae* it has been possible to isolate this stage in mutants lacking the membrane nuclease, EndA (Lacks *et al.*, 1974) and to show that concomitant with this binding is the appearance of single-strand breaks (nicks) in the DNA. These breaks apparently occur randomly along the DNA, every 2.5 kb on average, to give single-strand segments of 5 kb (weight average; Lacks & Greenberg, 1976); they may result from enzymatic action at entry sites. Proteins with strand nicking activity have been isolated from the cell surface of competent streptococci (Fujii *et al.*, 1987;

Lunsford, Nguyen & London, 1996), but they have not been well characterized. That the breaks are required for binding is indicated by the occurrence of at least one break on every molecule bound, as demonstrated for a covalently closed circular DNA donor (Lacks, 1979b). The reduction by half of DNA strand length upon uptake by cells (Morrison & Guild, 1972) also indicates that a strand break is required for DNA uptake.

In leaky mutants of *endA* that produce some nuclease, the nicks become double-strand breaks (Lacks & Greenberg, 1976); therefore, it can be presumed that in wild-type cells, the single-strand breaks are quickly converted to double-strand breaks, as the opposite strand is degraded and the initially nicked strand begins to enter the cell. In *B. subtilis*, donor DNA can be detected on the surface of wild-type cells, irreversibly bound, and nicked to give single-strand fragments of ~5kb (weight average; Dubnau & Cirigliano, 1972). Although the molecular processing of DNA during binding and entry has been investigated more fully in *S. pneumoniae*, these stages of uptake are considered to be similar in *B. subtilis* (Dubnau, 1991).

Binding is not sequence-specific in these Gram-positive species, but it is specific for unencumbered double-stranded DNA: Neither single-stranded DNA, RNA-DNA hybrids, nor glycosylated phage T4 DNA show appreciable binding or ability to compete with native DNA (Lacks, 1977a). Since the single-strand breaks render the binding irreversible, it is possible that a covalent link is formed with a surface protein during breakage. This protein might serve as a pilot, leading the DNA strand into the cell, or it might have another function, perhaps as the putative helicase proposed to assist DNA entry in *B. subtilis* (Londono-Vallejo & Dubnau,

1993). The TraI protein of the F-factor is reported to nick the transferred DNA, bind to it, and act as a helicase (Matson *et al.*, 1993). In the conjugative plasmid RP4, its TraI protein also accompanies single-stranded DNA into the recipient cell (Rees & Wilkins, 1990).

Single strand entry

Only single strands of donor DNA enter into *S. pneumoniae* cells (Lacks, Greenberg & Carlson, 1967). Similarly, newly entered DNA in *B. subtilis* is detectable as single strands (Piechowska & Fox, 1971). An amount of DNA approximately equal to the amount taken up is released in an acid-soluble form outside the cells (Lacks & Greenberg, 1973; Morrison & Guild, 1973). The released product consists of oligonucleotides of size distribution identical to that produced by the action of the EndA membrane nuclease of *S. pneumoniae* (Lacks *et al.*, 1974). It appears that one strand of DNA enters the cell as the complementary strand is degraded. Selective end-labeling of donor DNA showed that the 3'-end of the strand must enter first, whereas the opposite strand is degraded from the 5'-end (Mejean & Claverys, 1993). A reasonable view of events during entry is that a DNA strand with its 3'-end formed by the nick made on binding, enters the cell at the same time that the EndA nuclease makes an endonucleolytic break on the opposite strand and continues to degrade that strand from 5' to 3' as its complement enters 3' to 5'. A rate of 100 nt per sec was calculated for DNA entry in *S. pneumoniae* (Mejean & Claverys, 1993).

Null mutants of EndA are reduced in transformability to <0.1% of the normal level (Puyet, Greenberg & Lacks, 1990), and they are correspondingly reduced in

DNA uptake (Lacks, Greenberg & Neuberger, 1975). This accords with their role in DNA entry. The residual transformation, which is low but real, and the low level of transformation demonstrable with single-stranded DNA (Barany & Boeke, 1983) may result from an alternative mechanism for entry in the absence of EndA activity. A possible candidate for catalyzing such entry is the putative helicase encoded by *comFA* in *B. subtilis* (Londono-Vallejo & Dubnau, 1993) and its homologue, *cflA*, in *S. pneumoniae* (Table 1). An interesting set of mutants found in *S. pneumoniae* do not take up DNA into the cell but still bind DNA and degrade the same amount as wild-type cells (Morrison *et al.*, 1983). They are presumably normal for EndA but lack another factor necessary for entry, which might be the helicase, a pilot protein, or a protein forming the entry pore.

In Gram-negative species, DNA entry into the cell proper from the transformasome may occur by a mechanism similar to the Gram-positive species. DNA apparently enters the cell as a single strand, with its 3'-end entering first (Barany *et al.* 1983).

Ionic and energy requirements

Irreversible, external binding of DNA to *S. pneumoniae* cells requires an energy source (Lacks *et al.*, 1974) and potassium ions (Lacks, 1979a). The energy is required, possibly, for covalent linking of the DNA to a surface protein at the nick site. No divalent cations are necessary for binding. Potassium ions are necessary, perhaps for orienting the binding/entry complex in the membrane by maintaining the electrostatic potential across it.

When cells were preloaded with bound DNA in the presence of a chelating

agent, uptake into the cells was observed upon its removal, and that uptake was partially inhibited by a sugar analogue (Seto & Tomasz, 1974). So it appears that energy is needed for entry, too, perhaps to power a helicase that unwinds the entering DNA or a translocase that draws it into the cell. Calcium and magnesium ions are both required for entry (Seto & Tomasz, 1976; Lacks, 1977a). The role of calcium ions is unknown, but magnesium ions are required for the membrane nuclease activity (Lacks *et al.*, 1974).

The contribution to DNA uptake of the protonmotive force at the membrane, both with respect to electrostatic potential and pH gradient across the membrane, has been examined in both *S. pneumoniae* and *B. subtilis*. The data were reviewed recently (Palmen *et al.*, 1994), but no firm conclusions were drawn with respect to a requirement or function in uptake. However, some data with *S. pneumoniae* suggest that uptake depends on the ATP level in the cells (Clavé & Trombe, 1989). Possible roles for ATP are to activate a protein that is linked covalently to donor DNA during binding or to power helicase or translocase activity during entry.

Consequences of the entry mechanism

Nonhomologous DNA

DNA that enters into *S. pneumoniae* is coated with a single-stranded DNA binding protein that is induced during competence (Morrison, 1978). Nonhomologous DNA that is taken up persists in the cell with a half-life of ~30min at 30 C (Lacks *et al.*, 1967). It is apparently degraded slowly by exonucleases, and its breakdown products are reused for DNA synthesis.

Homologous recombination

DNA that is homologous to the recipient genome has a much shorter half-life as single strands, ~6 min at 30 C. This corresponds to the time required for integration of donor DNA into the chromosome (Ghei & Lacks, 1967). The single-stranded form of the donor DNA facilitates its interaction with the duplex DNA of the chromosome, and recombination occurs with an efficiency of at least 50% (Fox, 1957; Lerman & Tolmach, 1957).

Plasmid establishment

Because strand breaks occur on DNA binding and one strand is degraded during entry, the most plasmid material that can be introduced in a single entry event is a linearized single strand. Inasmuch as such a strand cannot circularize to form a replicon, establishment of a plasmid requires the interaction of complementary strand segments introduced in two separate events (Saunders & Guild, 1981)). However, if a plasmid carries a DNA segment homologous to the chromosome, pairing can occur in this segment between plasmid and chromosome to enable a single plasmid strand to establish the plasmid by a process called chromosomal facilitation (Lopez *et al.*, 1982).

Restriction of donor DNA

Restriction endonucleases generally do not act on either single-stranded or hemimethylated DNA. Therefore, chromosomal transformation by unmodified DNA is not affected by restriction enzymes in the host (Trautner *et al.*, 1974; Lacks & Springhorn, 1984). However, plasmid transfer requires interaction of two donor strands and is susceptible to restriction. In the case of *S. pneumoniae*, its *Dpn*II restriction system circumvents the plasmid susceptibility by expressing an accessory methyltransferase, *Dpn*A, that can methylate single-stranded DNA, so that plasmid

DNA is already methylated at potential restriction sites by the time that it becomes a double-stranded replicon (Cerritelli, Springhorn & Lacks, 1989).

Regulation of competence for DNA uptake

Streptococcal species

Competence-stimulating polypeptides

Early investigators of streptococcal transformation realized that the bacteria were not always transformable. They became competent for transformation only when grown in particular media and then only at a certain stage in the growth cycle, namely late logarithmic growth. Under these conditions an extracellular factor that stimulated competence in noncompetent cultures was isolated from cultures of *S. gordonii* (Pakula & Walczak, 1963) and *S. pneumoniae* (Tomasz & Hotchkiss, 1964). The factor from *S. pneumoniae* appeared to be a small protein (Tomasz & Mosser, 1966). This competence stimulating polypeptide (CSP) was recently shown to be a 17-mer polypeptide, EMRLSKFFRDFILQRKK (Ha'verstein, Coomaraswamy & Morrison, 1995). Other strains of *S. pneumoniae* and related streptococcal species have CSPs varying slightly in size and composition (Morrison, 1997).

Quorum-sensing mechanism

It was realized that a factor accumulating extracellularly could act as a signal of adequate bacterial culture density to render it likely that DNA leaking out of one cell would reach another cell of the species (Tomasz, 1965). The molecular mechanism of such quorum sensing is now understood (reviewed in Morrison, 1997). The precursor of CSP is encoded by *comC*, a small gene in an operon containing also *comD* and *comE*. (An operon is a cluster of genes transcribed on the same

messenger RNA.) Another operon contains *comA* and *comB*, the membrane-located products of which act to process the *comC* protein and export CSP out of the cell. When CSP accumulates, it signals ComD, which is a membrane-located histidine protein kinase, to phosphorylate ComE. The phosphorylated ComE activates transcription, possibly after another stage in the cascade, of competence-specific proteins.

Induction of competence genes

At least a dozen proteins are preferentially synthesized during the development of competence. Several operons encoding these proteins were found to have the "combox" sequence, TACGAATA (Campbell, Choi & Masure, 1998; Pestova & Morrison, 1998), in place of the typical pneumococcal *SigA* promoter (TTGACAN₁₂TNTGNTATAAT; Sabelnikov, Greenberg & Lacks, 1995). Thus, induction of these late competence genes may depend on a competence-specific sigma factor ultimately elicited by CSP. In a *trt* mutant strain of *S. pneumoniae*, which is transformable in the presence of trypsin and does not require CSP for competence (Lacks & Greenberg, 1973), combox genes are constitutively expressed (Lacks & Greenberg, unpublished data), perhaps because the putative sigma factor is no longer regulated.

Bacillus species

Competence-stimulating polypeptides

Cells of *B. subtilis* also become competent late in the culture cycle. In this species, two different polypeptides are released by the cell (reviewed by Solomon & Grossman, 1996). The ComX pheromone is a polypeptide of 9 or 10 aa, apparently

containing a modified tryptophan residue, that is processed from a 55-aa precursor. Similarly to CSP of *S. pneumoniae*, it acts on a histidine protein kinase (ComP) in the cell membrane, which in turn phosphorylates a transcription factor, ComA. Another competence-stimulating factor, which appears to be a small peptide, may act by blocking a phosphatase that restores ComA to an unphosphorylated and inactive state.

Regulation of gene expression

Phosphorylated ComA increases transcription of *srfA* (*comS*), which leads eventually to the accumulation of ComK (van Sinderen *et al.*, 1995). ComK acts to induce the late competence genes, not as a sigma factor, but as a transcription activator. Operons that are induced have a consensus SigA promoter, but they also have a characteristic pattern of A- and T-rich sequences upstream of the promoter, which apparently allow binding and activation by ComK (Hamoen *et al.*, 1998).

Gram-negative species

Regulated competence in H. influenzae

Competence appears in cells of *H. influenzae* when a culture is nutritionally depleted but can still carry on protein synthesis (Goodgal & Herriott, 1961). It appears to be mediated by the accumulation of cyclic AMP in the cell, which triggers the expression of a gene called *sxy* (Dorocicz, Williams & Redfield, 1993) or *tfoX* (Zulty & Barcak, 1995), the product of which, in turn, induces late competence genes (Karudapuram & Barcak, 1997). Null mutations in *sxy/tfoX* prevent competence, but certain missense mutants are constitutively competent, but at a lower level than normally attained (Redfield, 1991; Karudapuram & Barcak, 1997). The *sxy/tfoX*

product acts as a transcription activator for late competence genes, which generally have a 26-nt dyad symmetry element upstream from a SigA-like promoter (Karudapuram & Barcak, 1997). In the latter respect the induction system is more similar to *B. subtilis* than to *S. pneumoniae*.

Perpetual competence

A number of bacterial species, including the Gram-negative species *Neisseria gonorrhoeae*, *Acinetobacter calcoaceticus*, *Pseudomonas stutzeri*, the cyanobacterium *Synechococcus*, and the Gram-positive *Deinococcus radiodurans*, are competent throughout the culture cycle (see review by Lorenz & Wackernagel, 1994). This constitutive ability to take up DNA is mimicked by the *trt* mutant of *S. pneumoniae* (Lacks & Greenberg, 1973) and by some *sxy* mutants of *H. influenzae* (Redfield, 1991), which no longer regulate the late functions of competence.

Competence-inducible genes

Genes relating to competence have generally been classified as having early or late functions. Genes required for the *development* of competence, such as those involved in quorum sensing and signalling are the early genes. In *S. pneumoniae*, this class includes *comA,B,C,D*, and *E*. There may be additional genes mediating between these early genes and late genes. Late genes appear to constitute a regulon, that is, a set of coordinately induced operons. The late gene functions are needed both for DNA uptake and for allowing the introduced donor strands to persist in the cell and recombine with host DNA. The three best studied inducible systems, those of *S. pneumoniae*, *B. subtilis*, and *H. influenzae*, differ markedly in their early genes and their regulatory systems. However, their late genes show many similarities. The

fundamental mechanisms of uptake, therefore, may be quite alike.

Genes unique to the uptake mechanism

Inasmuch as the earliest and most complete characterization of late genes has been in *B. subtilis* (reviewed in Dubnau, 1997), the nomenclature originally used for them (*i.e.* *comGA*, *comGB*, etc.) will be taken to describe the general type. Their homologues in other species will derive their name from the *B. subtilis* gene, for example, *comG*-like: *cglA*, *cglB*, etc. Some exceptions based on the historical record, such as *cilC* for a *comC* homologue in *S. pneumoniae* (Campbell *et al.*, 1998), will be made. Since this review is centered on *S. pneumoniae*, reference will frequently be made to *cgl*, *cel* and *cfl* genes of this species.

The cgl family and cilC

The *comG* operon of *B. subtilis* contains seven genes, which encode proteins related to type IV pilins of Gram-negative bacteria (Albano *et al.*, 1989). Similar operons are found in other transformable species, and defects in any of the genes prevent binding and entry of DNA. The *cgl* operon of *S. pneumoniae* (Pestova & Morrison, 1998; see TIGR database for full sequence) also contains seven genes, which are homologous to those of *B. subtilis* (partly shown in Fig. 2). The *comG* proteins have been characterized with respect to their primary sequences, processing, and membrane location (Chung, Breidt & Dubnau, 1998). Based on those findings, the *cglA* and *cglB* products are larger proteins, 313 and 347 residues in length, respectively, an energy-transducing protein and a membrane-spanning protein responsible for transporting through the cell membrane the products of the five downstream *cgl* genes, which are all smaller polypeptides with hydrophobic segments

at their N-termini. In *B. subtilis*, ComGC, ComGD and ComGE contain the residues KGFT preceding the hydrophobic segment; in *S. pneumoniae*, a similar sequence, KAFT, precedes the segment in CglC, CglD and CglF (Fig.2). These N-terminal sequences are characteristic of prepilin-like proteins, typified by the type IV pilus protein of *Pseudomonas aeruginosa* (Sastry *et al.*, 1986). The earliest evidence connecting pilins to transformation was from *N. gonorrhoeae*, where strains unable to make pili were defective in transformation (Biswas *et al.*, 1977).

The *B. subtilis* *comC* gene (Dubnau, 1991) and its *S. pneumoniae* homologue, *cilC* (Campbell *et al.*, 1998; see TIGR database for full sequence), are both stand-alone genes unlinked to the *comG* or *cgl* locus. They encode a peptidase-methyltransferase that processes prepilins by cleaving the polypeptide at the F residue and methylating the new N-terminus (Strom, Nunn & Lory, 1993). In *H. influenzae*, the *cgl* locus contains only four genes. They encode proteins HI0299, HI0298, HI0297, and HI0296 (Fleischmann *et al.* 1995), which are homologues of the *P. aeruginosa* prepilin, ComGA, ComGB, and ComC, respectively.

In the three species with different patterns of competence regulation, the *cgl* operon is under late gene control. In *S. pneumoniae* the operon is preceded by the CATGAATA combox (Pestova & Morrison, 1998); in *B. subtilis* by the ComK recognition site (Hamoen *et al.*, 1998); in *H. influenzae* by the dyad symmetry element (Karudapuram & Barcak, 1997). *S. gordonii* uses a similar regulatory mechanism to *S. pneumoniae*, and a CATGAATA combox precedes its *comY* operon, which contains genes closely homologous to at least four *cgl* genes (Fig. 2; Lunsford & Roble, 1997). Although *S. gordonii* is transformable, *Streptococcus pyogenes* has

never been shown to be transformable, yet it has a homologous *cgl* locus (Fig. 2), preceded by a combox (Ha'verstein & Morrison, 1999). *Lactococcus lactis* also has a *cgl* operon (Fig. 2; sequence from Breunder, A. & Hammer, K., GenBank Acc. No. Y15043, 1997), but it is apparently transcribed from a typical SigA promoter. These observations raise the possibility that both species are transformable. Perhaps *S. pyogenes*, which lacks genes homologous to *comA-E* in *S. pneumoniae*, has a different system for signalling the development of competence than other streptococci, and maybe in *L. lactis* the entire regulatory mechanism is different, as in *B. subtilis*. Alternatively, the *cgl* pathway may have functions other than enabling DNA uptake; for example, in bacterial adhesion to surfaces.

How *cgl* products act in DNA uptake is unknown, but several possibilities can be envisioned (Chung *et al.*, 1998). By analogy to pilins, the products may form an appendage outside the cell membrane. This structure may act as a scaffold on which other proteins that bind and process DNA for entry are arranged. This scaffold may penetrate the peptidooglycan layer of the cell wall to form a passage for external DNA. In this connection some remodeling of the cell wall may occur, as discussed below. It is possible that *cgl* proteins are directly involved in binding or processing DNA, but no evidence for such functions have been adduced. An intriguing hypothesis, however, is that the *cgl* system, in a manner akin to pilus extrusion and retraction, serves as a motor for bringing DNA into the cell (Dubnau, 1997).

cel (comE-like) genes

Two genes, *comEA* and *comEC*, of the three in the *comE* operon of *B. subtilis* are essential for DNA uptake (Hahn *et al.*, 1993). Homologues of these two genes,

celA and *celB*, compose a late competence operon in *S. pneumoniae* (Pestova & Morrison, 1998). In *H. influenzae*, corresponding homologues, HI1008 (Fleischmann *et al.*, 1995) and *rec2* (Clifton, McCarthy & Roe, 1994), are transcribed separately, each under control of the dyad symmetry element. Relationships between products of these genes are shown in Fig. 3.

ComEA and CelA are proteins of ~20kDa with a long stretch of hydrophobic residues at the N-terminus. The 12-kDa *H. influenzae* homologue has a similar hydrophobic stretch but otherwise matches only the C-terminal half of the two other proteins. A resemblance between the C-terminus of ComEA and that of *E. coli* UvrC was previously noted (Dubnau, 1997). Mutations in *comEA* prevent binding of DNA to the cell, and it was shown recently that the ComEA protein *in vitro* binds tightly to DNA in a nonspecific fashion, but does not nick DNA (Provvedi & Dubnau, 1999). This protein, therefore, is very likely responsible for the initial, reversible binding of DNA to the cell.

ComEC, CelB, and Rec2 are 80-kDa proteins with multiple hydrophobic stretches that could be transmembrane segments. ComEC is not required for DNA binding, but it is necessary for DNA entry (Inamine & Dubnau, 1995). These proteins may form channels in the membrane for passage of DNA (Dubnau, 1997).

cfl (comF-like) genes

The *comF* operon of *B. subtilis* consists of three genes, two of which are essential for transformation and have counterparts in the two-gene *cfl* operon of *S. pneumoniae* (sequence from TIGR). ComFA mutants are reduced in transformability ~1000-fold (Londono-Vallejo & Dubnau, 1994). CflA mutations are

similarly defective, and they fail to take up DNA (D. Morrison, personal communication). These 50-kDa proteins show sequence similarity to ATP-dependent DNA helicases (Londono-Vallejo & Dubnau, 1993). Although they lack hydrophobic regions, ComFA was associated with the cytoplasmic face of the cell membrane (Londono-Vallejo & Dubnau, 1994).

With respect to function, it was suggested that ComFA acts as a helicase to help separate the DNA strands and propel one strand into the cell. It may also have a nicking activity, inasmuch as TraI of the F-factor conjugative system acts as a helicase after nicking the donor DNA (Matson *et al.*, 1993). Comparison of protein sequences indicates some resemblance of ComFA and CflA to UvrB of *E. coli* (Fig. 4A). UvrB contains a latent ATPase that is activated when it interacts with UvrC, and the UvrB-UvrC complex is able to nick DNA (Orren & Sancar, 1989). Conceivably, CflA, either by itself or together with CelA (which resembles UvrC), has the dual functions of nicking and unwinding donor DNA. If it binds covalently to the nicked DNA, as TraI is thought to do, it might also assist its entry into the cell.

ComFC, CflB, and the *comI01A* (*=comF* [Tomb *et al.*, 1991]) gene product of *H. influenzae* (Larson & Goodgal, 1991) are homologous 25-kDa proteins. Mutations in *comFC* reduced transformation only ~10-fold, but the reduction was much greater in *comF*. Mutants in *comF* bound normal amounts of DNA, but whether these proteins function in entry or in a later stage of transformation is unknown. The proteins contain two putative zinc fingers at their N-termini (Fig. 4), which may enable them to bind other components of the entry complex.

Post-entry genes

Examination of their genomes indicates that both *S. pneumoniae* and *B. subtilis* harbor two *ssb* loci that encode proteins homologous to the single-stranded DNA binding protein of *E. coli*. One is expressed as a late competence gene, and the other is constitutively expressed and presumably functions in DNA replication. In *S. pneumoniae*, it is apparently the competence-induced protein that is found bound to newly entered donor strands (Morrison, 1978). Whether such binding assists the uptake process is unknown, inasmuch as no mutants in the competence induced *ssbB* gene (= *cilA* [Campbell *et al.*, 1998]) have been tested.

There is only a single *recA* gene in *S. pneumoniae*, but its expression is increased ~5-fold during competence (Martin *et al.*, 1995), as a result of the presence of a combox (Campbell *et al.*, 1998) preceding its upstream neighbor, *exp10* (Pearce *et al.*, 1995; = *cinA* [Martin *et al.*, 1995]). DNA appears to enter normally in *recA* mutants.

Genes of unknown or accessory function

Three late competence genes in *S. pneumoniae* of unknown significance for DNA uptake are *exp10/cinA*, *coiA*, and *cilB*. Mutations in the first two genes reduce transformation ~10-fold (Pearce *et al.*, 1995; Pestova & Morrison, 1998). Homologues of *exp10/cinA* are found in many bacteria, but aside from a hydrophobic N-terminal region indicating a membrane location, nothing is known about the function of its protein product. As deduced from the genomic sequence of *S. pneumoniae*, the *coiA* gene is the first in an operon containing four genes. The products of *coiA* and *coiB* both appear to be peptidases, *coiC* encodes a putative

methyltransferase, and *coiD* encodes a putative cell-wall serine proteinase. It may be speculated that these enzymes play a role in cell wall remodelling during the development of competence.

CilB is homologous to the *H. influenzae* gene, *dprA*, in which mutations reduce chromosomal transformation 10,000-fold, with no effect on plasmid transfer (Karudapuram, Zhao & Barcak, 1995). Mutations in *dprA* may affect only chromosomal DNA entry because plasmids enter *H. influenzae* cells by a different mechanism (Pifer, 1986). Otherwise, the gene may be required only for chromosomal integration, as appears to be the case for *recP* in *S. pneumoniae* (Morrison *et al.*, 1973).

The case of *dpnA* is unusual in that it is regulated as a late competence gene (Lacks & Greenberg, in preparation), but it is not normally required for either chromosomal or plasmid transformation. It encodes a DNA methyltransferase that protects unmodified incoming plasmid DNA by methylating it while it is in a single-stranded form, thereby allowing plasmid establishment by unmodified donor DNA in a cell containing the *DpnII* restriction system (Cerritelli, Springhorn & Lacks, 1989).

Role of a membrane nuclease in DNA uptake

The *endA* gene

The only gene known to function in DNA uptake that is not induced during the development of competence is *endA* of *S. pneumoniae*. This constitutive gene is expressed from several upstream promoters, all with a SigA consensus sequence (Puyet *et al.*, 1990). The strongest promoter precedes the first gene of the operon

in which *endA* is the third and final gene. EndA is a 30-kDa polypeptide with a hydrophobic segment near its N-terminus (Puyet *et al.*, 1990), and it is located in the cell membrane from which it can be solubilized by mild detergents (Lacks & Neuberger, 1975). *In vitro*, it acts as a general endonuclease, producing oligonucleotide fragments from both single- and double-stranded DNA and from RNA (Rosenthal & Lacks, 1977). When isolated from membranes, EndA is found in a larger structure of ~250 kDa (Rosenthal & Lacks, 1980).

Mutants of *endA* that retain at least 5% of wild-type nuclease activity are fully transformable, but those with less than 1% activity are defective. Null mutants bind DNA but do not transport it into the cell; they are reduced in transformability more than 1000-fold (Lacks *et al.*, 1975; Puyet *et al.*, 1990). It is interesting that neither *endA* nor *cflA* null mutants reduce transformability as much as defects in *cgl* genes or those required for development of competence. Perhaps both EndA, which hydrolyses the complementary strand, and CflA, which may unwind donor DNA and/or translocate a donor strand into the cell, are required for efficient DNA strand entry, but in the absence of one, the other activity still manages to introduce some DNA.

No-entry degraders

Normally, one donor strand is degraded as the other enters the cell. In certain mutants, however, the expected amount of donor DNA is degraded and found outside the cell, but no DNA enters the cell (Morrison *et al.*, 1983). The single strand produced by EndA action presumably remains outside the cell. The mutations that give rise to these "no-entry degraders" occur in both the *cel* and *cfl* loci (D.

Morrison, personal communication). If CelB composes the pore for DNA entry through the cell membrane and CflA is needed for DNA passage through the pore, then defects in either could block entry without blocking degradation.

B. subtilis nucleases

Cells of *B. subtilis* contain at least ten enzymes with DNase activity (Rosenthal & Lacks, 1977). One of them, a 16-kDa protein encoded by *nucA*, is membrane-located (Smith *et al.*, 1983). This gene is under competence control, but mutations of it do not prevent transformation (van Sinderen, Kiewiet & Venema, 1995). However, the plethora of DNases in *B. subtilis* might allow redundancy of the entry nuclease. A possible candidate for this redundant function is the 30-kDa product of the gene *yurl* of *B. subtilis* (Kunst *et al.*, 1997), which has a potential membrane-binding segment and is homologous (28% identity; see Fig. 5) to the periplasmic nuclease, EndA of *E. coli*, a protein with enzymatic activities very similar to EndA of *S. pneumoniae*. A form of the *yurl* gene, called *bsn*, was cloned from a strain of *B. subtilis*, and its product was characterized as an RNase with no DNase activity (Nakamura *et al.*, 1992). However, only a truncated, extracellular form of the enzyme was examined, and that form may have lost its DNase activity. In addition, the nucleotide sequence of *bsn* was only 73% identical to the gene in the transformable strain 168 (shown in Fig. 5).

Is *S. pyogenes* transformable?

In addition to the *cgl* operon, many homologues of *S. pneumoniae* late competence genes are under combox control in *S. pyogenes* (Håverstein & Morrison, 1999). This suggests that bacteria of this species may be transformable, albeit with

a system different from *S. pneumoniae* for eliciting competence development. It was therefore of interest to look for the presence of *endA* in the *S. pyogenes* genomic sequence (SGSP database). A three-gene operon with the first two genes very closely homologous to those of *S. pneumoniae* (77% and 50% identity, respectively) is present. However, the *endA* homologue, although very similar to that of *S. pneumoniae* (61% identity, see Fig. 5), is truncated and missing 94 residues from its N-terminus. Because the hydrophobic segment of the protein is absent, the protein could not be bound to the membrane so as to function in DNA uptake. But *S. pyogenes* contains yet another nuclease gene, which we designate *endB*, unlinked to the first, and also encoding a product similar to *S. pneumoniae* EndA (29% identity), albeit less so than the *S. pyogenes* *endA* product (Fig. 5). The EndB protein, however, does contain a potential membrane-binding segment. It could support transformation in *S. pyogenes*.

Cell wall lysis and rebuilding

The first gene in the operon encoding EndA in *S. pneumoniae* is a *murA*-like gene that encodes UDP-N-acetylglucosamine 1-carboxyvinyltransferase, an enzyme required for cell wall synthesis. It is one of two *murA*-like genes in *S. pneumoniae*, and it may participate in cell wall synthesis and reconstruction during competence development. *B. subtilis* similarly contains two such genes, *murA* and *murZ*. Only one such gene is present in *E. coli* (Blattner *et al.*, 1997). The lytic fragility of cells of *S. pneumoniae* during competence and the ability of such cells to spontaneously form protoplasts have been reported (Lacks & Neuberger, 1975). The latter propensity was shown not to depend on the *cwl* (= *lytA*) gene that encodes the major

pneumococcal autolysin. It could depend on a competence-induced gene, which may be one of those with unknown function. CoiA, a putative peptidase, and CoiD, a putative cell-wall serine proteinase, are possible candidates. The purpose of such remodeling, would be to install into the cell wall an appendage, perhaps formed of *cgl*-encoded proteins, that could allow attachment of DNA and its transport through the peptidoglycan layer of the cell wall and the cell membrane.

Structure and function of the uptake machinery

A speculative scenario of how the late-competence gene products and other cell envelope components might interact to enable DNA uptake is depicted in Fig. 6. As a first step, CglA and CglB may act together to export CglC, D, E, F, and G, with CilC processing the N-terminii of CglC, D, and F. The five downstream *cgl* products would form an appendage extending into the cell wall layer and anchored in the cell membrane. Prior action of *coi*-encoded peptidases and MurA might disrupt and rebuild the cell wall in the vicinity of the pilin-related appendage. The Cgl structure would act as a scaffold for other proteins involved in DNA uptake, such as CelA, CelB, CflA, CflB and EndA. The scaffold would essentially trap CflA and CflB, which are not directly bound to the membrane. This structure would allow access of donor DNA to CelA, which could bind it reversibly. Combined action of CelA and CflA might then nick the DNA in an ATP-driven reaction, which would leave the DNA irreversibly bound to either CflA or CflB and, hence, the cell. Under normal conditions, helicase activity of CflA would separate the donor strands, and EndA would degrade one strand, both of which actions facilitate entry of the other strand into the cell through a membrane pore formed by CelB. Presumably, CflA or CflB, attached to the incoming strand at its 3'-end, acts as a pilot protein to bring

it through the transmembrane pore. One or more of these proteins may be able to pass through the transmembrane pore in either direction.

Once inside the cell, the entering strand is covered by SsbB protein. This may or may not assist its entry into the cell. Analysis of mutants defective in SsbB should answer this question. If EndA is defective, only a small amount of DNA is internalized by action of CflA. Similarly, if CflA is defective, only a small amount is internalized by action of EndA. When EndA is functional and CelB or CflA is defective, degradation of a donor strand could proceed, but the complementary strand would be unable to enter the cell, as found for mutants exhibiting degradation but no entry of DNA. Null mutants of CelA would not bind DNA at all.

ACKNOWLEDGEMENTS

I thank Donald Morrison and David Dubnau for generously providing information prior to publication. Genomic sequence data for *S. pneumoniae* was obtained from The Institute for Genomic Research (TIGR) website at <http://www.tigr.org>. Sequence data for *S. pyogenes* was obtained from the Streptococcal Genome Sequencing Project (SGSP) conducted at the University of Oklahoma by B. A. Roe, S. P. Linn, L. Song, X. Yuan, S. Clifton, M. McShan and J. Ferretti, funded by USPHS/NIH grant AI38406, and accessed at <http://dna1.chem.ou.edu/strep.html>. GenBank sequence data was obtained from the National Center for Biotechnology Information. This chapter was written at Brookhaven National Laboratory under the auspices of the U. S. Department of Energy Office of Biological and Environmental Research.

REFERENCES

Albano, M., Breitling, R. & Dubnau, D. A. (1989). Nucleotide sequence and genetic organization of the *Bacillus subtilis* *comG* operon. *Journal of Bacteriology*, 171, 5386-5404.

Avery, O. T., MacLeod, C. M. & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of Experimental Medicine*, 89, 137-158.

Barany, F. & Boeke, J. D. (1983). Genetic transformation of *Streptococcus pneumoniae* by DNA cloned into the single-stranded bacteriophage f1. *Journal of Bacteriology*, 153:200-210.

Barany, F., Kahn, M. E. & Smith, H. O. (1983). Directional transport and integration of donor DNA in *Haemophilus influenzae* transformation. *Proceedings of the National Academy of Sciences, USA*, 80, 7274-7278.

Biswas, G. D., Sox, T., Blackman, E. & Sparling, P. F. (1977). factors affecting genetic transformation of *Neisseria gonorrhoeae*. *Journal of Bacteriology*, 129, 983-992.

Blattner, F. R., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, 277, 1453-1474.

Campbell, E. A., Choi, S. Y. & Masure, H. R. (1998). A competence regulon in *Streptococcus pneumoniae* revealed by genome analysis. *Molecular Microbiology*, 27, 929-939.

Cerritelli, M. E., Wall, J. S., Simon, M. N., Conway, J. F. & Steven, A. C. (1996). Stoichiometry and domainal organization of the long tail-fiber of bacteriophage T4:

a hinged viral adhesin. *Journal of Molecular Biology*, 260, 767-780.

Cerritelli, S., Springhorn, S. S. & Lacks, S. A. (1989). DpnA, a methylase for single-strand DNA in the *DpnII* restriction system, and its biological function. *Proceedings of the National Academy of Sciences, USA*, 86, 9223-9227.

Chang, D. C., Chassy, B. M., Saunders, J. A. & Sowers, A. E., eds. (1992). *Guide to Electroporation and Electrofusion*. Academic Press, New York.

Chung, Y. S., Breidt, F. & Dubnau, D. (1998). Cell surface localization and processing of the ComG proteins required for DNA binding during transformation of *Bacillus subtilis*. *Molecular Microbiology*, 29, 905-913.

Citovsky, V & Zambryski, P. (1993). Transport of nucleic acids through membrane channels: snaking through small holes. *Annual Review of Microbiology*, 47, 167-197.

Clavé, C. & Trombe, M.-C. (1989). Intracellular pH, calcium, and ATP involvement in DNA uptake by *Streptococcus pneumoniae* competent for genetic transformation. *FEMS Microbiological Letters*, 65, 113-118.

Clifton, S. W., McCarthy, D. & Roe, B. A. (1994). Sequence of the *rec-2* locus of *Haemophilus influenzae*: homologies to *comE*-ORF3 of *Bacillus subtilis* and *msbA* of *Escherichia coli*. *Gene*, 146, 95-100.

Cohen, A., Fisher, W. D., Curtiss, III, R., & Adler, H. I. (1968). DNA isolated from *Escherichia coli* minicells mated with F+ cells. *Proceedings of the National Academy of Sciences, USA*, 61, 61-68.

Danner, D. B., Deich, R. A., Sisco, K. L. & Smith, H. O. (1980). An 11-base pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. *Gene*, 11, 311-318.

Dorocicz, I. R., Williams, P. M. & Redfield, R. J. (1993). The *Haemophilus influenzae* adenylate cyclase gene: cloning, sequence and essential role in competence. *Journal of Bacteriology*, 175, 7142-7149.

Dreiseikelmann, B. (1994). Translocation of DNA across bacterial membranes. *Microbiological Reviews*, 58, 293-316.

Dubnau, D. (1991). Genetic competence in *Bacillus subtilis*. *Microbiological Reviews*, 55, 395-424.

Dubnau, D. (1997). Binding and transport of transforming DNA by *Bacillus subtilis*: the role of type-IV pilin-like proteins - a review. *Gene* 192, 191-198.

Dubnau, D. & Cirigliano, C. (1972). Fate of transforming DNA following uptake by competent *Bacillus subtilis*. III. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. *Journal of Molecular Biology*, 64, 9-29.

Dürrenberger, F., Crameri, A., Hohn, B. & Koukolikova-Nicola, Z. (1989). Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proceedings of the National Academy of Sciences, USA*, 86, 9154-9158.

Fleischmann, R. D., M. D. Adams, O. White *et al.* 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269, 449-604.

Fox, M. S. (1957). Deoxyribonucleic acid incorporation by transformed bacteria. *Biochimica et Biophysica Acta*, 26, 83-85.

Fox, M. S. & Allen, M. K. (1964). On the mechanism of deoxyribonucleate integration in pneumococcal transformation. *Proceedings of the National Academy of Sciences*,

USA, 52, 412-419.

Fujii, T., Naka, D., Toyoda, N. & Seto, H. (1987). LiCl treatment releases a nickase implicated in genetic transformation of *Streptococcus pneumoniae*. *Journal of Bacteriology*, 169, 4911-4906.

Ghei, O. K. & Lacks, S. A. (1967). Recovery of donor deoxyribonucleic acid marker activity from eclipse in pneumococcal transformation. *Journal of Bacteriology*, 93, 816-829.

Glaser-Wuttke, G., Keppner, J. & Rasched, I. (1989). Pore-forming properties of the adsorption protein of filamentous phage fd. *Biochimica et Biophysica Acta*, 985, 239-247.

Goodgal, S. H. & Herriott, R. M. (1961). Studies of transformation of *Haemophilus influenzae*. 1. Competence. *Journal of General Physiology*, 44, 1201-1227.

Goodman, S. D. & Scocca, J. J. (1988). Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences, USA*, 85, 6982-6986.

Gray, C. W., Brown, R. S. & Marvin, D. A. (1981). Adsorption complex of filamentous fd virus. *Journal of Molecular Biology*, 146, 621-627.

Hahn, J., Inamine, G., Kozlov, Y. & Dubnau, D. (1993). Characterization of *comE* a late competence operon of *Bacillus subtilis* required for the binding and uptake of transforming DNA. *Molecular Microbiology*, 10, 99-111.

Hamoen, L. W., Van Werkhoven, A. F., Bijlsma, J. J. E., Dubnau, D. & Venema, G. (1998). The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes &*

Development, 12, 1539-1550.

Harlander, S. K. (1987). Transformation of *Streptococcus lactis* by electroporation. In *Streptococcal Genetics*, Ferretti, J. J. & Curtiss, R., III eds), pp.229-233. American Society for Microbiology, Washington, DC.

Håverstein, L. S., G. Coomaraswamy & D. A Morrison. 1995. An unmodified heptadecapeptide induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proceedings of the National Academy of Sciences, USA*, 92, 11140-11144.

Håverstein, L. S. & D. A Morrison. (1999). Quorum-sensing and peptide pheromones in streptococcal competence for genetic transformation. In *Cell-cell signalling in Bacteria*, Dunny, G. M. & Winans, S. C., eds., pp. 9-26. ASM Press, Washington, D.C.

Hershey, A. D. & Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology*, 36, 39-56.

Inamine, G. S. & Dubnau, D. (1995). ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *Journal of Bacteriology*, 177, 3045-3051.

Kado, C. I. (1991). Molecular mechanism of crown gall tumorigenesis. *Critical Reviews in Plant Science*, 10, 1-32.

Kahn, M. E., Maul, G. & Goodgal, S. H. (1982). Possible mechanisms for donor DNA binding and transport in *Haemophilus*. *Proceedings of the National Academy of Sciences, USA*, 80, 6927-6931.

Karudapuram, S. & Barcak, G. J. (1995). The *Haemophilus influenzae dprABC* genes constitute a competence-inducible operon that requires the product of the *tfoX* (*sxy*) gene for transcriptional activation. *Journal of Bacteriology*, 179, 4815-4820.

Karudapuram, S., Zhao, X. & Barcak, G. J. (1995). DNA sequence and characterization of *Haemophilus influenzae dprA*⁺, a gene required for chromosomal but not plasmid DNA transformation. *Journal of Bacteriology*, 177, 3235-3240.

Kunst, F. et al. (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* 390, 249-256.

Lacks, S. (1962). Molecular fate of DNA in genetic transformation of pneumococcus. *Journal of Molecular Biology*, 5, 119-131.

Lacks, S. (1977a). Binding and entry of DNA in pneumococcal transformation. In *Modern Trends in Bacterial Transformation and Transfection*, Portoles, A., Lopez, R. & Espinosa, M. (eds.), pp. 35-44. Elsevier/North-Holland, Amsterdam.

Lacks, S. (1979a). Steps in the process of DNA binding and entry in transformation. In *Transformation 1978*, Glover, S. W. & Butler, L. O., eds., pp. 27-41. Cotswold Press, Oxford.

Lacks, S. (1979b). Uptake of circular deoxyribonucleic acid and mechanism of deoxyribonucleic acid transport in genetic transformation of *Streptococcus pneumoniae*. *Journal of Bacteriology*, 138, 404-409.

Lacks, S. A. (1977b). Binding and entry of DNA in bacterial transformation. In *Microbial interactions*, Reissig, J. L., ed., pp. 179-232. Chapman and Hall, London.

Lacks, S. A. (1988). Mechanisms of genetic recombination in gram-positive bacteria. In *Genetic Recombination*, Kucherlapati, R. & Smith, G. R., eds., pp. 43-85. American Society for Microbiology Publications, Washington, D. C.

Lacks, S. & Greenberg, B. (1973). Competence for deoxyribonucleic acid uptake and deoxyribonuclease action external to cells in the genetic transformation of *Diplococcus*

pneumoniae. *Journal of Bacteriology*, 114, 152-163.

Lacks, S. & Greenberg, B. (1976). Single-strand breakage on binding of DNA to cells in the genetic transformation of *Diplococcus pneumoniae*. *Journal of Molecular Biology*, 101, 255-275.

Lacks, S. & Neuberger, M. (1975). Membrane location of a deoxyribonuclease implicated in the genetic transformation of *Diplococcus pneumoniae*. *Journal of Bacteriology*, 124, 1321-1329.

Lacks, S., Greenberg, B. & Carlson, K. (1967). Fate of donor DNA in pneumococcal transformation. *Journal of Molecular Biology*, 29, 327-347.

Lacks, S., Greenberg, B. & Neuberger, M. (1974). Role of a deoxyribonuclease in the genetic transformation of *Diplococcus pneumoniae*. *Proceedings of the National Academy of Sciences, USA*, 71, 2305-2309.

Lacks, S., Greenberg, B. & Neuberger, M. (1975). Identification of a deoxyribonuclease implicated in genetic transformation of *Diplococcus pneumoniae*. *Journal of Bacteriology*, 123, 222-232.

Lacks, S. A. & Springhorn, S. S. (1984). Transfer of recombinant plasmids containing the gene for *DpnII* DNA methylase into strains of *Streptococcus pneumoniae* that produce *DpnI* or *DpnII* restriction endonucleases. *Journal of Bacteriology*, 158, 905-909.

Larson, T. G. & Goodgal, S. H. (1991). Sequence and transcriptional regulation of *com101A*, a locus required for genetic transformation in *Haemophilus influenzae*. *Journal of Bacteriology*, 173, 4683-4691.

Lederberg, J. & Tatum, E. (1946). Gene recombination in *E. coli*. *Nature, London*, 158, 558.

Lerman, R. S. & Tolmach, L. J. (1957). Genetic transformation. I. Cellular incorporation of DNA accompanying transformation in pneumococcus. *Biochimica et Biophysica Acta* 28, 68-82.

Levengood, S. K. & Webster, R. E. (1989). Nucleotide sequences of the *tolA* and *tolB* genes and localization of their products, components of a multistep translocation system in *Escherichia coli*. *Journal of Bacteriology*, 171, 6600-6609.

Lipinska, B., Krishna Rao, A. S. M., Bolten, B. M., Balakrishnan, R. & Goldberg, E. B. (1989). Cloning and identification of bacteriophage T4 gene 2 product gp2 and action of gp2 on infecting DNA in vivo. *Journal of Bacteriology*, 171, 488-497.

Londono-Vallejo, J. A. & Dubnau, D. (1993). *comF*, a *Bacillus subtilis* late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases. *Molecular Microbiology*, 9, 119-131.

Londono-Vallejo, J. A. & Dubnau, D. (1994). Membrane association and role in DNA uptake of the *Bacillus subtilis* PriA analog ComF1. *Molecular Microbiology*, 13, 197-205.

Lopez, P., Espinosa, M., Stassi, D. L. & Lacks, S. A. (1982). Facilitation of plasmid transfer in *Streptococcus pneumoniae* by chromosomal homology. *Journal of Bacteriology*, 150, 692-701.

Lorenz, M. G. & Wackernagel, W. (1994). Bacterial gene transfer by natural genetic transformation in the environment. *Microbiological Reviews*, 58, 563-602.

Lunsford, R. D. (1998). Streptococcal transformation: Essential features and applications of a natural gene exchange system. *Plasmid*, 39, 10-20.

Lunsford, R. D., Nguyen, N. & London, J. (1996). DNA-binding activities in *Streptococcus gordonii*: Identification of a receptor-nickase and a histonelike protein.

Current Microbiology, 32, 95-100.

Lunsford, R. D. & Roble A. G. (1997). *comYA*, a gene similar to *comGA* of *Bacillus subtilis* is essential for competence-factor-dependent DNA transformation in *Streptococcus gordonii*. *Journal of Bacteriology*, 179, 3122-3126.

Mandel, M. & Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, 53, 159-162.

Manoil, C. & Rosenbusch, J. P. (1982). Conjugation-deficient mutants of *E. coli* distinguish classes of functions of the outer membrane OmpA protein. *Molecular and General Genetics*, 187, 148-156.

Martin, B., P. Garcia, M. P. Castanie & J. P. Claverys. (1995). The *recA* gene of *Streptococcus pneumoniae* is part of a competence-induced operon and controls lysogenic induction. *Molecular Microbiology*, 15, 367-379.

Matson, S. W., Nelson, W. C. & Morton, B. S. (1993). Characterization of the reaction product of the *oriT* nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *Journal of Bacteriology*, 175, 2599-2606.

Mejean, V. & Claverys, J. P. (1993). DNA processing during entry in transformation of *Streptococcus pneumoniae*. *Journal of Biological Chemistry*, 268, 5594-5599.

Miao, R. & Guild, W. R. (1970). Competent *Diplococcus pneumoniae* accept both single-and double-stranded deoxyribonucleic acid. *Journal of Bacteriology*, 101, 361-364.

Morrison, D. A. (1978). Transformation in pneumococcus: protein content of eclipse complex. *Journal of Bacteriology*, 136, 548-557.

Morrison, D. A. (1997). Streptococcal competence for genetic transformation: Regulation by peptide pheromones. *Microbial Drug Resistance* 3, 27-37.

Morrison, D. A. & Baker, M. F. (1979). Competence for genetic transformation in *pneumococcus* depends on synthesis of a small set of proteins. *Nature, London* 282, 215-217.

Morrison, D. A. & Guild, W. R. (1972). Transformation and deoxyribonucleic acid size: extent of degradation on entry varies with size of donor. *Journal of Bacteriology*, 112, 1157-1168.

Morrison, D. A. & Guild, W. R. (1973). Breakage prior to entry of donor DNA in *pneumococcus* transformation. *Biochimica et Biophysica Acta*, 299, 545-556.

Morrison, D. A., Lacks, S. A., Guild, W. R. & Hageman, J. M. (1983). Isolation and characterization of three new classes of transformation-deficient mutants of *Streptococcus pneumoniae* that are defective in DNA transport and genetic recombination. *Journal of Bacteriology*, 156, 281-290.

Nakamura, A., Koide, Y., Miyazaki, H., Kitamura, A., Masaki, H., Beppu, T. & Uozumi, T. (1992). Gene cloning and characterization of novel extracellular ribonuclease of *Bacillus subtilis*. *European Journal of Biochemistry*, 209, 121-127.

Notani, N. & Goodgal, S. H. (1966). On the nature of recombinants formed during transformation in *Hemophilus influenzae*. *Journal of General Physiology*, 49, part 2, 197-209.

Orren, D. K. & Sancar, A. (1989). The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. *Proceedings of the National Academy of Sciences, USA*, 86, 5237-5241.

Pakula, R. & Walczak, W. (1963). On the nature of competence of transformable streptococci. *Journal of General Microbiology*, 31, 125-133.

Palmen, R., Driessen, A. J. M. & Hellingwerf, K. J. (1994). Bioenergetic aspects of the translocation of macromolecules across bacterial membranes. *Biochimica et Biophysica Acta*, 1183, 417-451.

Pearce, B. J., Naughton, A. M. Campbell, E. A. & Masure, H. R. (1995). The *rec* locus, a competence-induced operon in *Streptococcus pneumoniae*. *Journal of Bacteriology*, 177, 86-93.

Pestova, E. V. & Morrison, D. A. (1998). Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a lacZ reporter insertion vector. *Journal of Bacteriology*, 180, 2701-2710.

Piechowska, M. & Fox, M. S. (1971). Fate of transforming deoxyribonucleate in *Bacillus subtilis*. *Journal of Bacteriology*, 108, 680-689.

Pifer, M. L. (1986). Plasmid establishment in competent *Haemophilus influenzae* occurs by illegitimate transformation. *Journal of Bacteriology*, 168, 683-687.

Provvedi, R. & Dubnau, D. (1999). ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Molecular Microbiology*, 31, 271-280.

Puyet, A., Greenberg, B. & Lacks, S. A. (1990). Genetic and structural characterization of EndA, a membrane-bound nuclease required for transformation of *Streptococcus pneumoniae*. *Journal of Molecular Biology*, 213, 727-738.

Redfield, R. J. (1991). *sxy-1*, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. *Journal of Bacteriology*, 173, 5612-5618.

Rees, C. E. D. & Wilkins, B. M. (1990). Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. *Molecular Microbiology*, 4, 1199-205.

Rosenthal, A. L. & Lacks, S. A. (1977). Nuclease detection in SDS-polyacrylamide gel

electrophoresis. *Analytical Biochemistry*, 80, 76-90.

Rosenthal, A. L. & Lacks, S. A. (1980). Complex structure of the membrane nuclease of *Streptococcus pneumoniae* revealed by two-dimensional electrophoresis. *Journal of Molecular Biology*, 141, 133-146.

Sabelnikov, A. G. (1994). Nucleic acid transfer through cell membranes: towards the underlying mechanisms. *Progress in Biophysics and Molecular Biology*, 62, 119-152.

Sabelnikov, A. G., Greenberg, B. & Lacks, S. A. (1995). An extended -10 promoter alone directs transcription of the *DpnII* operon of *Streptococcus pneumoniae*. *Journal of Molecular Biology*, 250, 144-155.

Sastry, P. A., Pasloske, B. L., Paranchych, W., Pearlstone, J. R. & Smillie, L. B. (1985). Nucleotide sequence and transcriptional initiation site of two *Pseudomonas aeruginosa* pilin genes. *Journal of Bacteriology*, 164, 571-577.

Saunders, C. W. & Guild, W. R. (1981). Monomer plasmid DNA transforms *Streptococcus pneumoniae*. *Molecular and General Genetics*, 181, 57-62.

Seto, H. & Tomasz, A. (1974). Early stages in DNA binding and uptake during genetic transformation of pneumococci. *Proceedings of the National Academy of Sciences, USA*, 71, 1493-1498.

Seto, H. & Tomasz, A. (1976). Calcium-requiring step in the uptake of deoxyribonucleic acid molecules through the surface of competent pneumococci. *Journal of Bacteriology*, 126, 1113-1118.

Shirasu, K. & Kado, C. I. (1993). Membrane location of the Ti plasmid VirB proteins involved in the biosynthesis of apilin-like conjugative structure on *Agrobacterium tumefaciens*. *FEMS Microbiological Letters*, 111, 287-294.

Solomon, J. M. & Grossman, A. D. (1996). Who's competent and when: regulation of natural genetic competence in bacteria. *Trends in Genetics*, 12, 150-155.

Smith, H., Wiersma, K., Bron, S. & Venema, G. (1983). Transformation in *Bacillus subtilis*: purification and partial characterization of a membrane-bound DNA-binding protein. *Journal of Bacteriology*, 156, 101-108.

Strom, M. S., Nunn D. N. & Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proceedings of the National Academy of Sciences, USA*, 90, 2404-2408.

Sugawara, E. & Nikaido, H. (1992). Pore-forming activity of OmpA protein of *Escherichia coli*. *Journal of Biological Chemistry*, 267, 2507-2511.

Sun, T. P. & Webster, R. E. (1986). *fii*, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant *tolA* and *tolB*. *Journal of Bacteriology*, 165, 107-115.

Tarahovsky, Y. S., Khusainov, A. A., Deev, A. A. & Kim, Y. V. (1991). Membrane fusion during infection of *Escherichia coli* cells by phage T4. *FEBS Letters*, 289, 18-22.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673-4680.

Tomasz, A. (1965). Control of the competent state in *Pneumococcus* by a hormone-like cell product: an example for a new type of regulatory mechanism in bacteria. *Nature, London*, 208, 155-159.

Tomasz, A. & Hotchkiss, R. D. (1964). Regulation of the transformability of

pneumococcal cultures by macromolecular cell products. *Proceedings of the National Academy of Sciences, USA*, 51, 480-487.

Tomasz, A. & Mosser, J. L. (1966). On the nature of the pneumococcal activator substance. *Proceedings of the National Academy of Sciences, USA*, 55, 58-66.

Tomb, J.-F., El-Hajj, H. & Smith, H.O. (1991). Nucleotide sequence of a cluster of genes involved in the transformation of *Haemophilus influenzae* Rd. *Gene*, 104, 1-10.

Trautner, T., Pawlek, B., Bron, S. & Anagnostopoulos, C. (1974). Restriction and modification in *Bacillus subtilis*. Biological aspects. *Molecular and General Genetics*, 131, 181-191.

van Sinderen, D., Kiewiet, R. & Venema, G. (1995). Differential expression of two closely related doxyribonuclease genes, *nucA* and *nucB*, in *Bacillus subtilis*. *Molecular Microbiology*, 15, 213-223.

van Sinderen, D., Luttinger, A., Kong, L. Dubnau, D., Venema, G. & Hamoen, L. (1995). *ComK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. *Molecular Microbiology*, 15, 455-462.

Wilkins, B. M. (1995). Gene transfer by bacterial conjugation: diversity of systems and functional specializations. In *Population Genetics of Bacteria*, Baumberg, S., Young, J. P. W., Wellington, E. M. H. & Saunders, J. R., eds., pp. 59-88. University Press, Cambridge.

Zulty, J. J. & Barcak, G. J. (1995). Identification of a DNA transformation gene required for *com101A* expression and supertransformer phenotype in *Haemophilus influenzae*. *Proceedings of the National Academy of Sciences, USA*, 92, 3616-3620.

FIGURE LEGENDS

Fig. 1. Modes of DNA uptake. A. Viral infection: gp2, phage-encoded protein; r, cell membrane receptor; tf, phage tail fiber. B. Conjugation: d, TraD, i, TraI, and t, other plasmid *tra* gene products; om, OmpA, and x, unknown, host proteins; h, donor cell helicase; s and ts, single-stranded DNA binding proteins. C. Transformation: End, membrane-bound endonuclease releases oligonucleotides (Oligos); other proteins, EA, EB, FA, FB, and G, are induced during competence development and have structural and functional roles in uptake; ts is a competence-induced single-stranded DNA binding protein. The mechanisms shown are in part hypothetical.

Fig. 2. Comparison of homologous ComG-like (Cgl) proteins related to type IV pilins. A. ComGC-like (CglC) proteins. B. *H. influenzae* HI0299 (PilA) compared to PaePilA, the type IV prepilin from *P. aeruginosa*. C. ComGD-like (CglD) proteins. Only the N-terminal half of SgoComYD was cloned. LlaCglD was translated from the reported sequence data after insertion of single nucleotide before position 1509. D. ComGF-like (CglF) proteins. Key to species: Spn, *S. pneumoniae*; Sgo, *S. gordonii*; Spy, *S. pyogenes*; Lla, *L. lactis*; BsU, *B. subtilis*. Key to symbols indicating similarity of residues: *, identical; :, very similar; ., similar. This Figure and Figs. 3, 4 and 5 were constructed with the aid of the CLUSTAL W program (Thompson, Higgins & Gibson, 1994).

Fig. 3. Comparison of *S. pneumoniae* CelA and *B. subtilis* ComEA to each other and to the carboxyl half of *E. coli* UvrC. Symbols as in Fig. 2.

Fig. 4. Comparison of homologous ComF-like (Cfl) proteins. A. *S. pneumoniae* CflA and *B. subtilis* ComFA proteins compared to each other and to *E. coli* UvrB. B. Comparison of *S. pneumoniae* CflB, *B. subtilis* ComFB, and *H. influenzae* HI0434 proteins. Symbols as in Fig. 2.

Fig. 5. Comparison of homologous endonucleases. A. *S. pneumoniae* EndA compared to putative *S. pyogenes* nucleases SpyEndA and SpyEndB. B. Comparison of *E. coli* and *B. subtilis* EndA nucleases. Symbols as in Fig. 2.

Fig. 6. Proposed mechanism for DNA uptake in *Streptococcus pneumoniae*. A. DNA binding. All components except EndA, a membrane nuclease, and MurA, a cell wall synthetic enzyme are induced during competence development. CoiA, a peptidase, and MurA reconstruct the cell wall for insertion of the DNA pore structure. CglA and CglB are responsible for export of CglC-G, prepilin like components of the uptake structure. The latter are processed by CilC; their final arrangement in the cell envelope is unknown and arbitrarily drawn in the Figure. CelA binds to donor DNA, and in conjunction with CflA makes a single strand break in it, possibly attaching the new 3'-end covalently to CflA or CflB. B. DNA entry. CflB may act as a pilot protein bringing the single strand through a membrane pore formed by CelB. Entry of the donor strand may be facilitated by degradation of the complementary strand by EndA and helicase action of CflA. A competence-induced protein, Ssb, binds to the incoming donor strand. Potassium ions are needed for binding, and calcium and magnesium ions for entry.

Table 1. *Genes implicated in DNA uptake*^a

<i>B. subtilis</i>	<i>S. pneumoniae</i>	<i>H. influenzae</i> ^b	Possible functions
	<i>coiA</i>		Cell wall degradation
	<i>murA</i> ^c		Cell wall resynthesis
<i>comGA</i>	<i>cglA</i>	[HI0298] <i>pilB</i>	Prepilin membrane transport pore
<i>comGB</i>	<i>cglB</i>	[HI0297] <i>pilC</i>	Prepilin transport ATPase
<i>comGC</i>	<i>cglC</i>	[HI0299] <i>pilA</i>	Pilin-like wall structure
<i>comGD</i>	<i>cglD</i>		Pilin-like wall structure
<i>comGE</i>	<i>cglE</i>		Pilin-like wall structure
<i>comGF</i>	<i>cglF</i>		Pilin-like wall structure
<i>comGG</i>	<i>cglG</i>		Pilin-like wall structure
<i>comC</i>	<i>cilC</i>	[HI0296] <i>pilD</i>	Prepilin processing
<i>comEA</i>	<i>celA</i>	[HI1018] ^c	Donor dsDNA binding
<i>comEC</i>	<i>celB</i>	[HI0061] <i>rec2</i>	DNA membrane transport pore
<i>comFA</i>	<i>cflA</i>		Nickase, translocase, helicase?
<i>comFC</i>	<i>cflB</i>	[HI0434] <i>com101A</i> (<i>comF</i>)	Pilot protein?
<i>bsn</i> ^c (<i>endA</i>)	<i>endA</i>		Degradates donor DNA strand
<i>nucA</i> ^c	-		Alternative degradative nuclease
<i>ssbB</i> ^c	<i>cilA</i> ^c		ssDNA binding protein
	<i>cilB</i> ^c	[HI0985] <i>dprA</i>	Late DNA processing

^a See text for references. All genes except *murA*, *bsn* and *endA* are competence-inducible.

^b Genome designations in brackets; additional gene names in parentheses.

^c No experimental evidence for uptake.

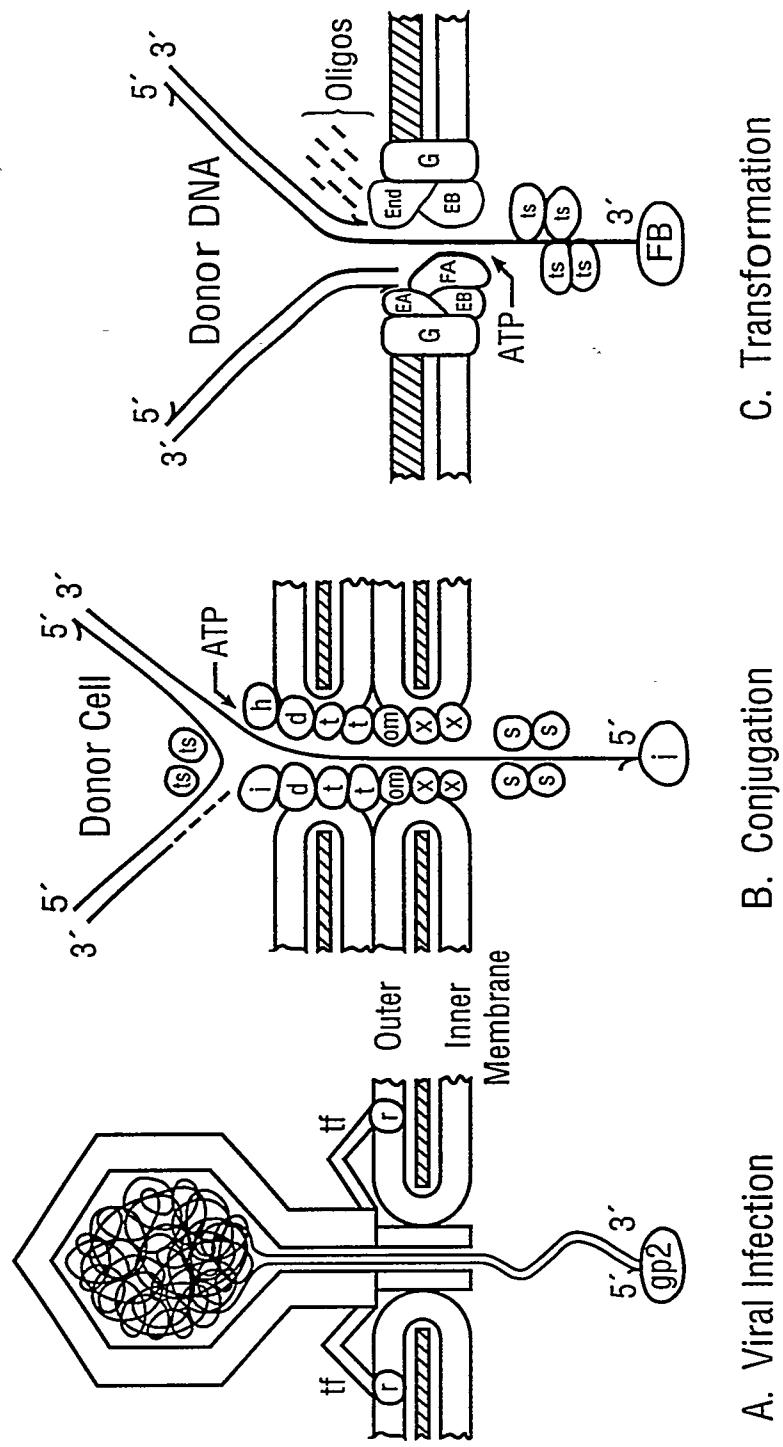


Fig. 1

A.

SpnCg1C	-----MKKMMTFLKKAKVKAFTLVEMLVVLIIISVLFLLFVPNLTQKEAVNDKGKAAVVKVVESQAElysle	68
SgoComYC	-----MNKLKKLRLVKRAFTLVEMLVVLIIISVLMILFVPNLTQKEAVSDTGNAAVVKVVESQAElyElk	64
SpyCg1C	-----MINQWNNLRRHKKLKGFTLLEMLLVLVILVISVLMILFVPNLSKQKDRVTETGNAAVVKLVENQAElyEl	67
LlaCg1C	MKIENITLIMIKALSLIKIHGRKLWQKKQKAFTLLEMLLVLVLAISIILLLFVPNLIKEAKQVKTGEAAVVKVVESQAElyEl	84
BsuComGC	-----MNEKGFTLVEMLVLFIIISILLITIPNVTKHNTIQKKCEGLQNMVKAQMtaFELD	58

SpnCg1C	KN-EDASLRKLQADGRITEEQAKAYKEYNDKNGGANRKVN	108
SgoComYC	NTGQATLSQLVAGNISQKQADSYKAYYGKNNSETQAVAN	105
SpyCg1C	QG-SKPSLSQLKADGSITEQEKAQYDYYDKHKNEKARLSN	108
LlaCg1C	HNDKPNLSELLSAGMITQKQVTAYDDYYDKHNNQERNFND	125
BsuComGC	HEGQTPLSLDLQSGEQYVKKDAPCVPNGKR1IITGGEVK-VEH	98

B.

C.

SpnCg1D	MDASRKNRLLKLKNTMIKMEQIVKSMIKAFTMLESLLVGLVSI	LA	LGSGSVQSTFSAVEEQIFFMEFEELYRETQKRSVAS	84	
SgoComYD	-----MVRTIVRLRQLPIKAFTVLESLLVLMISSFILLALSSSVQATF	FE	EQI	QAKIFFLEFEHFYQESQKLSVSS	69
SpyCg1D	MVVSLRNKKLTIKMTNIKMKKPVLAIKAFTLLETLLSLSVMSFI	IL	LGSLSPVTKSYQKVEEHLFFSHFEHLYRHQQKLALIQ	84	
LlaCg1D	-----MMITMTRIKMNSAISMIRAFTLLESLLVVLLIVSFI	LL	FSFAELTQTVHLFKGRLFVLFQLENLYKISQE-NAAL	73	
BsuComGD	-----MNIIKLNEEKGFTI	LL	LSASILLVAVFTLPPAYDNTAVRQAASQLKNDIMLTQQT	64	

SpnCg1D	QQKTSNLNDGQT--LSNGSQKLPVPGIQAPSGQSTTFDRAGGNSSLAKVCFQT--SKGAIY--QLYLNGKIKRIKETKN--	160
SgoComYD	QRKLVLEISSQE--ISNGYARL-----	89
SpyCg1D	QKQRVLDISSTK--IVTEGNSLTVPKSITVNHPYRLVIDQMGGNHSLSAKIIFDM--TDRRFKY--QFYLGSGNYQKTSQSLHSP	162
LlaCg1D	QSSSNLGSENKG--LIYENKEIDIPKEVEMA--EFLIIFDEDGGNSSLQKIKVYLPHYEKKTILY--QMEMGSGKYKKKIN-----	146
BsuComGD	QORTKILFHKKEYQLVIGDTVIERPYATGLSIELLTLKDRLFNEKGHPNAGGKIRVKHGAVDITVYLGSGRVNVERK--	143

10

SpnCg1F	-----MVQNSCWQSKSHVKVAKTLLESLLALIVISGGLLLQFQAMSQLLISEVRYQQQSEQKEWLLFVDQLEVELDRSQQF	75
SpyCg1F	-----MSKQLSNIKAFTLLLEALIALLVISGSLLVYQGLTRTLLKHSYHLARHDQDNWLLFSHQLREELSGARFY	69
LlaCg1F	-----MTMERKFCDLKLKIRAFTLLECLVALLAISGSVLVISGLTRMIEEQMKIQSNDRSRKDWQFCEQMRSELSGAKLD	75
BsuComGF	MKTYVSKPQLIKKNHFASAFCRQNGYTLNVLFLSLSVFLLISGSLAAIHLFLSRQEQHDGFTQQEWMSIEQMMNECKESQAV	84

SpnCg1F	KVEGN-RLYMKQ--DGKDIAGKSDDFRKTNARGRGYQPVMYGLKSVRITEDNQLVRFHQFQKG--LEREF-IYRVEKEKS	153
SpyCg1F	KVADN-KLYVEK--GKKVLAQGQFKSHDFRKSASNGKGYQPMLFGISRSHIIEQSQICITLWKWSG--LERTF-YYAFQD--	144
LlaCg1F	NVNQN-FLYVTK--DKK-LRFG-LVGDDFRKSDDKGQQGYQPMLYDLGAKIQAEEENLIKITIDFDNG--GERVF-IYRFTDTK-	150
BsuComGF	KTAEHGSVLICTNLSQGDIFRD-IYHSMIRKVD-GKGVPTLDHITAMKADIENGVLLKIESEDQKVYQTAPVPSYSLGGG-	165

Fig. 2

SpnCelA	MEAIIEKIKIKEYKIIIVICTGLGLLVGGFFLLKAPQTPVKETNLQAEVAA---VSKDLVSEKEVNKEEKEEPEQDL	73
BsuComEA	MNWLNQHKKA--IIIAASAAVETAIMIFLATGKNEPVKQA----VPT---ETENTVVKQEANNDESNE-----T	61
EcoUvrC	KLPEVKRMCECPDISHTMGEQTVASCVVFANGPLRAEYRRYNTGITPGDDYAAMNQVLRRYRGKAIDDSKIP-DV	462
SpnCelA	ITVD-VKGAVKSPGIYDLPVGSRINDAVQKAGGLTPEQADSKSLNLAQKVSDEA--LVYVPTKGEEAVSQQTGLGTA	146
BsuComEA	IVID-IKGAVQHPGVYEMRTGDRVSQAIEKAGGTSEQADEAQVNLAELQDGT--VYIPIKKGEETAVQQGGGSV	134
Hin1008	-----MKTLLFTSVVLCGALVSSSFAEK--ATXQTAQSVVTTQAEAQVAPA	45
EcoUvrC	ILIDGGKGQLAQAKNVFAELDVSWDKNHPLLGVAKGADRKAAGLETLFFPEPEEGFSLPPDSPALHVIQHIRDESH	538
SpnCelA	SSIS-KEKKVNLNKASLEELKQ-VKGLGGKRAQDIIDHR--EANGKFKSVDELKKVSGIGGKTIIEKLKDVTVD	216
BsuComEA	QSDGGKGALVNINTATLEELQG-ISGVGPSKAEAIAYR--EENGRFQTIEDITKVSGIGEKSFEKIKSSITVK	205
Hin1008	VVSD---KLNINTATASEIQLSITGIGAKKAEAVQYR--EKGHNFXNAEQLLEVQGIGKATLEKVRDRIF-	112
EcoUvrC	DHAIGGHRKKRRAVKNTSSLET-IEGVGPKRRQMLLKYMGLQGLRNASVEEIAKVPGISQGLAEKIFWSLKH-	610

Fig. 3

A.

SpnCflA	-MVKNLVDYLGLRLFTENELTEERQLA-----EKLPMAMRKE-----KGKLFQCRCNST-----ILEEWYLP	54
BsuComFA	-MNVPVEKNSFSKELQQTLRSRHLRLTLEFSDEMIIEWHKNGYITAENSISINKRYYRCNRQGTDQ-----RYFSFYHS	76
EcoUvrB	YVPSSDTFIEKDAVSNEHIEQMRLSAT-----KAMLERRDVVVVAWSAIYGLGDPDLYLKMLHLTVGMIIDQRAILR	174

SpnCf1A	IG--AYYCR--EC1LLMK--RVRSDQTLYYFPQE--DFP--KQDVLKWRGQLTPFQEKVSEGGLQVVDKQKPTLVHAVTGA	124
BsuComFA	SGKKNKLYCR--SCVMMG--RVSEEVPLYSWKEE--NESWNKSIKLTDWGKLSSGQQKAANVIEAISKKEELLIWAVCGA	150
EcoUvrB	RLAELQYARNDQAFQRGTFRVRGEV--IDIFPAESDDIALRVELFDEEVERLSLFDPLTGQIVSTIPRFTIYPKTHYVTPR	253

SpnCflA	GTKEMIYQVVAKVINAGGAVALCLASPRIDVCLELYKRLQQDFS-CGIALLHGESEPYFR-TPLVVA--TTHQLLKFYQAFD	200
BsuComFA	GKTEMLFPGIESALNQGLRVCIAPTRTDVVLEAPRLKAAFQGADISALYGGSSDDKGRLSPLMIS--TTHQLLRYKDAID	228
EcoUvrB	ERIVQAMEEIKEELAARRKVLLENNKKLLEQRLTQRTQFDLEMMNELGYCSCGIENYSRFLSGRPGEPPPTLFDYLPADG	333

SpnCf1A	LLIVDEVDAFPYVDNPMLY--HAVKNSVKENGRLRIF--LTATSTNEKKVRLGELKRLN-LPRRFHGNP---LIIPKPI	272
BsuComFA	VMTIIDEVDAFPYSADQTLQ--FAVQKARKGNSTLVY--LSATPPKELKRKALNGQLHSVR-IPARHHRKP---LPEPRFV	300
EcoUvrB	LLVVDESHVTIPQIGGMYRGDRARKETLVEYGFRLPSALDNRPLKPEEEFAEALAPQTIYVSATPGNYELEKSGGDVVQDVQV	413

SpnCf1A	WLSDFNRYLDKNR-LSPKLKSYIE--KQRKTAYPILLIFASEIK-KGEQLAELQEQFPNEKIGFVSS--VTEDRLEQVQA	346
BsuComFA	WCGNWKKKLNRNK-IPPAVKRWIE--FHVKEGRPVFLFVPSVS-ILEKAAACFKG--VHCRTASVHA--EDKHRKEVQQ	372
EcoUvrB	RPTGLLDPFIEVRPVATODV DLLSEI RORA INERVL VTTLTKRMAE DLT EYLEE--HGERVRYLHSIDT DVERMEIIRD	491

SpnCf1A	FRDGELETILISTTILERGVTFPCV-DVFFVVEANHRLFTKS--SLIQIGGRVGR-----	-SMDRPTGDLFF	408
BsuComFA	FRDGQLDLLTTTILERGVTFPKV-QTGVLGAESSIONTES--ALVQIAGRGTR-----	-HKEYADGDIVYF	434
EcoUvrB	LRIGEFDVLVGINLLREGIDMPEVSLVAILDADKEGFLRSERSLIQTIGRAARNVNGKAILYGDKITPSMAKAIGETERR	-	471

SpnCf1A	- - - - - H - - DGLNASIKKAIKEIQMMNKEAGL - - - - -	432
BsuComFA	- - - - - H - - FGKTKSMLDARKHIKEMNELAAKVECTD	463
EcoUvrb	REKQQKYNEEHGITPQGLNKKVVLDILALCQNIAKTKAKGRGKSRPIVEPDNVPMDMSPKALQQKIHLEGLMMQHA	647

B.

SpnCf1B	MKCLLCGQTMTKTVLTFSSLLLLRNDDSCLCSDCDSTFERIGEENCPNCM-KTE-LSTKQCDQCLW---CKEGVEVS-HRAI	75
BsuComFC	MICLLCDSQFSQDVWTWRLFLLKPDE-KVCYCSRSKLLKKITGHICPLCG-RPQSVHAVCRDCEVWRTRIRDSSLRQNRSV	79
HI0434	MMNFFN---FRCIHCGRGNLHIAKNG---LCSGCKQKIKSFP--YCGHCGSELQYYAQHCGNCLKQ----EPSWDK-MVII	67

SpnCf1B	FTYNQAMKDFFSRYKFDGDFLLRKVFA\$ - FLSEELKKYKEYQF - - VVIP - - LSPDRYANRGFNQVEGLVEAAG - FEYLD	148
BsuComFC	YTYNNDMMKETLSRKFGRDAEIIINAFKSDFSSTFSKVPDKHF - - VLVPIPLSKEREERGFNFQAHLLAECLDRPSHHP	156
HI0434	GHYIEPLSILIJQRKFQPNQFWIDRTLARLLYLAVRDAKRTQHLKPEAIIPVPLVYHFRQWRRGYNQADLLSQQLSRWLIDP	158

```

SpnCf1B  LLE-----KREERASSKKNRSERLGTTELPIFIKSGVTIPKKILLIDDIYTTGATINRVKKLLEEAG-AKDVKTFSLVR- 220
BsUComFC LIR-----LNNEKQSK-KKKTERLLESECIFDTKNNSAEGMNIIILIDDLYTGTGATLHFAAARCLLEKGKAASVSSFTLIRS 229
HI0434  NLNNIVKRVKVKHTYTORGLSAKDRRONIKNAPESLAWSKNEEPYRRVAIVDDVITTGSTI.NEISKLRLKIG-VEETOWGLARA 229

```

Fig. 4

A.

SpnEndA MNKKTRQTLIGLLVLLLSTGSYYIKQMPSAPNSPKTNLSQQASEAPSQALAESVLTDAVKSQIKGSLEWNGSGAF 78
SpyEndB MSKSNRRTWQGLVVLIAILTTFTTSTVTAARKIR-----NFPDTTEILLGTAKATEPGILPTGSYQL 64
..*: * **: *: : : . . . : : : : * : : * * : ** :

SpnEndA IVNGNKTNLDAVSSKPYADNKTKTVGKETVPTVANALLSKATRQYKNRKTGNGSTS WTPPGWHQVK--NLKGSYTH 154
SpyEndA MLTRKKRKRVQEQQVKPKRANALLTKSTRQYRNRYQTDNAYRNFKPAGWHRLH--HLKGSYDH 60
SpyEndB VL---GDLDN-LQRPTFAHIQLKDQE-----PNIKRKG--LKFNPWHNYKLTDANGKTTW 116
*: : ** : : : : * : : : : * : : : : * : : : : * :

SpnEndA AVDRGHLGYALIGGLGFDASTSNPKNIAVQTAWANQAQAEYSTGQN-YYES-----KVRKALDQN 215
SpyEndA AVDRGHLGYALVGGLGFDASTGNPDNIATQLSWANQANKPYLTGQN-YYEGLVSTELEYCT----- 122
SpyEndB LMDRGLVGYQF-SGLN-----DEPKNLVTMTKYLNTGFSKDNPCLMYYENRLDSWLALHNFWLDYKVTVPVYHKN 189
*****: ** : : **: : : * : : : : * : : : : * : : : : * :

SpnEndA KRVRYRVTLYYA--SNEDLVPSASQIEAKSSDGELEFNVLPNVQKGLQLDYRTG-EVTVTQ----- 274
SpyEndB ELVPRQVVLQYVGIDENGDLLQIKLGSEKEVDNFGVTSVLDNVSPLAELDYQTGMMLDSTQNEEDSNLEEEFEEAA 268
* : *.*. . * **: : * : * . . * : * . . * : * : : * :

B.

EcoEndA MYR---YLSIAAVVLSAAFSGPALAEG-----INS--FSQAKAAAVKVHADAPGTYCGCKINWQGKKGVV--DLQS 65
BsuEndA MTKKAFLPLVCVLLISGWLAPAASASAQTTLSLNDRLASSPGTGSLLSLAAPAAPYADTDIYYEGAEGKTGDSLKS 78
*: : : * . . . : : : * : : : * : : : * : : : * : : : * :

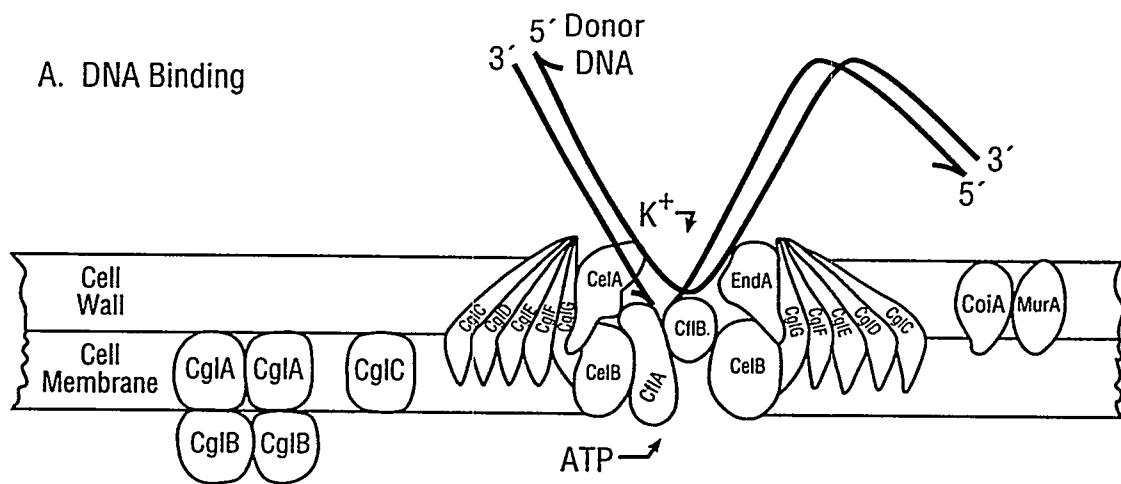
EcoEndA CGYQVRKNENRASRVEW-----EHVVP-----AWQFGHQRCWQDGGRKNCAKDPVYRKME-----SDM 119
BsuEndA TLHRIISGHTMLSYSEVWNALKETDEDPRNPNVNILLYTNERSRSKNLNGGNVGDWNREHVWAKSHGDFGTSKGPGTDI 156
*: : . . * . . : : * : : : : : : * . : : : * : : : * :

EcoEndA HNLQPSVGEVNGDRGNFMYSQWNGGEGQYQGQCAMKVDFKEKAAEPPARARGATARTYFYMR-----DQY----- 183
BsuEndA HHLRPADVQVNSARGNMDFD--NGG-TEY-AKAPGNYYDGDSWEPRDDVKGDVARMLFYMAVRYEGDDGYPDLENDK 230
*: : *: : * . . * **: : * : * . . . : * . : * : * : * :

EcoEndA --NLTLSRQQTQ-LFNAWNKMYPTWDWECERDERI-AKVQGNHNPYVQRACQARKS-- 235
BsuEndA TGNGSAPYHGKQSVILLEWNKQDPVDDRERKRNEIIYEKYQHNRNPFDIDHPEWADEIWP 288
*: : . . * : : * **: * * : * : * : * : * : * :

Fig. 5

A. DNA Binding



B. DNA Entry

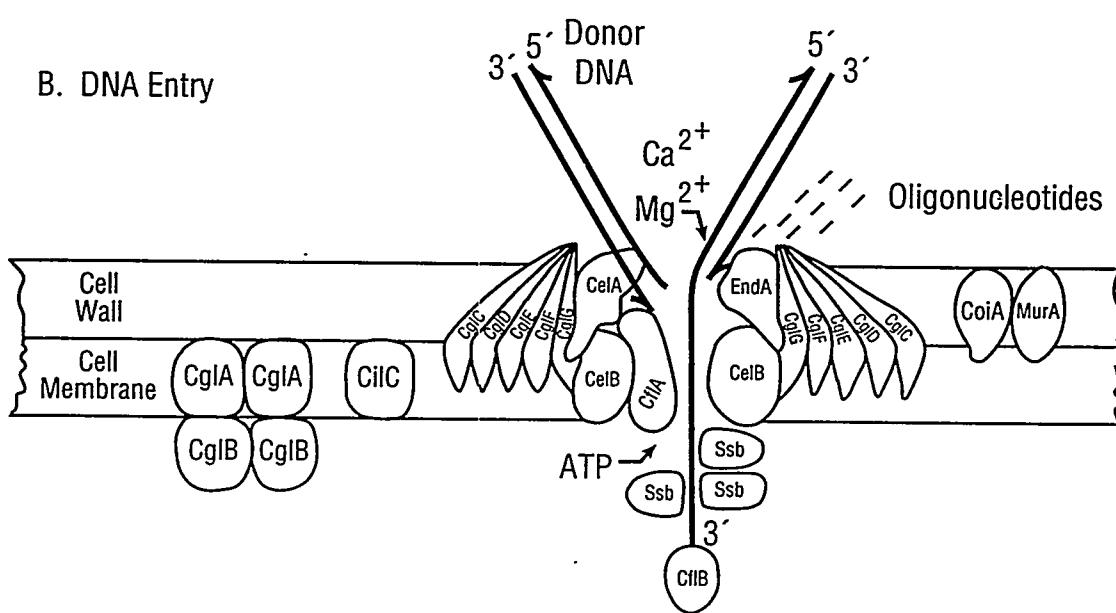


Fig. 6