BNL-71843-2003-BC

The Pneumococcus

Editor : E. Tuomanen

Associate Editors : B. Spratt, T. Mitchell, D. Morrison

To be published by ASM Press, Washington, DC

Chapter 9

Transformation

Sanford A. Lacks*

Biology Department

Brookhaven National Laboratory

Upton, NY 11973

Phone: 631-344-3369

Fax: 631-344-3407

E-mail: lacks@bnl.gov

Introduction

Transformation, which alters the genetic makeup of an individual, is a concept that intrigues the human imagination. In *Streptococcus pneumoniae* such transformation was first demonstrated. Perhaps our fascination with genetics derived from our ancestors observing their own progeny, with its retention and assortment of parental traits, but such interest must have been accelerated after the dawn of agriculture. It was in pea plants that Gregor Mendel in the late 1800s examined inherited traits and found them to be determined by physical elements, or genes, passed from parents to progeny. In our day, the material basis of these genetic determinants was revealed to be DNA by the lowly bacteria, in particular, the pneumococcus. For this species, transformation by free DNA is a sexual process that enables cells to sport new combinations of genes and traits.

Genetic transformation of the type found in *S. pneumoniae* occurs naturally in many species of bacteria (70), but, initially only a few other transformable species were found, namely, *Haemophilus influenzae*, *Neisseria meningitides*, *Neisseria gonorrheae*, and *Bacillus subtilis* (96). Natural transformation, which requires a set of genes evolved for the purpose, contrasts with artificial transformation, which is accomplished by shocking cells either electrically, as in electroporation, or by ionic and temperature shifts. Although such artificial treatments can introduce very small amounts of DNA into virtually any type of cell, the amounts introduced by natural transformation are a million-fold greater, and *S. pneumoniae* can take up as much as 10% of its cellular DNA content (40).

History

Despite its small size and relative simplicity, the pneumococcus, as the major causative agent of pneumonia, has been a scourge of humanity, particularly before the advent of antibiotics. In the early 1900s, pneumococci were the object of study in many laboratories worldwide, including those of Fred Griffith at the Public Health Ministry in London and of Oswald Avery at the Rockefeller Institute in New York. At the time, numerous serological types had been identified in *S. pneumoniae*, and the variation in capsule responsible for the serological differences was the focus of study in Griffith's laboratory.

Discovery of transformation. In 1928 Griffith reported that heat-killed encapsulated pneumococci could transfer the ability to make a capsule and, hence, to infect mice, when injected together with live,

unencapsulated (hence, nonpathogenic) pneumococci (31). Griffith termed the phenomenon *transformation*. He conjectured that a seed of capsular polysaccharide was perhaps transferred from the heat-killed bacteria, but he also wondered whether an enzymatically active protein might be the agent transferred. No one at the time supposed that bacteria contained genes, let alone that DNA was the genetic material.

DNA the transforming principle. Soon after Griffith's discovery, Oswald Avery at the Rockefeller Institute in New York took up the problem. Subsequent developments, such as a procedure to transform bacterial cells *in vitro* as opposed to in mice (17), and the extraction in soluble form of the active principle from heat-killed cells, enabled its further resolution (2). However, proof that the transforming principle was DNA awaited the landmark paper of 1944 (3). That work had enormous impact in demonstrating that the genetic material consisted of DNA.

Quantitation of transformation. Although capsular transformation was effective in proving that bacteria have genes and that genes are composed of DNA, it depended on screening for smooth (encapsulated) and rough (unencapsulated) colonies as opposed to selective growth of transformed cells. For precise quantitation in bacterial genetics, selectable markers are preferable because they allow counting of small numbers of transformants in a largely untransformed population. Rollin Hotchkiss, who continued the work of Avery's group at Rockefeller, obtained drug-resistance markers and devised methods for the quantitative measurement of transformation frequencies. The value of quantitative measurements for understanding molecular mechanisms of transformation, regulation of recipient cell competence, processes of DNA degradation, and genetic mapping by transformation is apparent from the results of Hotchkiss and those influenced by his laboratory. In 1957, using a streptomycin-resistance marker, he showed that transformants increased linearly with added marker DNA, until a saturation level was obtained, thereby demonstrating a discrete number of DNA uptake sites on the recipient cells (41). Competence for DNA uptake, measured as transformation frequency, was found to vary systematically during the culture growth cycle (40). Later it was shown that competence depended on accumulation of an extracellular polypeptide (34,109).

DNA structure. The intense interest in DNA that followed Avery's discovery led ultimately to the structural model for DNA proposed by Watson and Crick (114), which was based on the crystallographic data of Franklin and Gosling (27) and the chemical data of Chargaff (12). In this model two

complementary strands of DNA are helically wound around each other and attached by hydrogen bonds between complementary DNA bases. This structure was soon supported by additional studies of pneumococcal transformation. Using quantitative measurements of transforming activity, Marmur and Lane demonstrated strand separation on thermal denaturation of the native, double-stranded DNA and the renaturation of its transforming activity by annealing the separated strands at submelting temperatures (73). Thus, there was back-and-forth interplay between results from the transformation of *S. pneumoniae* and our knowledge of the role and structure of DNA.

Our current understanding of the mechanism of transformation and the genetics of *S. pneumoniae* has depended on a variety of experimental approaches: tracing the fate of isotopically labeled DNA, analysis of genetic recombination frequencies, isolation and characterization of transformation-defective and other mutants, DNA cloning and sequencing, and identification and use of the competence-inducing peptide to characterize the regulatory aspects of transformation.

Molecular fate of DNA in transformation

Eclipse and the entry nuclease. By labeling DNA with radioactive P atoms, it was shown that transforming DNA is physically incorporated into the cells of *S. pneumoniae* (24,67). The donor DNA must be in a native, double-stranded form; although single-stranded DNA can be taken up to a slight extent, its ability to transform is < 0.1% of native DNA (5,78). A curious and important finding, however, was the eclipse of donor markers immediately after DNA uptake; that is, DNA extracted from newly transformed cells was itself devoid of transforming activity when tested on other cells (25). Examination of the molecular state of the newly introduced DNA showed it to be present as single strands (44). An amount of donor DNA equivalent to the entered strands was released as oligonucleotides outside the cell (48). The eclipse in transforming activity is thus explained by the conversion of DNA upon entry to single strands, which are unable to efficiently enter tester cells. A membrane-located nuclease, EndA, was implicated in entry of the DNA strands (51-53), as shown in Fig. 1. Mutants in the *endA* gene are reduced in transformation to ~ 0.1% of the wild type (52,95). Terminal labeling of donor DNA has shown that the 3'-end of the incoming strand enters first (77).

Bound DNA. Instead of taking DNA into the cells, mutants lacking EndA bind large amounts of DNA to the outside of the cell (51). This DNA is irreversibly bound; it can be removed by treatment with DNase, but not by washing. Such externally bound DNA is always found to have undergone at least one single-strand break (47,49). Although, the DNA might be initially bound reversibly, it quickly undergoes strand breaks, perhaps at the sites of DNA entry, which may bind the DNA to a surface protein and render it irreversibly bound. These breaks limit the size of DNA strand segments entering the cell to a weight-average of ~5 kb. When EndA is present, it initially acts on the strand opposite the break made on DNA binding to give a double-strand break (49), as shown in Fig. 1. If the internal strand segments are homologous to the chromosome, they are rapidly integrated, largely intact, into the recipient chromosome (26,44). DNA lacking homology is taken up equally well into the cell, but it fails to be integrated and is eventually degraded (50).

Modes of genetic transformation

Chromosomal transformation. The pneumococcal chromosome corresponds to a circular genome containing 3.2 million base pairs (39,106). Spontaneous and chemically induced mutations in many genes have been obtained; they correspond to various single-site base changes and deletions and insertions of all sizes (45). Some chemical agents preferentially produce specific base changes (45,60). Ultraviolet light is not mutagenic in *S. pneumoniae*, unlike other bacteria (28). With current technology, designated mutations can be tailor-made *in vitro*, using synthetic oligonucleotides and the polymerase chain reaction, and introduced into the bacteria by transformation. Any selectable mutation or its wild-type allele constitutes a marker that can be analyzed by transformation. Classical transformation was concerned with changes in the recipient cell chromosome. For this to occur, single-stranded DNA must synapse with the double-stranded chromosomal DNA (45), as shown in Fig. 2B. As in most recombination systems, the protein RecA is required (74,87).

For chromosomal transformation, frequencies increase linearly with increasing DNA concentration, which indicates that entry of a single DNA fragment suffices (41). Markers located nearby on the chromosome will exhibit linkage, that is, show a co-transformation frequency greater than expected for two separate entry events. Recombination frequencies increase linearly with the genomic

distance between the markers. Fine structure analysis of recombination at two genetic loci—*mal* and *ami*—revealed a frequency of recombination of 0.02% per nucleotide (22,45) This analysis allowed a linear mapping of these two loci, which was confirmed by analysis of overlapping deletions and, subsequently, by DNA sequencing (15,60). The physical basis of recombination derives from double-strand breaks during DNA extraction, single-strand breaks on binding of DNA, and possible breaks or exchanges during strand integration (54).

Plasmid transfer and transformation. Chromosomal transformation is only one of several types of genetic transforming events; additional types relate to plasmids and other circular donor DNA, as shown in Fig. 2B. Although plasmids are found infrequently in natural strains of *S. pneumoniae*, they can be introduced from other streptococcal species. Plasmid transformation occurs in a cell containing a plasmid when it is treated with plasmid DNA containing a genetic marker. This is a process akin to chromosomal transformation, and it also has a linear dependence on DNA concentration. It is distinct from plasmid transfer, which follows introduction of monomeric plasmid DNA into a cell lacking the plasmid. Because a single entry event can provide only a linear strand, two entry events are required in this case to give a complementing pair of strands that can circularize into a plasmid capable of replication, and the frequency of establishment thus depends on the square of the DNA concentration (99).

Cloning in S. pneumoniae and chromosomal facilitation of plasmid transfer. Recombinant plasmids containing chromosomal DNA can be made and propagated in *S. pneumoniae*; the vector pMV158 and its derivatives have been particularly useful in this regard (103). Such cloning of pneumococcal DNA in *S. pneumoniae*, itself, has been helpful because many genes, such as *malM*, *hexA*, and *endA*, could not be cloned in systems using *E. coli*, where they exhibited toxic effects (102). Interestingly, transfer of a recombinant plasmid only requires a single entry event. Establishment of the recombinant plasmid apparently results from a circular synapsis of the linear plasmid strand with the chromosome, which allows replication of an intact complementary strand (Fig. 2C). This interaction, which elevates the frequency of plasmid establishment, is called chromosomal facilitation of plasmid transfer (69). During circular synapsis markers may be introduced from the chromosome into the cloned segment in the plasmid. This may or may not be desirable but can be easily monitored in the product.

Circular integration. Circular synapsis is also the basis for circular integration into the chromosome of nonreplicating circular DNA, which could be artifically constructed *de novo* or a recombinant plasmid incapable of replication in *S. pneumoniae* (Fig. 2D,E). Such integration, also called additive transformation, has been very useful for introducing marked mutations into genes to analyze their function and facilitate their cloning (56). Although the recombination event generally depends on homology, illegitimate recombination, at points lacking extensive homology, may occur more frequently than in linear chromosomal transformation and may produce chromosomal deletions adjacent to the point of insertion (110). Circular integration also allows the ectopic insertion of a pneumococcal or foreign gene at a place in the genome where it is not normally found (72). This can be accomplished by ligating a chromosomal DNA segment at the desired location to DNA containing either a pneumococcal or foreign gene and circularizing the product prior to its use as donor DNA (Fig. 2D).

Conjugative transposons. Although drug resistance genes are not usually found on plasmids in *S. pneumoniae*, they are frequently located on conjugative transposons, which are large chromosomal elements ranging from 15 to 60 kb that contain mobilization factors for their self-transmission to other cells (111). As inserts in the chromosomal DNA, these elements can also be transferred by the transformation mechanism.

Competence and its regulation

Initially, investigators found it difficult to reproduceably transform *S. pneumoniae*, because variations in strains used, growth media, and timing of DNA treatment affected results. Thus, it became evident that cells of *S. pneumoniae* are not always competent to take up DNA. Early observation of temporal variation of competence during a culture growth cycle (40) and of competence-stimulating activity of cell-free extracts (85,109), suggested that competence for transformation was under regulatory control. An important finding was that a specific set of proteins is induced during the development of competence (79). Extension of these studies coupled with analysis of transformation-negative mutants has identified a two-tiered regulatory mechanism for controlling competence.

Quorum sensing. A quorum-sensing mechanism constitutes the first tier of regulation, as illustrated in Fig. 3. It involves the products of five genes contained in two separate operons. ComC, the product of the first gene in one operon, is a polypeptide containing 41 amino acid residues and a GG motif past which cleavage occurs to give a carboxyl-terminal 17-mer oligopeptide (Fig. 4A) that is excreted from the cell (34). Similar leader sequences with GG motifs are commonly found in propeptide bacteriocins destined for excretion from the bacterial cell (34). The products of the other operon, ComA and ComB, which are, respectively, a transmembrane protein and an ATP-binding cytosolic protein of the ABC (ATP-Binding Cassette) transporter family, are responsible for processing and excreting the 17-mer (34,42). At sufficiently high concentrations in the medium, (which are nevertheless low in molar terms, that is, ~ 10 nM) this competence-stimulating polypeptide (CSP), can induce high levels of competence in an incompetent culture (34).

Roughly half of all pneumococcal strains isolated from patients encode the CSP1 sequence shown in Fig. 4A; nearly all the rest encode a distinct but similar sequence, CSP2, that differs in eight amino acid residues (93). Corresponding to these two alternative forms of CSP, the sequences of the corresponding *comD* products differ in 13 of the first hundred amino-terminal residues in the two classes (116), as shown in Fig. 4B. This supports the idea that residues in the amino-terminal half of ComD specifically recognize CSP (35). From its sequence, ComD is an integral membrane protein with six transmembrane helices in its amino-terminal half. Its carboxyl-teminal half corresponds to a histidine kinase. Together ComD and ComE function as a two-component signal-transduction system frequently

found in bacteria (13,88). In the present case, CSP acting externally on the ComD receptor presumably triggers the transfer of a phosphate residue from histidine in ComD to ComE, the response regulator, which would alter ComE and enable it to activate the genes under its control. CSP acts to increase transcription from the comAB and comCDE operons, themselves, thereby resulting in its autocatalytic production and a surge of competence.

The consensus binding site in DNA for ComE action, which was deduced by comparing the promoter regions of the comAB and comCDE operons (115), is shown in Fig. 4C. Both operons contain an extended –10 promoter motif typical for *S. pneumoniae* (98), which alone can afford a low basal level of transcription (59). This assures the production and release of CSP at a low rate so that it will accumulate in the medium. When external CSP reaches a sufficient concentration, it acts back on ComD, which in turn activates ComE. The latter acts as an enhancer to greatly increase transcription from both operons, which lack a typical –35 promoter motif. Instead, binding of activated ComE to the directly repeated sequences in the DNA elicits a high level of transcription. When the population of bacteria reaches a suitable density to allow interaction between released DNA and cells, that is, when a quorum is present, this system provides a surge of competence to facilitate transformation. Enabling genetic exchange during late exponential growth allows the selection of new genotypes at a moment late enough to recognize increasing stress on the population, yet early enough to synthesize a more adaptive response.

Alternative sigma factor. A most important third operon affected by ComE (Fig. 4C) contains the single gene, *comX*. In *S. pneumoniae* there are two copies this operon, each of which contains a single *comX* gene that encodes an alternative sigma factor (66). During competence, this sigma factor, ComX, replaces SigA in RNA polymerase. ComX, therefore, serves as the link between the two tiers of competence regulation. It does not recognize the usual promoter of *S. pneumoniae*, but rather a different –10 sequence upstream from the mRNA start site, TACGAATA. This sequence, called the "combox," is found upstream of operons containing nearly all other genes required for competence for DNA uptake and other functions associated with transformation (10), as shown in Fig. 5. The sequential expression of the two tiers causes the "early" competence gene transcripts of the first tier, which depend on the response regulator ComE, to reach a peak ~7.5 min after addition of CSP to a noncompetent culture,

whereas the "late" competence gene transcripts of the second tier, which depend on ComX, reach a peak ~12.5 min after such addition (1,91,97).

In addition to early and late classes, in which mRNA levels are generally increased ~50-fold and decline after reaching their peak values, induction of competence affects two additional classes of genes (91,97). Genes in the "delayed" response class are expressed more gradually, and mRNA increases only 2- to 10-fold. None of these genes appear to be involved in transformation. Rather, they encode chaperonins, heat shock proteins, and proteases implicated in protein folding or elimination of misfolded proteins. Thus, they may represent a response to stressful conditions. Genes in the "repressed" class, although numerous (over 60 have been reported), are transiently reduced only 2- to 4-fold in expression; they encode ribosomal proteins and various enzymes and transport systems (91,97). Regulation of the delayed and repressed genes is not affected by elimination of ComX (91), and the mechanisms of their control are unknown.

Approximately a dozen additional operons that show transcriptional behavior similar to early competence genes have been reported (6,91,97). Only half of them are preceded by a repeat sequence corresponding to the ComE binding site. With the exception of one gene, *comW* (originally called CPIP912; ref. 6), preceded by the binding site (Fig. 4C), none of these early genes are essential for transformation (91). The function of ComW is unknown, but it could be the factor additional to ComX that enhances synthesis of late competence gene products (71,91). Why the other operons are part of the first tier competence regulon and whether they have any function related to transformation has not been ascertained.

Combox regulon. Sixteen late competence genes that are required for DNA uptake or subsequent functions in transformation have been identified (for a review, see ref. 58). These genes, to be discussed in detail in the next section, are found in eight operons, each of which is preceded by a combox (Fig. 5). A search for additional combox sequences in *S. pneumoniae* revealed eight new operons expressed as late competence genes (90). However, they encode proteins that show no obvious relevance to transformation. Similarly, searches of transcripts elicited by CSP revealed eight more operons with kinetics similar to late competence genes, but again deletion of the genes or the associated combox did not affect transformation (6,91,97). Only four of the additional late operons

induced by CSP lacked a combox. Further investigation will be needed to see whether these putative members of the second tier competence regulon do function in transformation or whether they have roles in another cellular process, perhaps one requiring similar quorum sensing.

Strain variation and constitutive competence. In the laboratory strain Rx, in which competence regulation has been most intensively studied, competence appears as a sharp peak of ~30 min duration. The upsurge of competence is clearly due to the autocatalytic nature of the first tier regulation and its positive induction of competence genes. However, it is not known why synthesis abruptly ceases. It has been hypothesized that one of the late competence genes may block competence development, but no evidence for such a function has yet been adduced. Another laboratory strain, R6, exhibits a broad peak of competence lasting over several generations of bacterial growth. It is possible that the latter strain is deficient in the blocking function, but it is not known which expression pattern is typical of wild strains. It is conceivable that such strain variations occur naturally with a short-lived, high spike of competence being desirable under some circumstances and a more moderate but longer-lived level of competence under others. In the former case, rapid synthesis of competence proteins may interfere with normal protein synthesis, and this was in fact observed (79). However, it has long been known that a derivative of R6, the *trt* mutant, is constitutively competent (48). The *trt* mutation is located in *comD*; it corresponds to a change in ComD of Asp_{299} to Asn, which apparently alters the histidine kinase so that it can phosphorylate ComE even without activation of ComD by CSP (61). The *trt* mutant was selected on the basis of its transformability in the presence of trypsin, a proteolytic enzyme that degrades CSP. Additional mutations have recently been found in both *comD* (75) and *comE* (20) which similarly render their products constitutively active.

Molecular mechanism of DNA uptake

In addition to the constitutive *endA* gene discussed above, four inducible operons containing late competence genes, *celAB*, *cflAB*, *cglABCDEFG*, and *cilC*, are uniquely involved in DNA uptake. The genes in these operons are homologous to those genes found responsible for DNA uptake in another transformable Gram-positive bacterium, *Bacillus subtilis*, where their protein products have been characterized with respect to primary sequence, proteolytic processing, membrane location, and other

properties (reviewed in ref. 19). Mutations in the pneumococcal genes affect transformation similarly to those in *B. subtilis* (7,10,89), so the role of the encoded proteins can be assumed to correspond in the two species (Table 1).

The *cgl* operon of *S. pneumoniae* contains seven genes, which are adjacent to each other or slightly overlapping, as in *B. subtilis*. Together with *cilC*, the *cglABCDEFG* operon encodes a set of proteins similar to those responsible for the extrusion of type IV pilins and other proteins through the outer membrane and cell wall of Gram-negative bacteria (38). Mutations in any of these genes prevent binding and entry of DNA and reduce transformability to <0.01% of normal. Based on their homology, the *cglA* and *cglB* products are, 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 most bacterial species, including *B. subtilis*, the amino acid residues KGFT precede the hydrophobic segment, and homologues of CilC cleave the leader peptide between the G and F residues (19) during its transport. However, in *S. pneumoniae*, a similar sequence, KAFT, precedes the segment in CglC, CglD and CglF (58), and cleavage presumably occurs between A and F. The product of *cilC* and its homologues is a peptidase-methyltransferase that processes extruded proteins by cleaving the polypeptide at the F residue and methylating the new N-terminus (105).

How the extruded *cgl* products act in DNA uptake is conjectural, for their homologues in *B. subtilis* have not shown binding or other action on DNA, but possible functions can be envisioned (14). By analogy to their Gram-negative counterparts, they may form an appendage outside the cell membrane and passing through the cell wall. This structure may act both as a pore in the peptidoglycan layer of the cell wall through which external DNA can pass and as a scaffold on which other proteins that bind and process DNA for entry are arranged. It is also possible that the Cgl protein *complex* can bind external DNA, perhaps reversibly.

The *celA* and *celB* genes are essential for DNA uptake, and mutations in them reduce transformability to <0.01% (10,89). CelA is a membrane protein of ~20 kDa with a long stretch of hydrophobic residues at its N-terminus. Mutations of its *B. subtilis* homologue prevent binding of DNA to the cell, and the protein *in vitro* binds tightly to DNA (94). Binding is to the C-terminal portion of the protein, which is external to the cell. In *S. pneumoniae, celA* mutations greatly reduce DNA binding and completely eliminate strand degradation by EndA (7). The CelA protein thus may be responsible for nicking DNA,

thereby binding it irreversibly, or possibly binding it to another protein, such as CflB. CelA may also recruit EndA to degrade the complementary strand.

CelB is an 80-kDa protein with multiple hydrophobic stretches corresponding to transmembrane segments. The protein is not required for donor DNA binding or degradation, but it is necessary for DNA entry (7,19). It is likely that this protein forms a channel in the membrane for passage of single-stranded DNA.

The two genes in the *cfl* operon of *S. pneumoniae* correspond to the two genes of the *comF* operon of *B. subtilis* that are essential for transformation. CflA (and ComFA) mutants are reduced in transformability to 0.1 % of normal (19,65). They fail to take up DNA, but they do show binding and degradation (7). CflA and ComFA, both 50-kDa proteins, show sequence similarity to ATP-dependent DNA helicases (19). Although they lack hydrophobic regions, ComFA was associated with the cytoplasmic face of the cell membrane (19). With respect to function, it has been suggested that the protein acts as a helicase to help separate the DNA strands and propel one strand into the cell. However, it may have other, unknown functions. CflB is homologous to ComFC of *B. subtilis*. Mutants in the latter gene reduce transformability only to 10% of normal, and the role of the gene product in transformation has not been elucidated.

A provisional model of how EndA and the various components encoded by late competence genes may act to bring DNA into the cell and process it for chromosomal integration or plasmid establishment is shown in Fig. 5.

Release of donor DNA. For transformation to occur under natural conditions DNA must be released from donor cells as well as taken up by recipient cells. Early studies indicated that transforming DNA was released from cells in culture (84). Furthermore, an ability to form spheroplasts, which is indicative of cell wall fragility, was correlated with competence (53,101). Treatment of cells with crude preparations of CSP increased fragility (101); treatment with trypsin decreased fragility, except in the constitutively competent *trt* mutant (53). This fragility depended on the *lytA* gene, since it was blocked by the *cwl-1* mutation in that gene (53). Recently, it was shown that treatment of noncompetent cells with pure CSP can release up to 20% of a normally internal enzyme and also significant amounts of chromosomal DNA (104). It still is not known whether this external material leaks out of all competent cells or whether it represents the complete lysis of a small proportion of them. Also, it has not been

definitively shown that the *lytA* product, and not other cell wall lytic enzymes encoded by *lytB* or *lytC*, is uniquely required. In this connection, it is of interest that a late competence operon contains the *lytA* gene. This operon contains *exp10* (or *cinA*), a gene that encodes a membrane protein of unknown function, *recA*, *dinF*, a gene homologous to a DNA damage-inducible gene of *B. subtilis*, and *lytA* (74,87). Although *lytA* and *recA* have their own promoters and their products are constitutively produced, they are nevertheless upregulated ~5-fold during competence. Thus, competence regulation based on quorum sensing may enable both release and uptake of DNA by pneumococcal cells under conditions of cell density when they are most likely to achieve genetic exchange and to benefit from it.

Fate of DNA within the cell and recombination

DNA bound to protein. What happens when donor strands enter the cell? First, they are covered by a single-stranded DNA binding protein, which may protect them from degradation by nucleases (81). *S. pneumoniae* harbors two *ssb* loci that encode proteins homologous to the single-stranded DNA binding protein of *E. coli*. One, *ssbA*, is expressed constitutively and presumably functions in DNA replication, and the other, *ssbB*, is a late competence gene. The additional production is important because mutants defective in *ssbB* (=*cilA*) are reduced to 3% of normal transformability (10). Whether such binding also assists the uptake process is unknown.

Recombination. Binding of SsbB may facilitate recombination, as these proteins do in other systems. Essential to recombination, however, is the *recA* gene. As mentioned above, *recA* expression is increased during competence. The single-stranded donor DNA may synapse with chromosomal DNA to initially form a three-stranded structure (45), as has been proposed for RecA-mediated recombination, in general (9). Subsequently, the recipient DNA segment corresponding to the donor is replaced and eliminated to give hybrid chromosomal DNA containing donor segments of varied size, with a weight average of ~ 5 kb (26). Details of the mechanism of this recombination are not known. Finally, the donor segment, which initially has free ends (55), is ligated into the chromosome. Another competence-induced gene that may play a role in chromosomal recombination is *cilB*, which is homologous to the *H. influenzae* gene, *dprA*, in which mutations reduce chromosomal transformation to 0.01 %, with no effect on plasmid transfer (43). This behavior is reminiscent of mutations in *S. pneumoniae* called *recP*, in which chromosomal transformation, but not

plasmid transfer was blocked (80). The genes corresponding to the *recP* and *recQ* mutations were not fully characterized. It is possible that *recP* corresponds to *cilB*, and *recQ*, in which mutations blocked both chromosomal transformation and plasmid establishment, may correspond to *recA*. Action of the *cilB* gene product may be confined to chromosomal integration, inasmuch as plasmid establishment is a simpler process, not requiring displacement of existing chromosomal DNA.

Whether other competence-induced genes affecting transformation act in the recombination process is not known, but several late competence genes have not been characterized with respect to either DNA uptake or recombination. They include the aforementioned *cflB*, and, also, *coiA* (89). Mutations in these genes reduce transformability only 10- to 100-fold. 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 is conceivable that these enzymes play a role in cell wall remodeling during the development of competence, but singly mutating the three downstream genes does not affect transformation (91).

Gene redundancy and functional analysis. Many important cellular systems contain redundant genes or alternative pathways to assure retention of the system function despite the occurrence of spontaneous mutations. For example, there are duplicate *comX* genes in *S. pneumoniae* (66). Genes such as *cflB*, *exp10* and *coiA*, in which mutations reduce transformability minimally, or not at all, may represent redundant or alternative paths. *A fortiori*, the numerous newly reported genes that are induced during the development of competence (91,97) might correspond to this class of redundant or alternative genes. These genes were not identified in screening for transformation-defective mutations, presumably because single mutations have little impact. To determine whether such genes are required for transformation, double mutants in pairs of the genes must be tested.

Restriction and transformation. The case of *dpnA* is unusual in that it is regulated as a late competence gene (59), 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 (11), as shown in Fig. 6B, below. *S. pneumoniae* contains one or the

other of two complementary restriction systems, DpnI and DpnII, which appear to be designed to prevent infection by bacterial viruses that inject double stranded DNA. Chromosomal transformation, because it is mediated by DNA converted to single strands on entry, is not affected by restriction, as the nucleases act only on double-stranded DNA (112). In the case of plasmid establishment, two donor strands anneal to form the replicon. The facts that the pneumococcal restriction systems are designed to avoid degradation of transforming DNA, and that provision is made (i.e. the *dpnA* gene) even to counter restriction in plasmid transfer, indicate that genetic transformation plays an important role in survival of *S. pneumoniae* as a species.

Mismatch repair. Subsequent to uptake and synapsis with chromosomal DNA, but prior to its ligation to recipient DNA, the donor DNA is subject to mismatch repair. DNA mismatch repair is a mechanism, universally present in cells, that was first discovered in *S. pneumoniae* by analysis of transformation frequencies given by different mutational markers at the *malA* (45) and *amiA* (22) loci. DNA mismatches correspond to noncomplementary base pairings or deletions (or additions) in one strand to produce an aberration in the double-helical DNA structure. In chromosomal transformation, a single-stranded segment of donor DNA replaces the homologous segment of host DNA, thereby producing a mismatch wherever a genetic marker is present. It was found that different types of mismatch result in characteristic transformation efficiencies ranging from 0.05 to 1.0 (22,45). A repair system that differentially recognizes the various types of mismatch and eliminates the donor contribution to the mismatched heteroduplex accounts for the differences in integration efficiency (21,45,46). Products of the unlinked *hexA* and *hexB* genes are necessary for this repair; mutations in either gene give high integration efficiencies for all markers (4,16,46). Because the Hex system recognizes and corrects a variety of DNA mismatches it is called a *generalized* mismatch repair system (55). *S. pneumoniae* also contains a specialized repair system that recognizes only the mismatched sequence 5'-ATTAAT/TAAGTA- 5' and converts the A to C (86). For a more detailed description of mismatch repair in *S. pneumoniae* with references, see ref. 57.

Mutations in the *hex* genes also have a mutator effect (108). Thus, the system normally acts also to correct errors in newly replicated DNA. This is probably the main purpose of the repair system, and it explains why the *hex* genes are constitutively expressed. Their action after transformation may simply be

incidental to the process and have no survival value. Sets of genes homologous to the *hex* genes of *S. pneumoniae* are found in nearly all living species. The generalized mismatch repair system that they encode reduces spontaneous mutation rates a thousand-fold. As a consequence, the homologous system in human cells prevents various types of cancer (23).

Comparison with other Gram+ and Gram- Species

Homologues to the genes responsible for competence and DNA uptake in *S. pneumoniae* are found in many other species of bacteria. These species can be considered in four groups depending on their degree of divergence from *S. pneumoniae*. In group I are closely related streptococcal species, such as *S. mitis* and *S. gordonii*, among others, that contain homologues of both early and late genes, as well as *endA*, and they are all transformable. In this group, the arrangement of the *comCDE* genes is identical to that in *S. pneumoniae*, but the CSPs and their cognate receptor portions in the histidine kinase are different (36), as indicated in Fig. 4A for *S. mitis*. Group II comprises more distantly related streptococcal species, such as *S. mutans* and *S. pyogenes*, which also contain early and late genes, with the late genes under combox control. *S. mutans* is transformable (68), but there are no reports of transformation of *S. pyogenes*. Homologues of *comX* are found in groups I and II. *S. mutans* contains close homologues of *endA* and *comCDE*, except the arrangement of the latter genes differs from *S. pneumoniae* and group I (68), as shown in Fig. 4D. All the transformable species in groups I and II are similarly regulated by quorum sensing and depend on similar gene products for DNA uptake and processing. However, it is conceivable that some of these products and others induced by the same regulatory system may have additional physiological roles, including response to stress.

Bacterial species outside of the genus *Streptococcus*, such as *Lactococcus lactis* and *Enterococcus faecalis*, which are, nevertheless, closely related to *Streptococcus*, comprise group III. In these species homologues of early genes cannot be clearly discerned, but the late genes required for DNA uptake and processing are present; however, they are not under combox control. For example, in *L. lactis*, the *cglABCDEFG* operon is present (58), but it is apparently transcribed from a typical streptococcal SigA promoter (98), TTGAat..n=11..TnatnTATAtT. Although, no transformation of *L. lactis* has been reported, if it does occur, it must be regulated by a system very different from that in *S.*

pneumoniae. Alternatively, these "late" genes may have an entirely different physiological function. One possibility is that they serve to respond to physiological stresses, such as a dearth of nutrients, that may result from crowded conditions or other causes. Some genes induced during the development of competence have no role in transformation but may act to ameliorate the effects of stresses such as heat shock (91,97). It is possible either that the genes required for transformation acquired stress-related functions during evolution or that transformation co-opted structures previously evolved for such functions.

Distantly related bacteria that are naturally competent for transformation constitute group IV. There are many transformable species (70), but Gram-positive *Bacillus subtilis* and Gram-negative *Haemophilus influenzae* have been the most thoroughly studied. Interestingly, they both have regulatory mechanisms for induction of competence under specified conditions, but these mechanisms are different from the quorum sensing system of *S. pneumoniae* and from each other. However, for the most part, the late competence gene products are homologous throughout the bacterial kingdom (reviewed in refs. 58). In fact, the pioneering work with *B. subtilis* (reviewed in ref. 19) facilitated the characterization of these products in other species. Table 1 lists the genes in *B. subtilis* and *H. influenzae* that correspond to the genes in *S. pneumoniae* believed to be part of the transformation mechanism.

Relationship of transformation to other regulons and to virulence

Mutational analysis of function. On a practical level, transformation has provided the means for determining the functions of many genes in *S. pneumoniae*. Stable disruption of a gene can be accomplished by chromosomal transformation with DNA from that gene into which a drug-resistance marker is inserted. Circular integration of a nonreplicating plasmid containing a fragment of the gene and a selectable marker can similarly produce sufficiently stable insertion mutations, although the original plasmid is released at a low frequency. Properties of cells bearing these mutations can then be studied. However, the insertion may have polar effects on downstream genes in an operon. Introduction into the construct of a reporter gene can allow measurement of the target gene expression under various conditions. The possibilities of altering recipient cell genes are vastly increased by the use of oligonucleotide synthesis and the polymerase chain reaction and with the availability of the genomic

sequence of *S pneumoniae* (refs. 39 and 106; access at www.ncbi.nlm.nih.gov/genomes/MICROBES/Complete.html).

Two-component signal transduction systems. The quorum-sensing system for eliciting competence is an example of a two-component signaling (TCS) system, in which a sensor histidine-kinase phosphorylates a response regulator protein that carries out the effect. In most TCS systems, the histidine-kinase protein contains an N-terminal transmembrane domain that allows it to respond to external signals, and the response regulator interacts with DNA to stimulate or inhibit transcription of one or more operons, which constitute a regulon (30). TCS systems are found in many bacteria; the histidine kinase domains and response regulator proteins are sufficiently conserved to enable their recognition from sequence alone. Also, on the basis of sequence, they can be divided into subgroups, even across species. *S. pneumoniae* contains 13 TCS systems and a single isolated response regulator (63,107), which fall into four subgroups, as indicated in Table 2. All of its sensor proteins appear to have transmembrane domains (63).

With respect to function, only ComDE has been well characterized. In the 13 TCS systems, only one gene, the response regulator of TCS02/492, is essential for viability (63,107). Another system, TCS13/486, on the basis of its sequence similarity to ComDE and the presence of an adjacent operon encoding both a putative ABC transporter system and a polypeptide with a GG cleavage site, was surmised to be a quorum-sensing system similar to ComDE (63,107). On the basis of homology to PhoPR of *B. subtilis* , TCS04/481 may be involved in phosphate regulation (83). CiaRH is an interesting TCS system that can affect both competence and virulence. Its signal for kinase action and the DNA binding sequence for its regulator are unknown, but from the experimentally determined binding sites for the regulator that affect transcription, it appears that at least 10 operons are in the regulon (76). Activation of the regulator increases transcription in seven operons; elimination of the regulator increases transcription in the other three. Competence falls into the latter group, being inhibited by the regulator, which apparently acts at a binding site upstream from *comCDE*. The *C306* mutation (Thr_{230} to Pro) in CiaH, which (like *trt* in ComD) may render the kinase constitutively active, inhibits transformability (76). This mutation also increases beta-lactam resistance slightly (76). It was suggested that the CiaRH

regulon mediates the transition from exponential to stationary phase (76). Interestingly, the ciaRH operon shows "delayed" hyperexpression after CSP treatment (91).

TCS systems and virulence. The effect on bacterial virulence of mutations in components of the various TCS was examined with two different infection models. A "bacteremia" model, in which pneumococci were injected into the peritoneum of mice, revealed no attenuation in any of the mutants tested (63). However, a "pneumonia" model, in which pneumococci were inoculated intranasally in mice, and the bacterial load in lung tissue was analyzed after 48 hours, showed considerable attenuation for some mutants (107), as indicated in Table 2. Apparently, the pneumonia model, which requires nasal carriage and adhesion and colonization in the lung as well as growth and resistance to host defenses, is a more stringent and superior method for revealing virulence factors. For 9 of the 13 two-component systems and the isolated response regulator, deletion of the response regulator reduced the yield of bacteria in the lung by from 1 to 6 logs. The most attenuated strain was mutated in TCS04/481, which is presumably involved in phosphate metabolism. Inasmuch as the response regulator of TCS02/492 was essential for growth, it could not be deleted, but overexpression of the kinase decreased virulence in a bacteremia model (113). Deletion of *ciaRH* gave 5 logs attenuation. A gene, *htrA*, which appears to be normally upregulated by the CiaRH TCS, has been implicated in nasopharyngeal carriage of *S. pneumoniae* (100). This gene may encode a serine protease.

Virulence genes. Two large-scale investigations to identify virulence genes in *S. pneumoniae* were carried out using insertional mutagenesis of pathogenic strains. Mutants were produced by circular integration of a nonreplicating plasmid containing a drug-resistance marker. Both the pioneering study (92) and the later one (64) relied on signature tagging of individual mutants with a short, variable DNA sequence (37). This method allowed identification of bacteria that could not survive in a mixed infection with as many as one hundred different mutant strains pooled into a single inoculum. Screening of large numbers of mutants was thereby greatly facilitated. Both studies examined survival after intranasal inoculation by analyzing output from the lung, and one study also used a bacteremia model and analyzed output from the spleen (64). In the latter study, screening of 1786 mutants gave 186 that were attenuated, divided approximately equally among those defective only in the pneumonia model, those defective in the bacteremia model, and those defective in both. Among 56 mutants selected for further

study, 46 different genes were implicated and they were characterized with respect to genomic location and putative function. Similarly, in the earlier study (92), screening of 1250 mutants gave 200 attenuated strains mutated in 126 different genes, which were further characterized.

Results of the two independent studies of virulence are in surprisingly good agreement, even in some unexpected ways. Perhaps it is not surprising that as many as 10% of non-essential genes in *S. pneumoniae* are required for virulence. However, quite a few of these genes are implicated in rather mundane functions of transport, biosynthesis, and other metabolism. For example, of 46 genes implicated in one study, 11 were ABC-type transporters. Important metabolic systems in *S. pneumoniae* may be redundant. As an example, the bacterium has seven different transport systems for uptake of glutamine (39), its likely source of nitrogen. Both studies identified the glutamine transport gene *glnQ* as a virulence factor. Under particular conditions, as in infection, only one of the redundant functional forms may be operative, and hence essential for growth only under those conditions. Such regulatory effects might explain the contribution to virulence of so many genes not required for growth in culture. Both studies also identified several genes previously known to affect virulence, among them those encoding proteases (*iga*, *clpC*) and adhesins (*pavA*).

In both of the above studies, it was concluded that pathogenicity islands, that is, clusters of virulence genes, do not occur in *S. pneumoniae*. However, this conclusion may be questioned, because neither study obtained mutations in the capsular genes, which are known to be required for virulence (8). And the capsular gene cluster, consisting of ~ 15,000 base pairs, constitutes a hefty target. In fact, a later study using *in vitro* transposon mutagenesis, followed by insertion into the chromosome via transformation, revealed a possible small pathogenicity island of seven genes bracketed by insertion elements (33). The cluster encoded three surface proteins, three "sortases" for attaching the proteins to the cell wall, and a regulatory protein. Mutations in the regulatory gene or one sortase caused attenuation of virulence, apparently by interfering with nasal carriage (33). Otherwise, the results of the transposon study did not differ appreciably from the earlier studies. More complete characterization of the numerous "virulence" genes and their effects *in vivo* will be necessary to understand their role.

Quorums vs. crowds. Several observations suggest a connection between competence for transformation and virulence, but the evidence is weak. Although the CiaRH TCS can affect both

processes, virulence depends on action of the response regulator (107), whereas competence depends on inhibition of its action (76). Mutations in *comB* (64) and *comD* (6) were reported to reduce virulence but only gave tenfold attenuation. It is possible that many genes in the competence regulons function in stress response rather than transformation. The "quorum-sensing" mechanism that facilitates transformation can also serve as a "crowd-sensing" mechanism to facilitate growth in adverse conditions, which may include infected tissues. For this reason, genes affecting transformation, stress-response, and virulence may overlap in function.

Cassette mechanisms

Population genetics. Transformation plays an important role in the population genetics of *S. pneumoniae*. First and foremost, it allows the spread of mutant forms of genes among populations. It can produce mosaic genes, as shown for drug resistance (18,32). Second, it can introduce novel genes from genetically distant sources either in plasmids or in chromosomal fragments. A surprising number of pneumococcal genes appear to have originated in Gram-negative bacteria (39). For example, genes SP1467 and SP1468 exhibit 76% and 88% identity with genes from *Haemophilus influenzae* (106), but in this case transformation could have been in either direction. Third, transformation facilitates the maintenance of diverse populations of *S. pneumoniae* with respect to certain traits. Several systems of biological importance exist in two or more states in populations of pneumococci. This population diversity must have survival value for the species. In the following three cases, a similar mechanism of allelic substitution by a multigenic cassette is responsible for changes of state.

Restriction systems. Although the capsular genes, which constitute such a cassette, have been investigated since the time of Griffith (31), insight into the mechanism by which transformation transfers these cassettes came from analysis of the Dpn restriction systems of *S. pneumoniae* (62). Cells of *S. pneumoniae* contain either the DpnI or DpnII system (Fig. 6A). The DpnI endonuclease recognizes and cleaves the methylated DNA sequence 5'GmATC; cells that produce it contain unmodified DNA. The DpnII system is complementary to the DpnI system in that it recognizes the unmethylated sequence 5'GATC. Unlike other restriction systems, it encodes *two* methyltransferases, DpnM and DpnA, which methylate double- and single-stranded DNA, respectively. The DpnII endonuclease cleaves unmethylated,

double-stranded DNA. Both endonucleases cleave only double-stranded DNA, so incoming single strands of susceptible transforming DNA are not degraded. Furthermore, DpnA can methylate such strands, thereby protecting reconstituted plasmids (Fig. 6B), as described above. Thus, these systems are designed to block bacterial virus infection but not to interfere with genetic transformation between cells with different restriction systems. The dual systems may prevent viral epidemics from wiping out an entire pneumococcal population, inasmuch as an initial infecting agent would be either methylated or not (Fig. 6C). The genes encoding the two restriction systems are clustered at the same locus in the chromosome. Since the genes bordering them are identical, one system can replace the other by chromosomal transformation (62).

Capsular genes. The polysaccharide capsule that surrounds the pneumococcal cell is essential for its virulence. At least 90 different capsule types exist (8). The capsule protects the pneumococcus from destruction by phagocytes. Immunity to pneumococcal infection is directed mainly to the capsule, so the multiplicity of capsule types is clearly beneficial to the pathogen. Genetic investigation of several capsular types revealed that the genes for their biosynthesis are present at the same genetic locus, between *plpA* and *dexB* (106). Thus, capsular gene clusters of different specificity can be transferred between cells by chromosomal transformation just like the restriction gene cassettes. As many as 2% of pneumococcal genes appear to be truncated (39); these defective genes may be remnants of past transformations. In the case of the capsular locus, many such fragments at the capsular cassette borders correspond to genes of other capsular types (8). Such surrounding capsular gene remnants could facilitate transformation to those specific capsule types.

Competence systems. With respect to the regulation of competence, *S. pneumoniae* has two distinct, but closely related, systems for quorum sensing. As described above, the competence-stimulating peptides of the two systems, CSP1 and CSP2, differ by several amino acids, and the cognate receptor portion of the transmembrane histidine kinase, ComD, also differ. Approximately half of wild strains have the CSP1 system, and the other half have the CSP2 system. With mixed populations in nature, the result would be to remove part of the pneumococcal population from quorum sensing. Such inhibition of genetic exchange, however, should be deleterious rather than beneficial for the species. A possible resolution of this quandary is deducible from the prior examples of cassette mechanisms. In both of the above cases, diversity protected the species from noxious agents, bacteriophages and

immune antibodies, respectively. It is conceivable that CSP1 or CSP2 can prompt an antibody response from host organisms that would inactivate it and block the development of competence with respect to only cells of that type in the population. The persistence of two states of CSP in the species argues for the importance of competence induction for pathogenesis. Although other stress-response functions could be helpful, the ability for genetic exchange may be the most important stress-related response.

Methionyl-tRNA synthetases. All pneumococcal strains contain the *metS1* gene that encodes a methionyl-tRNA synthetase, which is an essential enzyme. Approximately half of strains isolated from patients contain a second such enzyme encoded by *metS2* (29). The second enzyme is resistant to certain synthetic inhibitors of the first enzyme. It was suggested that similar inhibitors occur as natural antibiotics (29). If so, this may be another case of population diversity conferring a selective advantage to the species.

Despite the considerable information presented in this chapter and additional information that space limitation precluded from consideration, many questions remain concerning the regulation of competence, DNA uptake, genetic recombination, and the ramifications of these processes, especially for virulence of the pneumococcus.

Acknowledgements

Donald Morrison generously provided information prior to publication. This chapter was written at Brookhaven National Laboratory during a Guest Appointment in the Biology Department, which operates under the auspices of the U. S. Department of Energy Office of Biological and Environmental Research.

References

1. **Alloing, G., B. Martin, C. Granadel, and J. P. Claverys**. 1998. Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. Mol Microbiol. **29:**75-83.
2. **Alloway, J. L.** 1931. The transformation in vitro of R pneumococci into S forms of different specific types by the use of filtered pneumococcus extracts. J. Exp. Med. **55:**91-99.
3. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 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. J. Exp. Med. **89:**137-158.
4. **Balganesh, T. S., and S. A. Lacks.** 1985. Heteroduplex DNA mismatch repair system of *Streptococcus pneumoniae*: Cloning and expression of the *hexA* gene. J. Bacteriol. **162:**979-984.
5. **Barany, F., and J. D. Boeke.** 1983. Genetic transformation of *Streptococcus pneumoniae* by DNA cloned into the single-stranded bacteriophage f1. J. Bacteriol. **153:**200-210.
6. **Bartilson, M., A. Marra, J. Christine, J. S. Asundi, W. P. Schneider, and A. E. Hromockyj.** 2001. Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide. Mol. Microbiol. **39:**126-35.
7. **Berge, M., M. Moscoso, M. Prudhomme, B. Martin, and J. P. Claverys.** 2002. Uptake of transforming DNA in Gram-positive bacteria: a view from *Streptococcus pneumoniae*. Mol Microbiol. **45:**411-421.
8. **Caimano, M. J., G. G. Hardy, and J. Yother.** 1998. Capsule genetics in *Streptococcus pneumoniae* and a possible role for transposition in the generation of the type 3 locus. Microb. Drug Resist. **4:**11-23.
9. **Camerini-Otero, R., and P. Hsieh.** 1995. Homologous recombination proteins in prokaryotes and eukaryotes. Annu. Rev. Genetics **29:**509-552.
10. **Campbell, E. A., S. Y. Choi, and H. R. Masure.** 1998. A competence regulon in *Streptococcus pneumoniae* revealed by genome analysis. Mol. Microbiol. **27:** 929-939.
11. **Cerritelli, S., S. S. Springhorn, and S. A. Lacks.** 1989. DpnA, a methylase for single-strand DNA in the *Dpn*II restriction system, and its biological function. Proc. Natl. Acad. Sci. USA **86:** 9223-9227.
12. **Chargaff, E.** 1951. Structure and function of nucleic acids as cell constituents. Fed. Proc. **10:**654-659.

13. **Cheng, Q., E. A. Campbell, A. M. Naughton, S. Johnson, and H. R. Masure.** 1997. The *com* locus controls genetic transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **23**:683-692.

14. **Chung, Y. S., F. Breidt, and D. Dubnau.** 1998. Cell surface localization and processing of the ComG proteins required for DNA binding during transformation of *Bacillus subtilis*. Mol. Microbiol. **29**:905-913.

15. **Claverys, J. P., V. Mejean, A. M. Gasc, and A. M. Sicard.** 1983. Mismatch repair in *Streptococcus pneumoniae*: relationship between base mismatches and transformation efficiencies. Proc. Natl. Acad. Sci. USA **80**:5956-5960.

16. **Claverys, J. P., H. Prats, H. Vasseghi, and M. Gherardi.** 1984. Identification of *Streptococcus pneumoniae* mismatch repair genes by an additive transformation approach. Mol. Gen. Genet. **196**:91-96.

17. **Dawson, M. H., and R. H. P. Sia.** 1931. In vitro transformation of pneumococcal types. I. A technique for inducing transformation of pneumococcal types in vitro. J. Exp.Med. **54:** 681-699.

18. **Dowson, C. G., T. J. Coffey, C. Kell, and R. A. Whiley.** 1993. Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. pneumoniae*. Mol. Microbiol. **9**:635-643.

19. **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.

20. **Echenique, J. R., S. Chapuy-Regaud, and M.-C. Trombe.** 2000. Competence regulation by oxygen in *Streptococcus pneumoniae*: involvement of *ciaRH* and *comCDE*. Mol. Microbiol. **36**:688-696.

21. **Ephrussi-Taylor, H., and T. C. Gray.** 1966. Genetic studies of recombining DNA in pneumococcal transformation. J. Gen. Physiol. **49**:part 2,211-231.

22. **Ephrussi-Taylor, H., A. M. Sicard, and R. Kamen**. 1965. Genetic recombination in DNA-induced transformation of pneumococcus. I. The problem of relative efficiency of transforming factors. Genetics **51**:455-475.

23. **Fishel, R., M. K. Lescoe, M. R. S. Rao, N. G. Copeland, N. A. Jenkins, J. Garber, M. Kane, and R. Kolodner.** 1993. The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. Cell **75**:1027-1038.

24. **Fox, M. S.** 1957. Deoxyribonucleic acid incorporation by transformed bacteria. Biochim. Biophys. Acta **26:**83-85.

25. **Fox, M. S.** 1960. Fate of transforming deoxyribonucleate following fixation by transforming bacteria. II. Nature **187:**1004-1006.

26. **Fox, M. S., and M. K. Allen.** 1964. On the mechanism of deoxyribonucleate integration in pneumococcal transformation. Proc. Natl. Acad. Sci. USA, **52:**412-419.

27. **Franklin, R. and R. Gosling.** 1953. Molecular configuration in sodium thymonucleate. Nature **171:**740-741.

28. **Gasc, A. M., N. Sicard, J. P. Claverys, and A. M. Sicard.** 1980. Lack of SOS repair in *Streptococcus pneumoniae*. Mutat. Res. **70:**157-165.

29. **Gentry, D. R., K. A. Ingraham, M. J. Stanhope, S. Rittenhouse, R. L. Jarvest, P. J. O'Hanlon, J. R. Brown, and D. J. Holmes.** 2003. Variable sensitivity to bacterial methionyl-tRNA synthetase inhibitors reveals subpopulations of *Streptococcus pneumoniae* with two distinct methionyl-tRNA synthetase genes. Antimicrob. Agents Chemother. **47:**1784-1789.

30. **Grebe, T. W., and J. B. Stock.** 1999. The histidine protein kinase superfamily. Adv. Microbial Physiol. **41:**139-227.

31. **Griffith, F.** 1928. The significance of pneumococcal types. J. Hyg. **27:**113-159.

32. **Hakenbeck, R., N. Balmelle, B. Weber, C. Gardes, W. Keck, and A. de Saizieu.** 2001. Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. Infect. Immun. **69:**2477-2486.

33. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389-1406.

34. **Håverstein, L. S., G. Coomaraswamy, and D. A Morrison.** 1995. An unmodified heptadecapeptide induces competence for genetic transformation in *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **92:**11140-11144.

35. **Håverstein, L. S., P. Gaustad, I. F. Nes, and D. A Morrison.** 1996. Identification of the streptococcal competence-pheromone receptor. Mol. Microbiol. **21:**863-869.

36. **Håvarstein, L. S., R. Hakenbeck, and P. Gaustad.** 1997. Natural competence in the genus *Streptococcus*: evidence that streptococci can change pherotype by interspecies recombinational exchanges. J. Bacteriol. **179:**6589-6594.

37. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400-403.

38. **Hobbs, M., and J. S. Mattick.** 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol. Microbiol. **10:**233-243.

39. **Hoskins, J., W. E. Alborn Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass.** 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709-5717.

40. **Hotchkiss, R. D.** 1954. Cyclical behavior in pneumococcal growth and transformability occasioned by environmental changes. Proc. Natl. Acad. Sci. USA **40:**49-55.

41. **Hotchkiss, R. D.** 1957. Criteria for the quantitative genetic transformation of bacteria, p. 321-335. *In* W. D. McElroy and B. Glass (ed.), The chemical basis of heredity. Johns Hopkins Press, Baltimore.

42. **Hui, F. M., and D. A. Morrison.** 1991. Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. J. Bacteriol. **173:**372-381.

43. **Karudapuram, S., X. Zhao, and G. J. Barcak.** 1995. DNA sequence and characterization of *Haemophilus influenzae dprA*$^+$, a gene required for chromosomal but not plasmid DNA transformation. J. Bacteriol. **177:**3235-3240.

44. **Lacks, S.** 1962. Molecular fate of DNA in genetic transformation of pneumococcus J. Mol. Biol. **5:**119-131.

45. **Lacks, S.** 1966. Integration efficiency and genetic recombination in pneumococcal transformation. Genetics **53:**207-235.

46. **Lacks, S.** 1970. Mutants of *Diplococcus pneumoniae* that lack deoxyribonucleases and other activities possibly pertinent to genetic transformation. J. Bacteriol. **101:**373-383.

47. **Lacks, S.** 1979. Uptake of circular deoxyribonucleic acid and mechanism of deoxyribonucleic acid transport in genetic transformation of *Streptococcus pneumoniae*. J. Bacteriol. **138:**404-409.

48. **Lacks, S., and B. Greenberg**. 1973. Competence for deoxyribonucleic acid uptake and deoxyribonuclease action external to cells in the genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **114:**152-163.

49. **Lacks, S., and B. Greenberg.** 1976. Single-strand breakage on binding of DNA to cells in the genetic transformation of *Diplococcus pneumoniae*. J. Mol. Biol. **101:**255-275.

50. **Lacks, S., B. Greenberg, and K. Carlson.** 1967. Fate of donor DNA in pneumococcal transformation. J. Mol. Biol. **29:**327-347.

51. **Lacks, S., B. Greenberg, and M. Neuberger.** 1974. Role of a deoxyribonuclease in the genetic transformation of *Diplococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **71:**2305-2309.

52. **Lacks, S., B. Greenberg, and M. Neuberger.** 1975. Identification of a deoxyribonuclease implicated in genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **123:** 222-232.

53. **Lacks, S., and M. Neuberger.** 1975. Membrane location of a deoxyribonuclease implicated in the genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **124:**1321-1329.

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

55. **Lacks, S. A.** 1989. Generalized DNA mismatch repair--its molecular basis in *Streptococcus pneumoniae* and other organisms, p. 325-339. *In* L. O. Butler, C. Harwood, and B. E. B. Moseley (ed.), Genetic transformation and expression. Intercept, Andover, Hants, England.

56. **Lacks, S. A.** 1997. Cloning and expression of pneumococcal genes in *Streptococcus pneumoniae*. Microb. Drug Resist. **3:**327-337.

57. **Lacks, S. A.** 1998. DNA repair and mutagenesis in *Streptococcus pneumoniae*, p. 263-286. *In* J. A. Nickoloff and M. F. Hoekstra (ed.), DNA damage and repair, vol. 1: DNA repair in prokaryotes and lower eukaryotes, Humana Press, Totowa, NJ.

58. **Lacks, S. A.** 1999. DNA uptake by transformable bacteria, p.138-168. *In* J. K. Broome-Smith, S. Baumberg, C. J. Stirling, and F. B. Ward, (ed.), Transport of molecules across microbial membranes. Cambridge University Press: Cambridge, UK.

59. **Lacks, S. A., S. Ayalew, A. G. de la Campa, and B. Greenberg.** 2000. Regulation of competence for genetic transformation in *Streptococcus pneumoniae*: expression of *dpnA*, a late competence gene encoding a DNA methyltransferase of the *Dpn*II restriction system. Mol. Microbiol. **35:**1089-1098.

60. **Lacks, S. A., J. J. Dunn, and B. Greenberg.** 1982. Identification of base mismatches recognized by the heteroduplex-DNA-repair system of *Streptococcus pneumoniae*. Cell **31:**327-336.

61. **Lacks, S. A., and B. Greenberg.** 2001. Constitutive competence for genetic transformation in *Streptococcus pneumoniae* caused by mutation of a transmembrane histidine kinase. Mol. Microbiol. **42:**1035-1045.

62. **Lacks, S. A., B. M. Mannarelli, S. S. Springhorn, and B. Greenberg.** 1986. Genetic basis of the complementary DpnI and DpnII restriction systems of S. pneumoniae: An intercellular cassette mechanism. Cell **46:** 993-1000.

63. **Lange, R., C. Wagner, A. de Saizieu, N. Flint, J. Molnos, M. Stieger, P. Caspers, M. Kamber, W. Keck, and K. E. Amrein.** 1999. Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. **Gene 237:**223-234.

64. **Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison, and D. W. Holden.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:**555-571.

65. **Lee, M. S., B. A. Dougherty, A. C. Madeo, and D. A. Morrison.** 1999. Construction and analysis of a library for random insertional mutagenesis in *Streptococcus pneumoniae*: use for recovery of mutants defective in genetic transformation and for identification of essential genes. Appl. Environ. Microbiol. **65:**1883-1890.

66. **Lee, M. S., and D. A. Morrison.** 1999. Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation. J. Bacteriol. **181**:5004-5016

67. **Lerman, R. S., and L. J. Tolmach.** 1957. Genetic transformation. I. Cellular incorporation of DNA accompanying transformation in pneumococcus. Biochim. Biophys. Acta **28**:68-82.

68. **Li, Y. H., P. C. Lau, J. H. Lee, R. P. Ellen, and D. G. Cvitkovitch.** 2001. Natural genetic transformation of *Streptococcus mutans* growing in biofilms. J. Bacteriol. **183**:897-908.

69. **Lopez, P., M. Espinosa, D. L. Stassi, and S. A. Lacks.** 1982. Facilitation of plasmid transfer in *Streptococcus pneumoniae* by chromosomal homology. J. Bacteriol. **150**:692-701.

70. **Lorenz, M. G., and W. Wackernagel.** 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. **58**:563-602.

71. **Luo, P., H. Li, and D. A. Morrison.** 2003. ComX is a unique link between multiple quorum sensing outputs and competence in *Streptococcus pneumoniae*. Mol. Microbiol., in press.

72. **Mannarelli, B. M., and S. A. Lacks.** 1984. Ectopic integration of chromosomal genes in *Streptococcus pneumoniae*. J. Bacteriol. **160**:867-873.

73. **Marmur, J., and D. Lane.** 1960. Strand separation and specific recombination in deoxyribonucleic acids: Biological studies. Proc. Natl. Acad. Sci. USA **46**:453-461.

74. **Martin, B., P. Garcia, M. P. Castanie, and J. P. Claverys.** 1995. The *recA* gene of *Streptococcus pneumoniae* is part of a competence-induced operon and controls lysogenic induction. Mol. Microbiol. **15**:367-379.

75. **Martin, B., M. Prudhomme, G. Alloing, C. Granadel, and J. P. Claverys.** 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **38**:867-878.

76. **Mascher, T., D. Zahner, M. Merai, N. Balmelle, A. B. de Saizieu, and R. Hakenbeck.** 2003. The *Streptococcus pneumoniae cia* regulon: CiaR target sites and transcription profile analysis. J. Bacteriol. **185**:60-70.

77. **Mejean, V., and J. P. Claverys.** 1988. Polarity of DNA entry in transformation of *Streptococcus pneumoniae*. Mol. Gen. Genet. **213**:444-448.

78. **Miao, R., and W. R. Guild.** 1970. Competent *Diplococcus pneumoniae* accept both single-and double-stranded deoxyribonucleic acid. J. Bacteriol. **101:**361-364.

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

80. **Morrison, D. A., S. A. Lacks, W. G. Guild, and J. M. Hageman.** 1983. Isolation and characterization of three new classes of transformation-deficient mutants of *Streptococcus pneumoniae* that are defective in DNA transport and genetic recombination. J. Bacteriol. **156:**281-290.

81. **Morrison, D. A., and B. Mannarelli.** 1979. Transformation in pneumococcus: nuclease resistance of deoxyribonucleic acid in the eclipse complex. J. Bacteriol. **140:**655-665.

82. **Mortier-Barriere, I., A. de Saizieu, J. P. Claverys, and B. Martin.** 1998. Competence-specific induction of *recA* is required for full recombination proficiency during transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **27:**159-170.

83. **Novak, R., A. Cauwels, E. Charpentier, and E. Tuomanen.** 1999. Identification of a *Streptococcus pneumoniae* gene locus encoding proteins of an ABC phosphate transporter and a two-component regulatory system. J. Bacteriol. **181:**1126-1133.

84. **Ottolenghi, E., and R. D. Hotchkiss.** 1962. Release of genetic transforming agent from pneumococcal cultures during growth and disintegration. J. Exp. Med. **116:**491-519.

85. **Pakula, R., and W. Walczak.** 1963. On the nature of competence of transformable streptococci. J. Gen. Microbiol. **31:**125-133.

86. **Pasta, F., and M. A. Sicard.** 1994. Hyperrecombination in pneumococcus: A/G to C.G repair and requirement for DNA polymerase I. Mutat. Res. **315:**113-122.

87. **Pearce, B. J., A. M. Naughton, E. A. Campbell, and H. R. Masure.** 1995. The *rec* locus, a competence-induced operon in *Streptococcus pneumoniae*. J. Bacteriol. **177:**86-93.

88. **Pestova, E. V., L. S. Håverstein, and D. A. Morrison.** 1996. Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. Mol. Microbiol. **21:**853-862.

89. **Pestova, E. V., and D. A. Morrison.** 1998. Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a *lacZ* reporter insertion vector. J. Bacteriol. **180:**2701-2710.

90. **Peterson, S., R. T. Cline, H. Tettelin, V. Sharov, and D. A. Morrison.** 2000. Gene expression analysis of the *Streptococcus pneumoniae* competence regulons by use of DNA microarrays. J. Bacteriol. **182**:6192-202.

91. **Peterson, S. N., R. Cline, B. V. Desai, E. Snesrud, P. Luo, J. Walling, H. Li, M. Mintz, G. Tsegaye, P. Burr, Y. Do, S. Ahn, J. Gilbert, R. Fleischmann, and D. A. Morrison.** 2003. Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. Mol. Microbiol., submitted.

92. **Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari, and D. Simon.** 1998. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:**5620-5629.

93. **Pozzi, G., L. Masala, F. Iannelli, R. Manganelli, L. S. Håvarstein, L. Picolli, D. Simon, and D. A. Morrison.** 1966. Competence for genetic transformation in encapsulated strains of *Streptococcus pneumoniae*: two allelic variants of the peptide pheromone. J. Bacteriol. **178:**6087-6090.

94. **Provvedi, R., and D. Dubnau.** 1999. ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. Mol. Microbiol. **31:**271-280.

95. **Puyet, A., B. Greenberg, and S. A. Lacks.** 1990. Genetic and structural characterization of EndA, a membrane-bound nuclease required for transformation of *Streptococcus pneumoniae*. J. Mol. Biol. **213:**727-738.

96. **Ravin, A. W.** 1960. The genetics of transformation. Advanc. Genet. **10:**61-163.

97. **Rimini, R., B. Jansson, G. Feger, T. C. Roberts, M. de Francesco, A. Gozzi, F. Faggioni, E. Domenici, D. M. Wallace, N. Frandsen, and A. Polissi.** 2000. Global analysis of transcription kinetics during competence development in *Streptococcus pneumoniae* using high density DNA arrays. Mol. Microbiol. **36:**1279-1292.

98. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended -10 promoter alone directs transcription of the *Dpn*II operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144-155.

99. **Saunders, C. W., and W. R. Guild.** 1980. Monomer plasmid DNA transforms *Streptococcus pneumoniae*. Mol. Gen. Genet. **181:**57-62.

100. **Sebert M. E., L. M. Palmer, M. Rosenberg, and J. N. Weiser.** 2002. Microarray-based identification of *htrA*, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization. Infect. Immun. **70:**4059-4067.

101. **Seto, H., and A. Tomasz.** 1974. Protoplast formation and leakage of intramembrane cell components: induction by the competence activator substance of pneumococci. J. Bacteriol. **121:**344-353.

102. **Stassi, D. L., and S. A. Lacks.** 1982. Effect of strong promoters on the cloning in *Escherichia coli* of DNA fragments from *Streptococcus pneumoniae*. Gene **18:** 319-328.

103. **Stassi, D. L., P. Lopez, M. Espinosa, and S. A. Lacks.** 1981. Cloning of chromosomal genes in *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **78:**7028-7032.

104. **Steinmoen, H., E. Knutsen, and L. S. Håvarstein.** 2002. Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population. Proc. Natl. Acad. Sci. USA **99:**7681-7686.

105. **Strom, M. S., D. N. Nunn, and S. Lory.** 1993. A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. Proc. Natl. Acad. Sci. USA **90:**2404-2408.

106. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson,** L. A. **Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson,** E. K. **Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293**:498-506.

107. **Throup, J. P., K. K. Koretke, A. P. Bryant,** K. **A. Ingraham, A. F. Chalker, Y. Ge, A. Marra, N. G. Wallis, J. R. Brown, D. J. Holmes, M. Rosenberg, and M. K. Burnham.** 2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. Mol. Microbiol. **35:**566-576.

108. **Tiraby, G., and M. S. Fox.** 1973. Marker discrimination in transformation and mutation of pneumococcus Proc. Natl. Acad. Sci. USA **70:**3541-3545.

109. **Tomasz, A., and R. D. Hotchkiss.** 1964. Regulation of the transformability of pneumococcal cultures by macromolecular cell products. Proc. Natl. Acad. Sci. USA **51:**480-487

110. **Vasseghi, H., J. P. Claverys, and A. M. Sicard.** 1981. Mechanism of integrating foreign DNA during transformation in *Streptococcus pneumoniae*, p.137-154. *In* M. Polsinelli and G. Mazza (ed.) Transformation 1980. Cotswold Press, Oxford.

111. **Vijayakumar, M. N., S. D. Priebe, and W. R. Guild.** 1986. Structure of a conjugative element in *Streptococcus pneumoniae*. J. Bacteriol. **166:**978-984.

112. **Vovis, G. F., and S. Lacks.** 1977. Complementary action of restriction enzymes Endo R.*Dpn*I and Endo R.*Dpn*II on bacteriophage fl DNA. J. Mol.Biol. **115:**525-538.

113. **Wagner, C., A. de Saizieu, H. J. Schonfeld, M. Kamber, R. Lange, C. J. Thompson, and M. G. Page.** 2002. Genetic analysis and functional characterization of the *Streptococcus pneumoniae vic* operon. Infect. Immun. **70:**6121-6128.

114. **Watson, J. D., and F. H. C. Crick.** 1953. Molecular structure of nucleic acids. Nature **171:**737-738.

115. **Ween, O., P. Gaustad, and L. S. Håvarstein.** 1999. Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. Mol. Microbiol. **33:**817-827.

116. **Whatmore, A. M., V. A. Barcus, and C. G. Dowson.** 1999. Genetic diversity of the streptococcal competence (*com*) gene locus. J. Bacteriol. **181:**3144-3154.

TABLE 1. Genes of *S. pneumoniae* implicated in transformation

Gene [a]	Regu-lation [b]	Homo-logue [c]	Product	Function	R6 genome	TIGR4 genome
endA	C	-	Membrane nuclease	Degrade donor strand	spr1779	SP1964
recA	C,L	Eco *recA*	Rec protein	DNA recombination	spr1757	SP1940
comA	E	-	ABC transporter	Process/export CSP	spr0043	SP0042
comB	E	-	ABC transporter	Process/export CSP	spr0044	SP0043
comC	E	-	CSP precursor	Quorum sensing	spr2043	SP2237
comD	E	-	Histidine kinase	Detect CSP signal	spr2042	SP2236
comE	E	-	Response regulator	Early gene transcripts	spr2041	SP2235
comW	E	-	81-aa protein	Unknown	spr0020	SP0018
comX1	E	-	Sigma factor	Late gene transcripts	spr0013	SP0014
comX2	E	-	Sigma factor	Late gene transcripts	spr1819	SP2006
celA/cilE	L	Bsu *comEA*	DNA binding protein	Bind donor DNA	spr0856	SP0954
celB	L	Bsu *comEB*	Transmembrane protein	DNA entry pore?	spr0857	SP0955
cflA	L	Bsu *comFA*	Putative helicase	Assist DNA entry?	spr2013	SP2208
cflB	L	Bsu *comFB*	221-aa protein	DNA pilot protein?	spr2012	SP2207
cglA/cilD	L	Bsu *comGA*	ABC transporter	Export CglC,D,E,F,G	spr1864	SP2053
cglB	L	Bsu *comGB*	ABC transporter	Export CglC,D,E,F,G	spr1863	SP2052
cglC	L	Bsu *comGC*	Membrane protein	Cell wall channel?	spr1862	SP2051
cglD	L	Bsu *comGD*	Membrane protein	Cell wall channel?	spr1861	SP2050
cglE	L	Bsu *comGE*	Membrane protein	Cell wall channel?	spr1860	SP2049
cglF	L	Bsu *comGF*	Membrane protein	Cell wall channel?	spr1859	SP2048
cglG	L	Bsu *comGG*	Membrane protein	Cell wall channel?	spr1858	SP2047
cilB/smf	L	Hin *dprA*	287-aa protein	Post-entry processing	spr1144	SP1266
cilC/pilD	L	Bsu *comC*	Peptidase-Mtase	Process CglC,D,F	spr1628	SP1808
coiA	L	-	318-aa protein	Unknown	spr0881	SP0978
dpnA	C,L	-	Methyltransferase	Protect donor DNA	- [d]	- [d]
exp10/cinA	L	-	Membrane protein	Unknown	spr1758	SP1941
lytA	C,L	-	Cell wall lysin	Release donor DNA?	spr1754	SP1937
ssbB/cilA	L	Eco *ssb*	Strand binding protein	Protect donor strands	spr1724	SP1908

[a] Alternative name indicated after slash mark.

[b] C, constitutive expression; E, early inducible expression; L, late inducible expression

[c] Species designations: Bsu, *B. subtilis*; Eco, *E. coli*; Hin, *H. influenzae*.

[d] Absent in strains R6 and TIGR4, which are DpnI strains.

TABLE 2. Two-component signal transduction systems in *S. pneumoniae*

Group[a]	R6 genome	TIGR4 genome	Ref. 62 Name	Ref. 107 Name	Com-ponent	Viru-lence[b]	Regulatory function
AgrA	spr2042	SP2236	ComDE	498	HK	ND	Quorum sensing (88)
/Agr	spr2041	SP2235			RR	ND	
	spr0464	SP527	TCS13	486	HK	ND	
	spr0463	SP526			RR	-4	
LuxR	spr1815	SP2001	TCS11	479	HK	ND	
/ Nar	spr1814	SP2000			RR	WT	
	spr0343	SP386	TCS03	474	HK	WT[c]	
	spr0344	SP387			RR	[c]	
AraC	spr0153	SP155	TCS07	539	HK	-4[c]	
/ Lyt	spr0154	SP156			RR	[c]	
	spr0579	SP662	TCS09	488	HK	-1[c]	
	spr0578	SP661			RR	[c]	
OmpR	spr1473	SP1632	TCS01	480	HK	-5[c]	
/Pho	spr1474	SP1633			RR	[c]	
	spr1106	SP1226	TCS02	492	HK	WT	
	spr1107	SP1227			RR	E	
	spr1894	SP2083	TCS04	481	HK	ND	Phosphate metabolism (83)
	spr1893	SP2082			RR	-6	
	spr1997	SP2192	TCS06	478	HK	-3[c]	
	spr1998	SP2193			RR	[c]	
	spr0077	SP0084	TCS08	484	HK	ND	
	spr0076	SP0083			RR	-3	
	spr0529	SP0604	TCS10	491	HK	ND	
	spr0528	SP0603			RR	-1	
	spr0708	SP0799	CiaRH	494	HK	-5[c]	Stationary phase? (76)
	spr0707	SP0798			RR	[c]	
	spr0336	SP0376	-	489	RR	-4	

[a] Group designations in ref. 62 /ref. 107.

[b] Virulence measured as $\log_{10}$ reduction in bacterial count in lung tissue 48 hr after intranasal instillation of mice (107), with mutant compared to wild-type strain 0100993. ND, not determined; E, essential for growth.

[c] Mutant deleted in both HK and RR genes.

Figure Legends

FIG. 1. Model for DNA uptake in transformation of *S. pneumoniae* . Double-stranded DNA is irreversibly bound to the cell surface and undergoes single-strand cleavage at random sites, possibly by action of a binding protein. A membrane-located nuclease, EndA, initiates entry of the bound strand by endonucleolytic cleavage of the complementary strand to give a double-strand break. Processive action of EndA 5' to 3' degrades the complementary strand to oligonucleotides, which remain outside the cell, while donor strands enter from their 3'-end (half-arrowhead). It is not known whether the strand enters without (a) or with (b) a pilot protein. The entering DNA is covered with a single-strand binding protein (c').

FIG. 2. A. Chromosomal transformation. Heavy line, donor DNA strand segment. Thin line, chromosomal DNA. M and m, marker difference between donor and recipient. For plasmid transformation, substitute resident plasmid for chromosomal DNA. 1. Linear synapsis. 2. Integration intermediate. 3. Covalent joining. B. Plasmid establishment. 1. Annealing of complementary strand fragments that entered separately. 2. Repair synthesis. 3. Completed replicon. C. Chromosomal facilitation of plasmid establishment. 1. Circular synapsis followed by repair synthesis and ligation to close the plasmid strand. 2. Synthesis of the complementary strand from the plasmid origin of replication. 3. Release of established plasmid. D. Ectopic integration of the *mal* marker in the vicinity of the *sul* locus. 1. Donor DNA consists of separately cloned *mal* and *sul* genes ligated together. 2. Circular synapsis of the donor strand fragment at the *sul* chromosomal locus; a gap is filled by repair synthesis. 3. A single-strand crossover integrates the donor strand into the chromosome. 4. Replication of the chromosome converts the integrated single-strand segment to a duplex form giving a *mal* segment inserted between duplicated *sul* segments. E. Mutagenesis of the *ami* gene by additive insertion of a nonreplicating plasmid. 1. Donor DNA consists of the *ami* gene joined to an *E. coli* plasmid containing an *erm* gene expressible in *S. pneumoniae*. 2. Circular synapsis of the donor strand at the *ami* chromosomal locus and repair synthesis. 3. A single-strand crossover integrates the donor strand into the chromosome. 4. Replication of the chromosome converts the integrated single-strand segment to a duplex form so that the *E. coli* plasmid segment is inserted between duplicated

ami segments, thereby producing an aminopterin-resistance mutation. Letters a-d and a'-d' in D and E designate parts of the *sul* and *ami* loci, respectively.

FIG. 3. Model of quorum sensing in the regulation of competence for transformation. Accumulated extracellular CSP signals ComD to phosphorylate ComE, which then enhances synthesis of CSP and ComX. ComX is needed to transcribe genes required for transformation. Relevant genes are shown at bottom. Open arrows point to gene products. Solid arrows show effects on promoters Operon control elements: black, SigA promoter; white, weak SigA promoter; horizontal hatch, binding site for ComE enhancer. Other symbols: P, protein phosphate; ComC', residual *comC* product after removal of CSP.

FIG. 4. Variation in competence regulatory components of streptococci. A. Competence-stimulating peptides. Species and strain are indicated. Vertical arrows indicate point of cleavage from precursor. Dots indicate identity to peptide sequence above. B. CSP-receptor regions of ComD in *S. pneumoniae* strains. Amino-terminal 96 residues are compared. C. Binding sites for ComE in *S. pneumoniae*. Consensus sequence from ref. 115. Upper case indicates correspondence to relatively invariant bases of the consensus. Numbers in brackets give distance between ComE-binding sequence repeats and between the second repeat and the extended –10 promoter site (98). D. Arrangement of genes *comCDE*. Boxes depicting genes point in the direction of transcription.

FIG. 5. Late competence genes and the construction and function of the DNA uptake apparatus in *S. pneumoniae*. DNA is depicted by heavy lines with half-arrowheads indicating the 3'-direction. The "translocasome" is a hypothetical structure extruding through the cell wall and formed by CglC-G proteins, which are exported by the CglA,B complex and processed by CilC. Other components of the translocasome are CelA, which binds DNA, EndA, which degrades one strand, CelB, which forms a membrane pore for entry of the other strand, CflA, which may unwind donor DNA. CoiA and CflB also may function in DNA uptake, possibly by nicking and attaching to DNA prior to entry. Calcium and magnesium ions are required for DNA uptake, with the latter needed by EndA. Upon entry, single strands are coated with Ssb. CilB, DpnA, and RecA act subsequent to DNA uptake Relevant genes are shown

at bottom. Open arrows point to gene products. Solid arrows show effects on promoters. Operon control elements: black, SigA promoter; horizontal hatch, ComE enhancer; crosshatch, ComX promoter. Other symbols: P, protein phosphate; m, methyl group to CglC-G on processing by CilC. Question marks indicate uncertain role in transformation.

FIG. 6. Restriction enzyme systems of *S. pneumoniae*. A. Restriction gene cassettes of *S. pneumoniae* and their products. Symbols: thin bar, *S. pneumoniae* chromosome; thick bar, Dpn cassette; open boxes, genes in the cassettes or in the adjacent chromosome, showing direction of transcription. B. Role of DpnA methylase in enabling unmethylated plasmid transfer into cells containing the DpnII restriction system. The degradative processing of DNA entering the cell by the transformation pathway, requires the reconstitution of a plasmid from complementary strands that separately enter the cell. In a host lacking the DpnA methyltransferase, unmethylated plasmid DNA, upon reconstitution to a double-stranded form, would be cleaved by the DpnII endonuclease. In a host containing DpnA, single strands are methylated upon entry, so that the reconstituted plasmid is protected from the DpnII endonuclease. C. Possible survival value of complementary restriction systems. I and II, Cells making DpnI and DpnII, respectively. Infection of a mixed population by a single viral particle would only destroy part of the population.

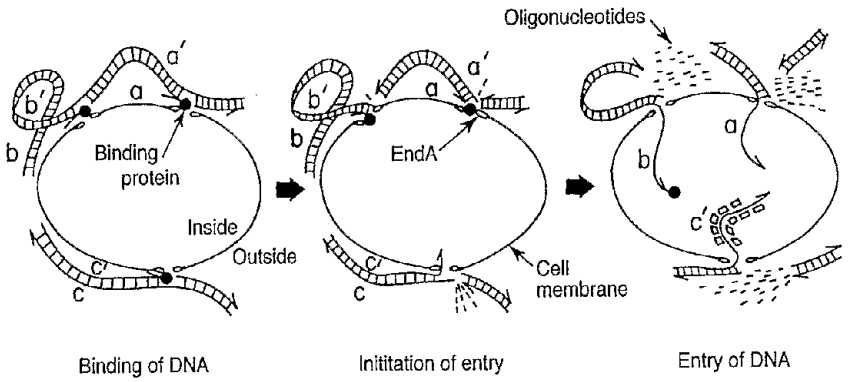

Oligonucleotides
a'
a
b'
b
Binding protein
Inside
Outside
c'
c
EndA
Cell membrane
Binding of DNA
Inititation of entry
Entry of DNA

Fig. 1

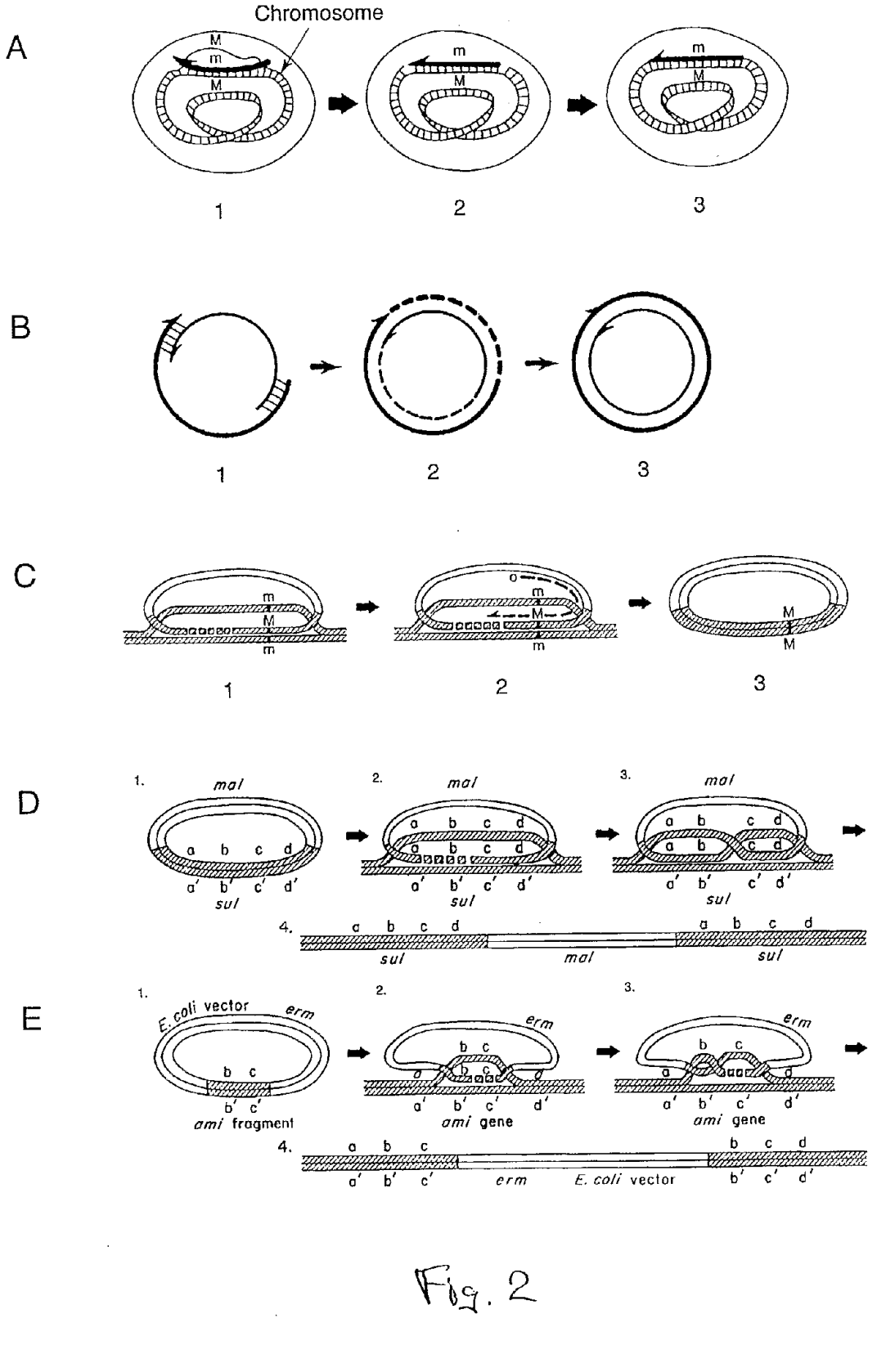

A
Chromosome
M
m
M
1
m
M
2
m
M
3
B
1
2
3
C
m
M
m
1
o
m
M
m
2
M
M
3
D
1.
mal
a b c d
a' b' c' d'
sul
2.
mal
a b c d
a b c d
a' b' c' d'
sul
3.
mal
a b c d
a b c d
a' b' c' d'
sul
4.
a b c d
a b c d
sul
mal
sul
E
1.
E. coli vector
erm
b c
b' c'
ami fragment
2.
erm
b c
a d
a' b' c' d'
ami gene
3.
erm
b c
a d
a' b' c' d'
ami gene
4.
a b c
b c d
a' b' c'
erm
E. coli vector
b' c' d'

Fig. 2

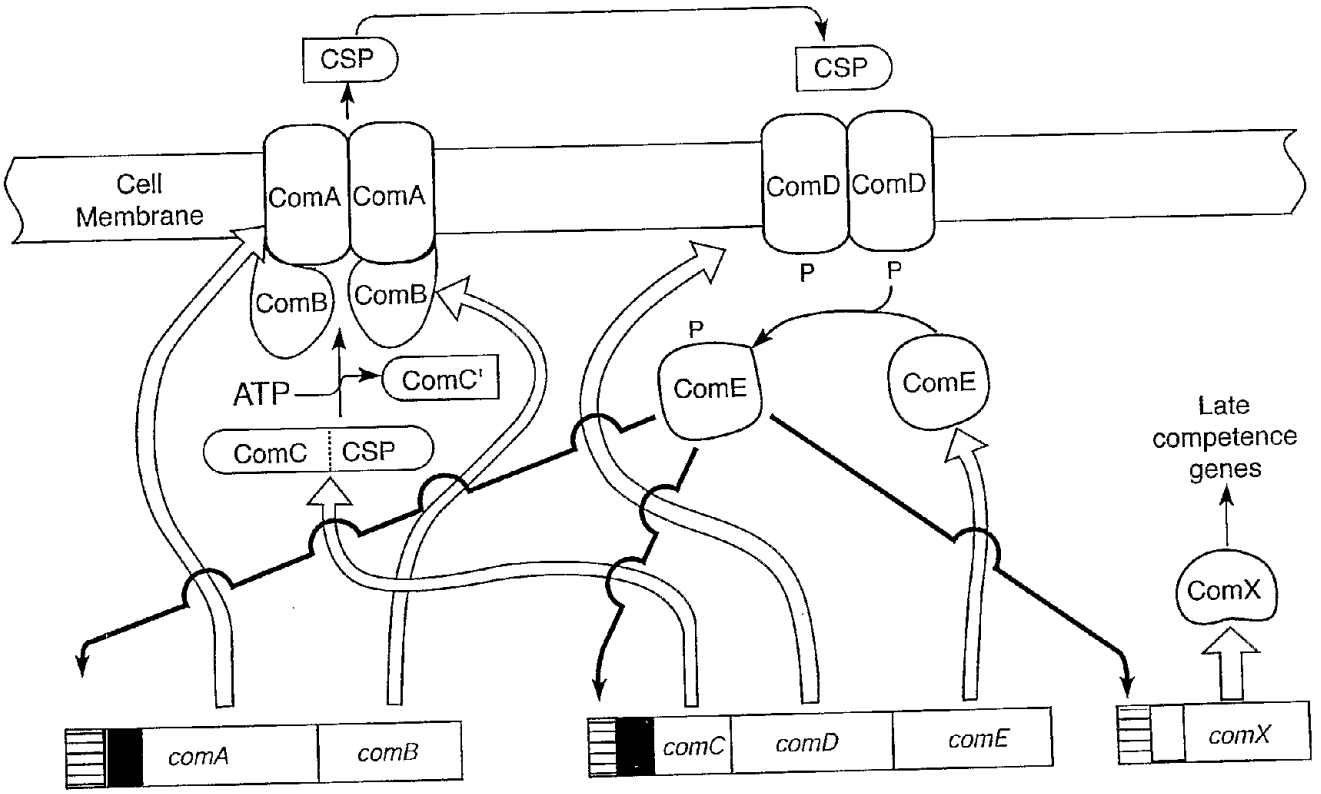

CSP
CSP
Cell
Membrane
ComA
ComA
ComD
ComD
P
P
ComB
ComB
P
ATP
ComC'
ComE
ComE
Late
competence
genes
ComC
CSP
ComX
comA
comB
comC
comD
comE
comX

Fig. 3

A. *S.pneumoniae* R6 MKNTVKL-EQFVALKEKDLQKIKGG↓EMRLSKFFRDFILQRKK (CSP1)

S.pneumoniae F5 .....-.....↓...I.RIIL..LFL... (CSP2)

S.mitis B5 .....-.....↓.S.LPKIRF..I.P...

S.mutans NG8 ..K.LS.KDD.KEI.TDE.EI.I..↓SGS.STFFRL.NRSFTQALGK

B. R6 ComD [1-48] MDLFGFGTVIVHFLIISHSYHFICKGQINRKELFVFGAYTLLTEIVFB

F5 ComD [1-48] ...L.....RL....R.....YI.....LE

R6 ComD [49-96] FPLYILYLDGLGIERFLFPLGLYSYFRWMKQYERDRGLFLSLLLSLLY

F5 ComD [49-96] .SF.L.....KI.....

C. Consensus: aCAtTTc(a)g(t)G [12] ACAgTTgAG [27] TnTGnTATAAT

comC aCAcTTtgGg [11] ACAgTTgAGa[26] TtTtgTATAAT

comA gCAgTTggGa [11] tCAtTTggGa[26] TggGgTAagAT

comX gCAgTTtaGg [11] ACAgaatgGa[26] TcTGtTAgAcT

comW cCAtTTttGa [9] gCAcTTaAaa[32] TtTttTATAcT

D. *S.pneumoniae* C D E

S.mutans C D E

Fig 4

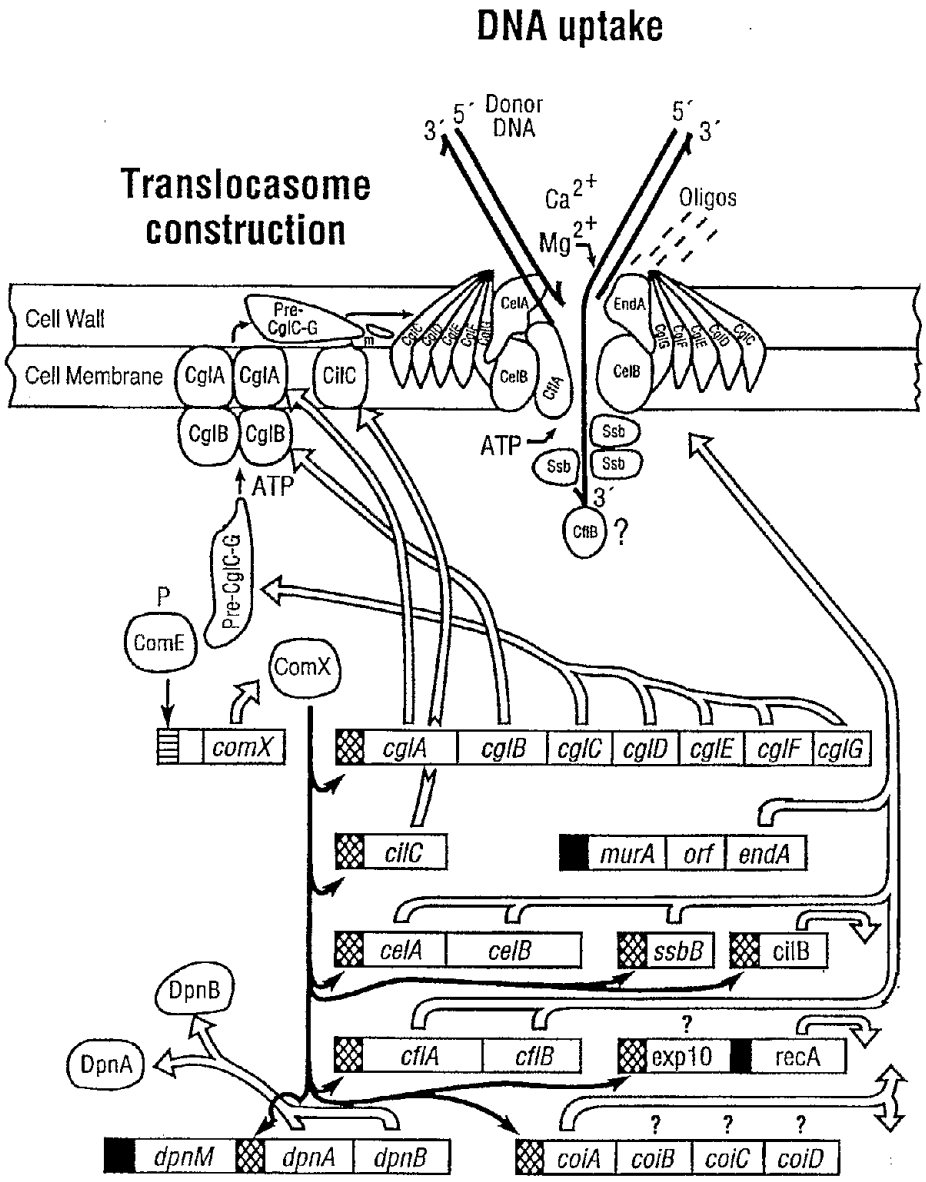
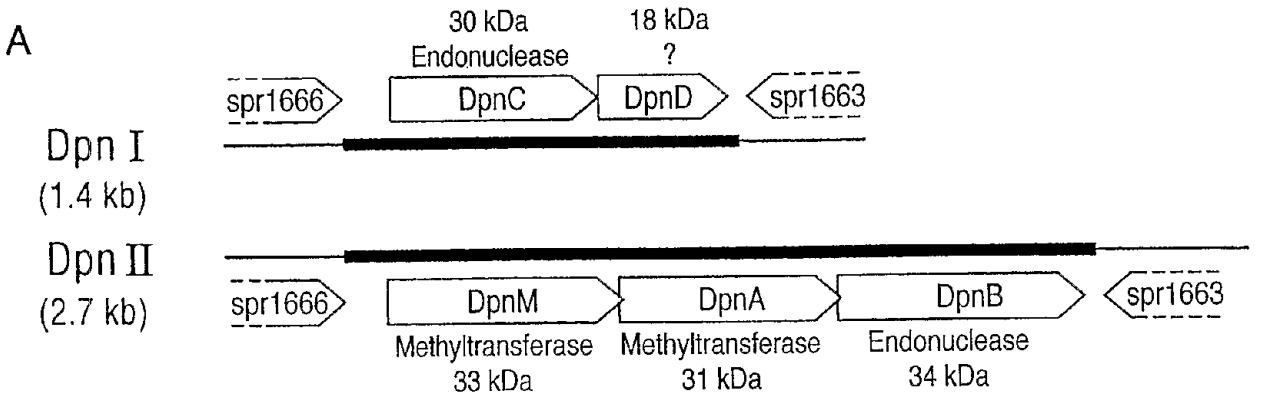
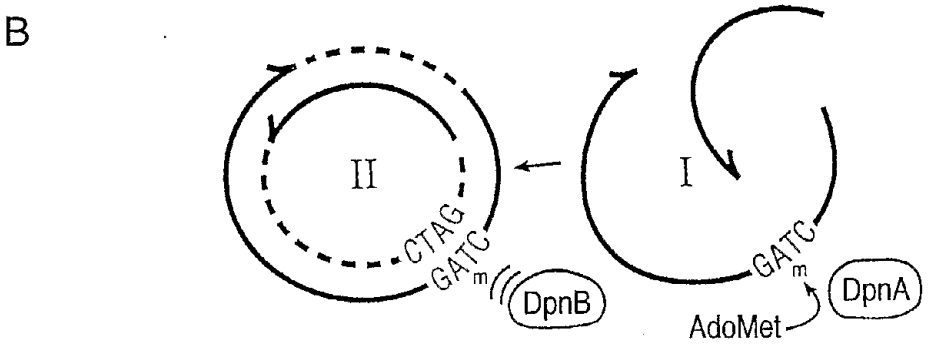

DNA uptake
Translocasome construction
Donor DNA
5´
3´
Ca2+
Mg2+
Oligos
Cell Wall
Cell Membrane
Pre-CglC-G
CglA
CglB
CilC
CelA
CelB
CilA
EndA
Ssb
CglC
CglD
CglE
CglF
CglG
ATP
CilB
?
P
ComE
ComX
comX
cglA
cglB
cglC
cglD
cglE
cglF
cglG
cilC
murA
orf
endA
celA
celB
ssbB
cilB
DpnB
DpnA
cflA
cflB
exp10
recA
dpnM
dpnA
dpnB
coiA
coiB
coiC
coiD

Fig. 5


A
30 kDa
Endonuclease
18 kDa
?
spr1666
DpnC
DpnD
spr1663
Dpn I
(1.4 kb)
Dpn II
(2.7 kb)
spr1666
DpnM
DpnA
DpnB
spr1663
Methyltransferase
33 kDa
Methyltransferase
31 kDa
Endonuclease
34 kDa


B
II
I
CTAG
GATC
m
DpnB
GATC
m
DpnA
AdoMet

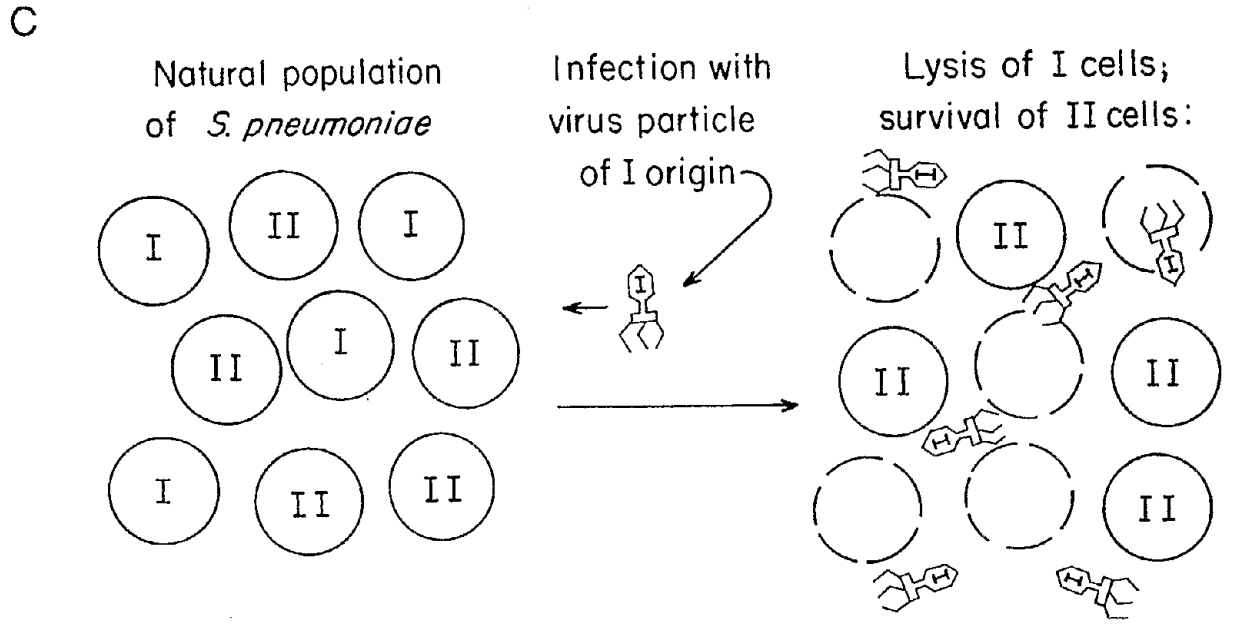

C
Natural population
of S. pneumoniae
Infection with
virus particle
of I origin
Lysis of I cells;
survival of II cells:
I
II
I
II
I
II
I
II
II
II
II
II
II

Fig. 6