MICROBIOLOGICAL REVIEWS, Sept. 1991, p. 395–424
0146-0749/91/030395-30$02.00/0

Vol. 55, No. 3

Genetic Competence in *Bacillus subtilis*

DAVID DUBNAU

Department of Microbiology, Public Health Research Institute, 455 First Avenue, New York, New York 10016

INTRODUCTION

Bacterial genetic competence may be defined as a physiological state which permits the uptake of exogenous DNA in macromolecular form. So-called natural competence, the subject of this review, is not the same as artificial competence. The latter results from the treatment of noncompetent cultures with chemical and physical agents to permit the uptake of transforming DNA. Natural competence is a physiologically and genetically determined property of a particular strain. It results from the growth of a bacterial culture under defined growth conditions or is constitutively expressed under all growth conditions. Natural competence (which will be referred to below simply as competence) is widespread among bacteria, in both gram-positive and gram-negative organisms. The best-studied systems are those of *Streptococcus pneumoniae*, *S. sanguis*, *Haemophilus influenzae*, *Neisseria gonorrhoeae*, and *Bacillus subtilis*. The emphasis here will be on the regulation of competence expression and on the machinery involved in the binding, processing, and transport of transforming DNA. Certain other topics, such as a discussion of the gene products involved in the integration of DNA, the mechanics of recombination, and the factors determining genetic linkage, will not be presented in detail. Although this review will focus on the *Bacillus* system, it will occasionally adopt a comparative approach when this promises to be informative. Competence and genetic transformation have been the subject of several recent reviews, to which the interested reader is referred (50–52, 60, 77, 103, 191, 203). The present article summarizes information available to me up to March 1991.

Why Study Competence?

In several organisms, including *Bacillus subtilis*, competence has provided an important means of genetic exchange, which has been used for genetic analysis. As such, an improved understanding of the regulation and mechanism of competence promises to enhance the usefulness and facilitate the interpretation of transformational crosses. Of more intrinsic interest are two additional aspects.

In *B. subtilis*, competence is usually expressed postexponentially. It constitutes an experimentally amenable global regulatory system and is one of several postexponential systems in *B. subtilis* that are of considerable current interest (including competence, sporulation, motility, the expression of degradative enzymes, and antibiotic production). Competent cells are able to efficiently bind, process, and internalize exogenous high-molecular-weight DNA. This poses interesting problems concerning macromolecular transport. Earlier work focused on the alterations in DNA molecular weight and strandedness that accompany the transformation process. Recently the characterization of gene products that are involved in DNA uptake has begun, allowing investigators to pose new questions concerning this process. In this review emphasis will be placed on the newer work, which concerns the regulation of competence and the nature of transformation-specific gene products.

Properties of Competent Cultures

Competent cultures of *B. subtilis* are heterogeneous. This was first suggested by the analysis of transformation for unlinked pairs of markers (152). The conclusion from this indirect analysis was then substantiated by the resolution of competent cultures on isopycnic gradients of Renografin into a majority noncompetent population of high buoyant density and a minority subpopulation of more buoyant competent cells (87, 91). An independent study revealed that the competent cells could also be resolved on the basis of sedimentation velocity in sucrose gradients (188).

Further study revealed that competent cells were in an altered metabolic state (47, 129, 152). They were found to be relatively dormant with respect to most forms of macromolecular synthesis and were reported to have completed a round of chromosomal replication, with the result that their chromosomes were lined up in a terminated configuration. It has also been suggested that competent cells may exhibit a block in deoxyribonucleotide synthesis (123). Relatively little has been reported concerning the ultrastructure of competent cells of *B. subtilis*, although Vermuelen and Venema have suggested that they contain a higher density of mesosome structures than do noncompetent cells (226). Although for the most part these early studies have not been confirmed and extended, they serve to underscore the metabolic and morphological uniqueness of the competent state.

Regulation of Competence

The earliest protocols described for the growth of competent cultures of *B. subtilis* involved the use of glucose minimal medium supplemented with a mixture of amino acids (7, 200). These protocols also involved a two-step procedure, in which the cultures were grown to early stationary phase and then diluted into a less rich version of the same medium. Competence developed after further incubation, with little cell multiplication. However, variation of the procedure demonstrated that a step down was not essential and that competence also developed following a single cycle of growth in a suitable medium (3, 242). Although the minimal requirements for competence have not been systematically worked out, all the published protocols involve the use of glucose minimal salts-based media, and substitution of glucose by other carbon sources decreases the level of competence achieved. Finally, it is interesting, and possibly of great importance, that the growth of *B. subtilis* in competence medium conditioned by the prior growth of a culture to competence results in markedly accelerated development of competence in the second culture (101). This was not due to exhaustion of the medium or to removal of an inhibitor by the first growth period, since the stimulating activity could be removed by passage through nitrocellulose filters or through a nitrocellulose column. These experiments were interpreted as indicating the participation of a soluble signaling factor.

It appears, then, that the development of competence in *B. subtilis* is subject to three types of regulation: nutritional, growth stage specific, and cell type specific. These regulatory modes will be discussed in detail below.

Regulation of Competence in Other Systems

In *S. pneumoniae*, competence develops during exponential growth rather than postexponentially, is transitory, and develops in all the cells in the culture (216, 217). In this organism a soluble competence factor has been isolated and partially characterized; it appears to be a basic protein with a molecular mass of about 10,000 Da (217, 218). It appears that accumulation of this factor to a critical level triggers the development of competence in all the cells of the culture. The development of competence is correlated with the appearance of a set of new proteins, several of which appear to be membrane localized (139, 227).

In *H. influenzae*, like in *B. subtilis*, competence is expressed poorly in exponential cultures. Again like in *B. subtilis*, conditions that limit growth appear to induce the development of competence, but in all the cells of the culture (reviewed in reference 103). No evidence for a soluble competence factor has been presented.

Finally, in *N. gonorrhoeae*, competence is constitutively expressed and therefore is presumably not subject to the growth stage and nutritional forms of control typical of the other systems discussed (16, 198).

These examples, which could be extended to additional organisms, illustrate the variety of patterns of behavior exhibited by competence development. It is not yet known to what extent the machinery involved in the binding and uptake of transforming DNA is related in these various organisms. It appears, however, that the regulatory apparatus that serves to turn on competence is highly variable and evidently has been selected to serve the special needs of each organism.

TRANSFORMATION PATHWAY

Much of the earlier work characterizing the transformation pathway traced the fate of transforming DNA during binding, processing, and uptake by the competent cell. The result of these studies was a description of the molecular weight and strandedness of transforming DNA at various stages of the transformation process, but relatively little understanding of the gene products involved in transformation and of the processing of the recipient chromosome. A detailed description of the experiments that have led to this descriptive pathway is beyond the scope of this review, and instead a relatively brief summary of the results of these studies will be presented. First, however, it may be useful to describe the various types of transformation that have been studied.

Transformation Classified by Type of Donor DNA

The transformation process may be subdivided according to the nature of the donor molecule. For instance, the donor may consist of phage DNA (transfection), plasmid DNA, or DNA derived from the bacterial chromosome. Although the outcome of the transformation process will differ in each case, the uptake of DNA probably proceeds by a common pathway. This underscores an important feature of the *B. subtilis* competence system, namely that the binding and uptake of DNA are not base sequence specific. In transfection, a successful event will result in a burst of phage particles, usually detected as an infectious center on a lawn of indicator cells. In plasmid transformation, a clone results in which the donor plasmid has been established as an autonomously replicating element. In chromosomal transformation, the donor DNA is rescued by recombination with a homologous resident DNA segment and the transformation event is usually detected by the permanent genetic alteration of the recipient.

A form of plasmid transformation in which the donor DNA can recombine with a homologous plasmid replicon has also been described (35, 133). This has been used to increase the efficiency of molecular cloning experiments in *B. subtilis* (84) and has served as a model for chromosomal transformation (35, 235, 236).

Not only does the fate of the DNA differ in these various types of transformation experiments, but so does the efficiency of the various events. For instance, nearly one transformant results from each molecule of chromosomal DNA taken up (18, 52, 54, 196). However, in transfection and plasmid transformation, several hundred molecules are taken up for each successful transformation event, measured as a colony or a plaque on a petri dish. The major consequence of this difference lies in the inapplicability of experiments that determine the biochemical fates of bulk plasmid or bacteriophage DNA in order to draw inferences concerning the pathways of these types of transformation, since only a very small proportion of the DNA taken up participates in a transfection or plasmid transformation event. Despite this difficulty, models to account for these varieties of transformation have been advanced, but will not be discussed further here (46, 57).

Stages during Uptake of Transforming DNA

Several stages have been discerned during the processing of transforming DNA: binding, fragmentation, uptake, and, in chromosomal transformation, integration and resolution.

Binding. Binding is usually measured by the wash-resistant association of radioactive DNA with competent cells, although a filter-binding assay may also be used (48). Noncompetent cells exhibit little or no DNA binding. The attachment of DNA to the competent cell proceeds with no discernible lag and is saturable (49, 54, 116). Two studies have concluded that about 50 binding sites are present on the surface of the average competent cell (54, 187). The mass of DNA bound at saturation is proportional to the molecular weight of the donor DNA molecules (49). This suggests that each binding site is occupied by a fixed number of DNA molecules (presumably one), rather than by a fixed mass of DNA. Immediately after binding, the DNA is sensitive to hydrodynamic shear and to exogenous DNase (54, 116). It is therefore extended into the aqueous phase and must be attached to the cell surface at relatively few points per molecule. This inference has been confirmed by electron microscopy (50, 160). All, or nearly all, of the bound DNA constitutes a precursor form for the generation of chromosomal transformants, since the addition of heterologous DNA to prevent further binding by competition permits the nearly quantitative conversion of previously bound radioactive DNA to either the integrated form or acid-soluble material (41, 54). The binding of DNA is not accompanied by immediate double-strand cleavage, since intact donor molecules can be recovered from the cell surface by treatment with phenol and detergent (49). It is possible that binding does occur concomitantly with nicking; this question has not been addressed for *B. subtilis*. The fact that bound DNA can be removed by phenol-detergent treatment indicates that its attachment is noncovalent. The precise location of the binding sites on the cell surface is not known; they may be localized on the cell wall, on an exposed portion of the

membrane, or on a specialized structure that traverses both. Although a careful quantitative study has not been reported, there appears to be little if any base sequence specificity in binding. All DNA samples tested appear to be bound (and taken up) with approximately equal frequency. This is evident when using DNA from *B. subtilis*, *Escherichia coli*, phage T7, and a variety of other phages and plasmids (54, 64, 70, 80, 195). On the other hand, it has been reported that glucosylated DNA, double-stranded RNA, and various synthetic polymers bind poorly if at all (31, 195).

Fragmentation. Very rapidly, the bound DNA suffers double-strand cleavage (11, 49, 55). The cleaved molecules are relatively resistant to removal by hydrodynamic shear, possibly because they are not as greatly extended into the liquid phase as the intact molecules initially bound. However, they are still completely susceptible to exogenous DNase and are therefore exposed on the cell surface (54, 55).

The pattern of cleavage has been studied by using radiolabeled phage T7 DNA as well as chromosomal DNA of known average molecular size (49). This was possible because bound DNA before and after fragmentation could be quantitatively recovered in soluble form by treatment of the DNA-cell complexes with Sarkosyl in the presence of phenol. The T7 molecules each underwent an average of one double-strand cleavage, at an essentially random position. This resulted in a halving of the initial number-average molecular weight to about 12.5×10^6. The only exception to the random location of the cleavage was an apparent bias against cleavage near the ends of the molecule. When chromosomal DNA with a starting weight-average molecular weight of 125×10^6 was used, a broad distribution of recovered fragment molecular weights was observed, with some starting molecules suffering more than one double-strand break, in a time-dependent manner. An accumulation of recovered material was observed at the low end of the distribution, with the same weight-average molecular weight as the distribution of cleaved T7 DNA fragments, together with polydisperse material of intermediate size. These observations can be at least partially rationalized by assuming that cleavage occurs at fixed points on the cell surface. These may most simply be assumed to correspond to binding sites. A characteristic distribution of cleavage sizes would then be observed, depending on the stiffness of the DNA, the average distance between the binding/cleavage sites on the cell surface, and other factors, such as the preferred range of angles at which a DNA molecule emerges from the binding/cleavage site. The average size of the fragmented molecules held on the competent cell surface may therefore be determined in part by the distribution of binding/cleavage sites on the cell surface and the probability, given the stiffness of DNA, that the end-to-end distance of any segment will coincide with the distance between two sites. As noted above, the observed number-average lower-molecular-weight limit of DNA recovered from the cell surface was 12.5×10^6, corresponding to a contour length of 6.4 μm. When DNA is free in solution, this would correspond to an average end-to-end distance of about 0.74 μm (26). The range of molecular weights recovered from bound T7 DNA was 5×10^6 to 21×10^6, corresponding to end-to end distances of 0.47 to 0.96 μm. Although these appear to be reasonable values, of the same order of magnitude as the dimensions of a *Bacillus* cell, it is not possible to precisely predict the observed distribution of sizes on the basis of current understanding. The geometric ideas proposed to account for the fragmentation of donor DNA therefore remain plausible but unproven.

Uptake. As mentioned above, donor material is attached to the cell surface in a form that is completely accessible to exogenous DNase. Beginning at 1 to 2 min following the addition of DNA to the competent cells, transformation becomes resistant to incubation with DNase (116). The lag preceding this resistance increases with decreasing temperature. The precise interpretation of DNase resistance is uncertain. It may represent transport across the cell membrane or penetration into a site at which the DNA is protected by the cell wall from access to the added nuclease. It was observed that transformation becomes cyanide resistant with a distinct time course following the acquisition of DNase resistance (209). This was interpreted as suggesting that cyanide resistance reflects transport, whereas DNase resistance measures some prior step. However, steps following transport, such as integration, might easily be energy dependent. In fact the cyanide sensitivity appears at 6 to 8 min, which is when integration occurs. It seems prudent to use a relatively neutral term such as uptake to describe the step resulting in DNase resistance and to defer a more precise interpretation.

Strauss (207, 208) has used the kinetics of DNase resistance to study the uptake of single and linked pairs of transforming markers. He has shown clearly that pairs of markers are taken up with a longer lag than are single markers and that the length of this lag is proportional to the genetic distance between the markers. This was interpreted as indicating that uptake is linear. Since several single markers were taken up with identical kinetics, it appears that the competent cell does not distinguish the "right" and "left" ends of transforming molecules, a conclusion that is in accord with the observation that there is no apparent base sequence specificity in the binding and uptake steps. The lag observed in the uptake of linked marker pairs was used to infer that about 55 bp/s are taken up at 28°C. Since the chromosomal DNA region used for these studies has now been completely sequenced (94), it is possible to arrive at a more reliable estimate. The extreme markers studied (*aroE1* and *mtr*) are separated by about 15 kb. Since the time required for entry of this marker pair is 1.4 min, DNA enters at a rate of about 179 bp/s, at least for this segment.

Donor DNA is recoverable from transformed cells as single-stranded fragments (158). The weight-average molecular weight of this material is 3×10^6 to 5×10^6 (41). Kinetic studies have shown that these single strands, which are resistant to the action of exogenous DNase, are derived from the surface-localized double-stranded fragments and are precursors of integrated donor DNA (41). The first appearance of single-stranded material coincides with the development of DNase-resistant transformation. In addition, the appearance of single-stranded DNA is coincident in time with the release into the cell medium of acid-soluble products derived from donor DNA (41, 55, 102). The acid-soluble and single-stranded products each make up about half of the material initially bound to the cell surface, which suggests that one strand may be completely degraded. When the transforming DNA was labeled with [^{3}H]thymidine, this acid-soluble material was recovered as TMP, thymidine, and thymine (55). It is tempting to conclude that these three events (the appearance of acid-soluble material, single-stranded DNA, and DNase-resistant transformants) all reflect a single concerted process. Similar reasoning (based on data obtained with *S. pneumoniae*) has led Rosenthal and Lacks to propose a model in which a nuclease, asymmetrically located in association with a membrane pore for DNA transport, degrades one strand (173). This would then result

in the release of acid-soluble material into the medium and could drive the concomitant transport of the complementary strand. This model will be discussed further below.

Recently, Vagner et al. (221) have investigated the chemical polarity of uptake. A variety of DNA molecules were constructed with a radioactive label on one strand, either centrally located or near the 5′ or 3′ terminus. After 15 min at 32°C the uptake of radioactivity by the competent cells was measured and found to be quite similar with the various donor DNA samples. This was interpreted as indicating either that single strands are taken up with no preferential polarity or that uptake of double-stranded DNA occurs initially, with the subsequent intracellular degradation of one strand. Although neither of these possibilities is ruled out by other data, the former would appear to be unlikely, since a single strand entering with 5′-to-3′ or 3′-to-5′ polarity would present very different aspects to the uptake machinery, requiring the existence of two uptake systems. The possibility was also considered that uptake of an entire strand took place in these experiments, thus obscuring the underlying polarity of uptake. This possibility was rejected, basically because *B. subtilis* cells introduce double-strand breaks into DNA from which uptake is thought to be initiated, and this was assumed to prevent the uptake of an entire strand. However, a reconsideration of these interesting experiments is warranted because of two factors. First, the donor molecules used were derived from M13 phage and had a molecular weight of about 4.5×10^6. The smallest molecules used in studies of double-strand cleavage on an competent cell surface had molecular weights of about 2×10^7 (49, 55). It is uncertain whether molecules as small as those used in the study of Vagner et al. (221) undergo efficient cleavage. Second, the samples were taken for determination of uptake after 15 min. By this time, uptake has usually been completed, and even if cleavage did occur, it is possible that an entire strand was taken up in segments, one from a preexisting molecular terminus and the other from the point of cleavage. It would be better to carry out such studies as a function of time. Finally, it should be noted that for *S. pneumoniae*, which exhibits a transformation pathway similar to that in *B. subtilis*, a 3′-to-5′ polarity of uptake has been indicated (131).

Energetics of uptake. The transport of DNA into the competent cell is extremely efficient. About half of the mass of donor DNA initially bound to the cell surface is found as DNase-resistant single strands. Obviously, a DNA-specific transport mechanism must exist in competent cells. As mentioned above, Rosenthal and Lacks (173) have suggested that transport occurs through a membrane channel that is associated with an asymmetrically located nuclease. Asymmetry of the nuclease active site ensures that only one strand is degraded during transport, and it is implicit in the model that the driving force for transport is provided by the action of the nuclease. In support of this model are two observations. First, in *S. pneumoniae*, like in *B. subtilis*, uptake of single-stranded DNA occurs after the binding of double-stranded DNA. Second, Rosenthal and Lacks (173) have identified a membrane-localized nuclease required for transformation and DNA entry in *S. pneumoniae*. This is an attractively simple model, which is consistent with the data so far presented for the *B. subtilis* system.

Grinius (81) has used uncoupling agents in a variety of transformation experiments and has concluded that DNA uptake in *B. subtilis* requires both the pH and the electrical components of the proton motive force. Poisoning with arsenate led Grinius to suggest that uptake is not driven directly by ATP. He proposed instead that the transforming DNA binds to proteins on the cell surface, that this complex acquires a net positive charge by binding protons, and that the entire complex then electrophoreses through a water-filled membrane channel. This electrogenic mechanism explains the need for both components of the proton motive force. Van Nieuwenhoven et al. (222) have criticized aspects of the Grinius model, suggesting that the ΔpH alone provides the driving force. Theirs is an electroneutral proton symport mechanism, and its continued operation would require pumps to prevent the intracellular accumulation of protons. Neither model would explicitly explain the requirement for a membrane nuclease, and neither relates the transport process to the appearance of single strands. Both of these studies measured ΔpH and $\Delta\Psi$, and Grinius also studied the ATP pool, in bulk competent cells. Since the competent cell fraction is physiologically distinct, and makes up about 10% of the culture, it would be important to repeat such measurements in the competent cell subfraction. A further, highly speculative model, which is consistent with all of the existing data, has been proposed (51). This proposal was suggested by the anion-exchange model of sugar phosphate transport (4, 5, 125). Perhaps duplex DNA enters a water-filled membrane channel, where it encounters a nuclease, with the result that one residue is cleaved and expelled, possibly by reorientation of a membrane exchange center. If the ΔpH is appropriate, the incoming base pair, consisting of two molecules of monovalent nucleotide, may release two protons in the cytoplasm. If a single divalent nucleotide is expelled, electroneutrality would be maintained, but the system would be continually coupled to proton circulation. Clearly, more data are needed to distinguish between these and additional models for the energetics of DNA transport.

Integration and resolution. As stated above, single-stranded donor molecules of 3×10^6 to 5×10^6 daltons are recoverable from transformed cells. These are direct precursors of integrated material (donor-recipient complex); about 72% of the single-strand DNA mass eventually is integrated (41).

The product of recombination is a heteroduplex consisting of paired donor and recipient strands (17, 53, 58, 210). The average size of the integrated donor segment has been determined to be 2.8×10^6 (weight-average molecular weight) or ca. 8.5 kb in one study, in which physical measurements were used (53), and 2.9×10^6 to 3.9×10^6 in another study, in which electron microscopy was used (72). It was predicted (54), on the basis of what was known about transformation in *B. subtilis*, that the integrated segments would be clustered. This follows from the fact that a donor DNA molecule is fragmented on the cell surface and that uptake and integration of the resulting fragments proceed efficiently. This should result in a number of clustered but independent integration events from each molecule that was initially bound. This prediction was confirmed by both studies.

Early during the integration process, a noncovalent association of donor and recipient DNA was detected, from which the donor moiety could be liberated at pH 12 (10, 56). The freed donor single-strand moiety had a weight-average molecular weight of 2.5×10^6 to 3×10^6, in good agreement with the measured sizes of both the precursor single strands and the final integrated moiety. Arwert and Venema (10) have suggested that nicks rather than gaps separate the donor and recipient moieties of the joint molecules, since DNA ligase alone increases the transforming activity of the joint molecule preparation as much as does ligase plus DNA

polymerase. The structure of the joint molecules is not known; it may consist of a branched triple-stranded molecule, a completely base-paired molecule, or some other form. Also, little is known about the mechanics of recombination in this system, although it is tempting to think in terms of a strand assimilation mechanism, since RecE, the *B. subtilis* analog of the *E. coli* RecA protein, is required for integration (59).

The final heteroduplex molecule formed by the integration step can be resolved by replication or by mismatch repair to form a homoduplex. Mismatch repair is known to occur following transformation of *B. subtilis*, and the probability of repair depends on the nature of the mismatch (23, 199). A precise description of this dependence and information about the probability of cocorrection of linked markers as a function of distance has not been reported for *B. subtilis*.

Transformation Pathway in Other Systems

The binding and further processing of transforming DNA in the *S. pneumoniae* system is similar to that in *B. subtilis*, at least superficially (73, 86, 108, 110, 113, 140, 141). In both gram-positive systems, binding to the cell surface is followed by fragmentation, uptake to yield intracellular single strands, and integration to form a heteroduplex molecule. The process appears to be somewhat slower in *S. pneumoniae*, and in this system binding to the cell surface is accompanied by nicking of the DNA prior to double-strand cleavage (109). This event has not been demonstrated to occur in the *B. subtilis* system.

In the gram-negative *Haemophilus* system the process of transformation appears to be quite different (103). First, binding exhibits base sequence specificity, in contrast to the *Bacillus* and *Streptococcus* systems. In fact, a binding recognition sequence has been characterized in *H. influenzae* (39, 40, 189). In addition, double-strand transforming DNA appears to associate first with discrete membranous vesicular structures located on the competent cell surface, which have been called transformasomes. This association renders the transforming DNA resistant to exogenous DNase. The DNA is then transported into the cytoplasmic compartment and integrates to form a heteroduplex. It is believed that transport from the transformasome into the cytoplasm is accompanied by degradation of one strand and that the 3′ end of the other strand participates preferentially in the integration event. This is the same polarity observed in *S. pneumoniae* (131).

The energetics of DNA transport in the *S. pneumoniae* system has been examined (32, 33, 120); in this system the interpretation of results is not complicated by the heterogeneity of the competent culture, as it is in *B. subtilis*. These studies have led to the conclusion that the uptake of DNA in *S. pneumoniae* is driven directly by ATP, in apparent contrast to the situation reported for *B. subtilis*.

LATE COMPETENCE GENE PRODUCTS

Types of Gene Products

Several approaches have identified gene products required for transformation. These can be classified into two groups. One type of product plays a regulatory role and is needed for the appropriate expression of other competence-specific proteins. These regulatory proteins have also been referred to as early competence products. Since some of them are not expressed constitutively, but only in certain media and at certain growth stages, the term "regulatory product" is preferable. Other proteins either are required directly for the binding and processing of DNA or are needed for the proper assembly of the competence machinery. These have been called late competence proteins. Since it is not yet possible to distinguish those that are directly involved in transformation from those that play a morphogenetic role in the assembly of the competence apparatus, we will retain the term "late proteins." The regulatory and late genes have usually been distinguished by using genetic tests for the dependencies of expression of one gene on another, most often with fusions of target gene promoters to the *E. coli lacZ* marker. All but one of the known late gene products have been identified by the characterization of competence deficient mutations followed by the cloning and sequencing of the cognate gene. In a number of cases the proteins encoded by these genes have been studied directly.

Description of Late Competence Gene Products

Competence-specific single-strand-binding protein. The single exception to the above noted generalization has been the discovery and preliminary characterization of a competence-specific single-strand-binding protein (65). This was detected as an activity that protected denatured radiolabeled DNA from degradation by DNase. This activity was missing from noncompetent cells and was dependent for its expression on an uncharacterized sporulation (*spo*) gene, probably an allele of *spo0A*. This is interesting, since *spo0A* is required for the expression of competence (*com*) genes. Unfortunately, no further characterization of the single-strand-binding activity has been reported. Since many gene products are expressed postexponentially, it remains to be demonstrated that this single-strand-binding protein is required for transformation.

Competence nuclease. Two other potential late competence gene products were studied biochemically and genetically by Smith et al. (192–194). These are a 75-kDa membrane-localized protein complex, which contains a 17-kDa nuclease, and an 18-kDa protein that seems to limit the activity of the nuclease. The genes encoding these two proteins (*comI* and *comJ*, respectively) have been cloned and sequenced, and a null mutation, at least in *comI*, was reported to be partially competence deficient (228). However, recent evidence has called this conclusion into question (224); a null mutation of the nuclease gene results in a mutant that exhibits normal or near-normal transformability. Therefore, no *B. subtilis* analog of the *S. pneumoniae* entry nuclease (173) has been clearly identified to date. However, it has been pointed out by Van Sinderen (224) that an alternative entry nuclease may exist in *B. subtilis*, in view of the report by Vagner et al. that uptake may occur with either polarity (221). If this is true, inactivation of only one of these nucleases would have a minor effect on transformability.

***div-341* (*secA*) locus.** An interesting mutation (*div-341*) has been described that results in a pleiotropic phenotype (178, 180). This mutation, which maps near the *degS-degU* genes (Fig. 1), results in filamentous growth above 45°C. At an intermediate temperature (37°C), competence and sporulation are strongly inhibited and exoenzyme production is also limited. At a permissive temperature (30°C), competence, sporulation, and exoenzyme production are normal. These pleiotropic effects would seem to suggest the classification of the *div-341* locus as regulatory. However, cloning (176) and sequencing (181) of the *div-341* gene have revealed the presence of an open reading frame (ORF) with striking similarity to that of the *E. coli secA* gene, at both the DNA

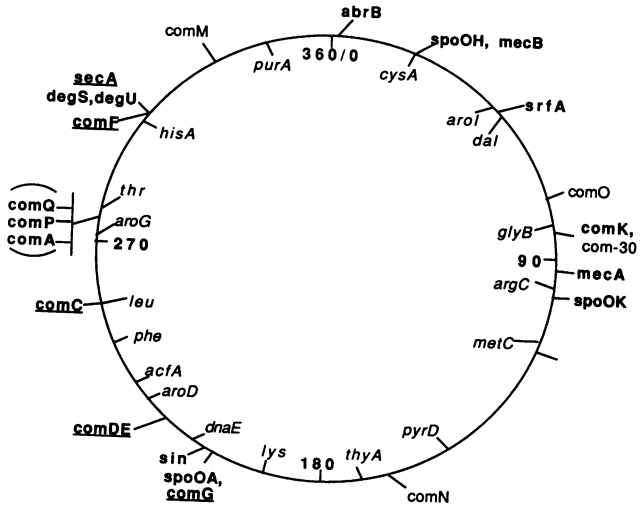

FIG. 1. Genetic map of *B. subtilis* competence genes. The known competence loci are indicated on the outside of the circle. The italicized markers on the inside are included for reference. The boldface type without underlining indicates regulatory loci. Boldface underlined type designates late competence loci. The plain typeface indicates loci that may include either regulatory or late competence genes. The order of the *comQ-P-A* cluster relative to flanking markers is not known.

and amino acid sequence levels. The *div-341* mutation results from a single amino acid change in this ORF (177). This finding provides an obvious explanation for the exoenzyme deficiency of the *div-342* mutants, since *secA* lies on the main secretion pathway in *E. coli* and presumably in *B. subtilis* as well. The deficiency in sporulation and competence may result from the failure to secrete products essential for these systems. Results of a further experiment lend support to this notion (177). The mutant strain was grown to late exponential phase at 30°C and then transferred to a second medium for 90 min to allow the development of competence. Competence was then measured by a third incubation with transforming DNA for 30 min. Transformability was high when competence was allowed to develop at the permissive temperature, whether incubation with transforming DNA was at the permissive (30°C) or the nonpermissive (37°C) temperature. When the second incubation occurred at the nonpermissive temperature, however, almost no transformants were obtained. This experiment is consistent with the idea that the *div-341* (*secA*) gene plays a role during the development of competence, rather than during the binding and uptake of DNA. However, these data may be interpreted in several ways. SecA may play either a regulatory or a morphogenetic role during the development of competence. It is also possible that the SecA product plays a direct role during transformation, but that the *div-341* mutation prevents the synthesis or proper folding of SecA rather than its activity. This would restrict the window during which incubation at the nonpermissive temperature can inhibit transformation to the period of formation of the competence machinery. It would be useful to determine the dependencies of late gene expression on SecA and to determine the effect of *div-341* on the assembly of known late competence proteins on the surface of competent cells. It is also possible that the effect of the *secA*(Ts) mutation on competence is indirect. For instance, the intracellular accumulation of precursor proteins or the effect of the SecA lesion on the regulation of other Sec proteins (183) may interfere with the development of competence.

Genetic identification of additional late competence genes. Several competence loci have been identified following the isolation of *com* mutants. In some cases it is not possible to firmly classify these mutants as having alterations in regulatory or in late genes. For instance, Fani et al. (67) used nitrosoguanidine to isolate four *com* mutants. These four mutants were reduced in transformation by 3 to 5 orders of magnitude, but were not reduced in transduction or protoplast transformation. Three of the mutations cannot yet be classified as defining regulatory or late genes. One of these (*comM104*) defines a new gene, based on its map position (Fig. 1). The fourth, *com-30*, was deficient in uptake, but bound nearly wild-type amounts of DNA. Therefore, it probably defines a late competence gene. Its reported map position is near that of *comK*, a regulatory gene (225).

Mastromei et al. (128) have used the transposon Tn*917* to identify at least four distinct *com* genes. Of these, two can be said to be distinct from other previously identified loci. One of these (*comO18*) was strongly blocked in DNA binding and was depressed in transformation about 10,000-fold. The other (*comN114*) was depressed about 200-fold and took up about two-thirds as much DNA as did the wild type. The approximate map locations of *comO* and *comN* are shown in Fig. 1. These are candidates for late competence genes, but their roles in competence cannot be specified with certainty.

The powerful modifications to Tn*917* now available (246–248) have permitted the isolation of a number of *com* mutants as gene fusions to a promoterless *lacZ* element and the ready cloning of the *com* loci so identified (3, 88). The availability of *lacZ* fusions has permitted the examination of the expression patterns of these new *com* determinants and their classification as regulatory or late genes. Several of the late *com* loci have been characterized by sequencing and transcriptional mapping, and the latter aspect will now be described.

Transcriptional Organization and Phenotypic Properties of Late Competence Loci

Identification of late *com* genes for further study. By the following criteria, *comG*, *comC*, and *comDE* have been identified as late *com* transcription units. First, they are expressed only at the time of transition to the postexponential state and in competence medium. Second (with one exception to be discussed below), they are not required for the resolution in Renografin gradients into competent and noncompetent subpopulations. Third, the expression of these genes is restricted to the competent Renografin-resolved subpopulation. Fourth, the expression of these genes is dependent on several other determinants identified as regulatory genes, but not on one another. Finally, the late *com* mutants appear otherwise normal in growth, colony morphology, and other late growth properties such as sporulation. This is in contrast to certain of the regulatory mutants, which exhibit pleiotropic phenotypes. The transcriptional organization of each of these loci will now be described, as well as the phenotypes conferred by their cognate mutations.

comG. Several closely linked insertions of the transposon Tn*917lacZ* were isolated that resulted in a complete loss of transformability; competence was reduced at least by 6 orders of magnitude (1–3, 88). The mutant *com* strains grown through the usual competence regimen were unable to bind transforming DNA. Transcription of the mutant locus oc-

curred normally as judged by β-galactosidase production from the *lacZ* fusion construct and by S1 nuclease and primer extension measurement of mRNA synthesis. The region of DNA in which the transposons had inserted was cloned, sequenced, and transcriptionally mapped; it was shown to consist of an operon (*comG*) with seven ORFs (1, 2). The end of ORF3 and the beginning of ORF4, the end of ORF4 and the beginning of ORF5, and the end of ORF5 and the beginning of ORF6 overlap to various extents. This configuration suggests that these four ORFs may be translationally coupled, as if their products are synthesized in fixed proportions. The map location of *comG* is shown in Fig. 1.

Although evidence was obtained for at least one and possibly two minor downstream promoters, these promoters were not able to support the expression of more than very limited transformability, and it is clear that a single major promoter drives the great bulk of *comG* expression. Upstream from the transcriptional start site, identified by S1 nuclease mapping and by primer extension analysis, was a sequence that resembles a vegetative *B. subtilis* promoter. Downstream from the final ORF was a dyad symmetry element that resembles a rho factor-independent terminator.

Although Tn*917* insertions in five of the seven *comG* ORFs were identified, it is possible that not all of the ORFs are required for transformation, since Tn*917* insertion would be expected to be polar on downstream genes. However at least ORF1, ORF3, and ORF7 are essential competence genes. Inactivation of ORF1 results in a competence-deficient strain that fails to resolve into two cell populations on gradients of Renografin (1). Inactivation of ORF2 does not prevent resolution in Renografin. Complementation in *trans* of a strain with a transposon insertion in ORF1, by a plasmid carrying only *comG* ORF1, restores Renografin separation but not competence. Thus, it appears that ORF1 is essential. ORF7 is also needed for transformability, since it is the most distal gene of the *comG* operon and an ORF7 insertion is competence deficient. Finally, an in-frame deletion of ORF3 has been isolated and shown to completely eliminate transformability (22). This deletant can be complemented by ORF3 in *trans*.

comC. Other insertions of Tn*917lacZ* defined the *comC* locus, which consists of a single ORF (3, 88, 137, 138). *comC* mutants are completely deficient in competence and cannot bind transforming DNA. The location of *comC* is shown in Fig. 1. The transcriptional start site was identified by primer extension and is located downstream from a sequence that resembles a vegetative promoter, *comC*, like *comG*, is followed by an apparent transcriptional terminator element.

comDE. Finally, four Tn*917lacZ* insertions were used to identify a locus that we provisionally call *comDE* (3, 88). Mutations in this locus result in a complete loss of competence. Its map position shown in Fig. 1. The sequence of the *comDE* region has recently been completed (90) and reveals the presence of four ORFs, three of which are read in one direction, almost certainly as an operon. The Tn*917* insertion in the third (most distal) ORF results in the complete loss of competence and in a mutant strain that binds DNA but cannot take it up. Insertion in the most promoter-proximal ORF (ORF1) also results in the complete loss of competence, but abolishes the ability to bind DNA. This is presumably not due to a polar effect of the Tn*917* insertion on ORF3, because of the distinct phenotype conferred by ORF3 inactivation. Since Tn*917* is polar on downstream genes, however, it is not certain whether ORF1 or ORF2 or both are essential competence genes. The promoter of this operon has been localized by low-resolution S1 nuclease mapping to a region about 1 kb upstream from the start of ORF1 translation. Contained within this leader sequence is a large ORF that would be read in a direction opposite to that of ORF1 to ORF3. It is not known what role, if any, is played by this reverse ORF, or even whether it is expressed.

Some Properties of Late *com* Proteins

The amino acid sequences of the seven *comG* proteins inferred from the DNA sequence demonstrated that all but ComG ORF1 possess at least one highly hydrophobic probable transmembrane segment (1). The ORF1 sequence exhibits one hydrophobic region with a marginal probability of being classified as a transmembrane segment. The ORF2 protein (323 residues) has three predicted transmembrane segments. The remaining ORFs (ORF3 to ORF7) are each predicted to contain a single N-terminally located transmembrane segment; they consist of 98, 115, 165, and 124 amino acid residues, respectively.

ComG ORF1 is predicted to be a protein of 356 residues. Located within ORF1 is a sequence that resembles a nucleotide-binding site, similar to those found in other proteins such as AMP kinase, bovine ATPase, and the RecA protein (229). Antiserum was prepared against a hydrophilic synthetic peptide derived from the ComG ORF1 sequence and used to study the subcellular location of this protein (21). The ComG ORF1 signal, detected on Western immunoblots, was localized to the membrane fraction of competent cells. It could be solubilized from the membrane by extraction with 0.1 M NaOH (175) and was therefore considered to be a peripheral membrane protein. Probing of membrane and protoplast preparations with protease suggested that the ORF1 protein was located on the cytoplasmic face of the membrane.

Antipeptide antiserum was also prepared against ComG ORF3 (22). This was used to demonstrate that in competent cells the protein was in part an intrinsic membrane protein, on the basis of its insolubility in 0.1 M NaOH. All of the signal detected on Western blots could be eliminated by prior treatment of protoplasts, or even of whole competent cells, with protease. These observations, together with the single predicted transmembrane domain located near the N terminus of the ORF3 protein, suggested that about half of the protein in competent cells was located with the N terminus in the cytoplasm and the C-terminal moiety exposed at the exterior of the membrane. This was also true when ORF3 was expressed under control of the regulatable *Pspac* promoter in noncompetent cells. However, in intriguing difference was noted between the localization of ORF3 in competent and noncompetent cells. In the latter case, none of the signal was removed from the membrane by 0.1 M NaOH, whereas in the competent cell membranes, about half was solubilized. Since all of the signal could be eliminated by protease treatment of protoplasts, it was concluded that in the competent cell a portion of the ComG ORF3 protein was on the exterior of the membrane and the remainder was intrinsically associated with the membrane. This points to a specific organization of ORF3 proteins in the surface layers of competent cells. No biochemical evidence is available for the localization or properties of the remaining ComG proteins (ORF2, ORF4, ORF5, ORF6, and ORF7), although, as noted above, they are probably intrinsic membrane components.

ComC is predicted to be a polytopic membrane protein of 248 residues, containing four or five transmembrane segments (137). The ORF3 protein of the *comDE* operon is also

likely to be a polytopic membrane protein (90). This gene product, 776 amino acid residues in length, is highly hydrophobic and contains about eight predicted transmembrane segments. The ComDE ORF1 and ORF2 proteins are hydrophilic, except that ORF1 has a single predicted N-terminal transmembrane region.

Transformation Machine

The presence of at least some of the late competent products in association with the membrane and the differences in the dispositions of ComG ORF3 in noncompetent and competent cells support the existence of a surface-localized multicomponent transformation machine. It is likely that such a machine would include a DNA-binding component, a channel through the peptidoglycan layer, and an aqueous pore that extends through the cell membrane. Since the transport process is apparently coupled to an energy source (see above), the machine probably includes proteins that are required for this coupling. Finally, a nuclease capable of processively degrading one strand of donor DNA may also be present in the machine. The existence of such a multicomponent transformation machine, in which the individual components are in intimate contact, has not been demonstrated. It is, however, an attractive hypothesis, which has heuristic value since it is testable biochemically as well as genetically.

Similarities of Late *com* Genes to Other Proteins

Identification of protein similarities. We have noted interesting similarities of several *com* gene products to other proteins in the data base. Many of the similarities discussed below have been discovered independently by Whitchurch et al. (241).

Certain *com* proteins exhibit similarity to proteins of the pullulanase secretion system of *Klebsiella pneumoniae* (reviewed in reference 161). This gram-negative organism secretes the enzyme pullulanase across the inner and outer membranes and into the extracellular space. To secrete pullulanase, the product of *pulA*, at least 12 gene products are specifically required. Of these, PulE, PulF, PulG, and PulO are similar to ComG ORF1, ComG ORF2, ComG ORF3, and ComC, respectively (Fig. 2 to 6). In addition, similarity of these Com proteins to several gene products required for the processing of the *Pseudomonas* PAK pilin and the assembly of this protein into a pilus structure (154), as well as to the primary sequence of the pilin protein itself (1), is evident. Specifically, PilB, PilC, and PilD are similar to ComG ORF1, ComG ORF2, and ComC, respectively, whereas ComG ORF3 resembles the pilin protein (Fig. 2 to 6). In Fig. 6, the Z values for these and other similarities are shown in tabular form. These indicate that the resemblances of the *com* products to those of the *pul* and *pil* systems are quite significant. In addition to participating in the assembly of pili, PilD is required for the export of several extracellular enzymes in *Pseudomonas aeruginosa* PAK, and a nearly identical protein (XcpA) is similarly required in *P. aeruginosa* PAO (13, 212). Other proteins required for extracellular secretion in *Peudomonas* spp. (XcpY and XcpZ) exhibit similarity to PulL and PulM, respectively (69). Secretion systems like those encoded by the *xcp* and *pul* genes may be widespread in gram-negative organisms. For instance, the *out* system of the plant pathogen *Erwinia chrysanthemi* is required for the export of various cell wall-degradative enzymes, and several *out* and *pul* products are quite similar (92). Finally, a protein required in *P. aeruginosa* for twitching motility (PilT) is similar to PilB, PulE, and ComG ORF1 (241).

In addition to these similarities, which will be discussed further below, we have noted a weaker resemblance between ComG ORF1 and the VirB ORF11 protein of *Agrobacterium tumefaciens* (1, 186, 231) (Fig. 2 and 6). The *virB* operon is thought to play a role in the transport of T-DNA from the bacterial cell to the plant cell. It has been shown that the purified VirB ORF11 protein is localized on the cytoplasmic face of the cell membrane (as is ComG ORF1) and that it possesses both ATPase and autophosphorylating activity (30). The similarities among ComG ORF1, VirB ORF11, PilB, PilT, and PulE are most marked in the region corresponding to the nucleotide-binding site proposed above.

The relationships among ComG ORF3, ComG ORF4, and ComG ORF5, PulG, and the *Pseudomonas*-type pilin molecules (illustrated in Fig. 4 and 6) are restricted to the N-terminal portions of these molecules and correspond approximately to a portion of pilin that is conserved among the so-called *N*-methylphenylalanine pilins. These proteins, found in *Neisseria*, *Moraxella*, and *Bacteroides* spp., in addition to *Pseudomonas* spp., are major subunits of pili, which are filamentous extracellular appendages. The pilins are processed by the removal of a few amino acid residues from their N termini, followed by N-methylation of the terminal Phe residue. Nonpilated variants of *Neisseria* and *Moraxella* spp. have been reported to be competence deficient (20, 184, 198), as are strains deficient in comG ORF3 (22). It is remarkable also that ComG ORF4 to ComG ORF7 each exhibit some resemblance to ComG ORF3. All are small proteins (98 to 165 residues), all have hydrophobic N-terminal segments and very similar hydropathy profiles (1), and there appears to be some conservation of amino acid sequence within the N-terminal segments of ORF3 to ORF5 (Fig. 4 and 6). Finally, it is worth noting that in addition to the resemblance of PulG to pilin and to ComG ORF3 already noted, the N-terminal region of the PulG protein also resembles that of PulI and PulJ (241). Thus, both the Pul and the Com systems seem to encode several small proteins belonging to the pilinlike family.

Strom and Lory (211) have examined the structural requirements for the processing of pilin derived from *P. aeruginosa* PAK-NP and have discovered that at least the Gly residue immediately preceding the conserved Phe (Fig. 4) is required for cleavage. Possibly the pilinlike Com proteins are also processed (although no direct evidence for this exists) and the sequence similarly reflects requirements for the processing site. The possibility that ComG ORF3 to ComG ORF5 are processed is supported by the observation that the ComC protein closely resembles PilD (Fig. 5 and 6) and that the latter is specifically required for the processing of *Pseudomonas* pilin (154). In addition (or alternatively), the similarities may reflect requirements for assembly of the pilinlike proteins into a higher-order structure.

What do the protein similarities mean? The striking amino acid sequence similarities noted above establish the relationships between the competence system of *B. subtilis* and the pilus assembly and protein export systems of *Pseudomonas*, *Xanthomonas*, *Erwinia*, and *Klebsiella* spp. Two types of roles for the proteins involved in these relationships may be considered. First, the *pul*, *pil*, *out*, and *xcp* products may be directly required for protein secretion in the gram-negative systems, whereas their homologs in the *B. subtilis com* system may be required for the uptake of DNA. This

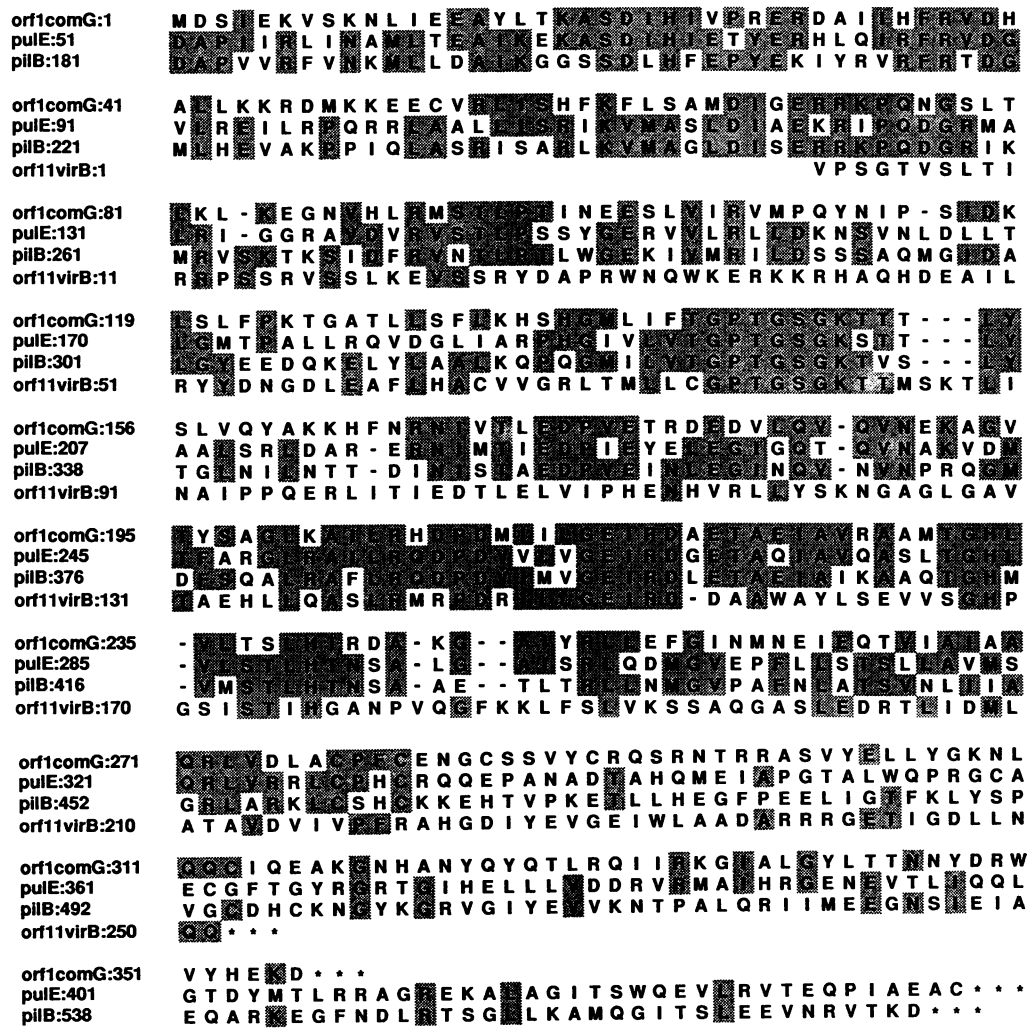

FIG. 2. Similarities of ComG ORF1 (1) to protein members of the pullulanase secretion (43), pilin assembly (154), and *A. tumefaciens* Ti plasmid (231) systems. Amino acid identities are indicated by shading whenever at least two of them occur at the same position. The residue number of the first amino acid on each line is indicated. The asterisks denote the end of the protein sequence.

hypothesis therefore proposes that many of these late competence proteins are directly involved in the transport of DNA. Alternatively, certain of these *com* products may fulfill morphogenetic roles, required for the transport or assembly of components of the transformation machine, without direct involvement in DNA processing and transport. The similarities of *com* products to those of the *pil*, *pul*, *xcp*, and *out* systems are certainly suggestive of such a morphogenetic role and strengthen the second hypothesis. The last four systems are ultimately required for the export of specific extracellular proteins through the inner and outer membranes of *Klebsiella*, *Xanthomonas*, *Erwinia*, and *Pseudomonas* spp. These export systems also require *sec* products, presumably for inner membrane transport. The latter requirement may be true of the *B. subtilis com* system as well, in view of the probable role for *secA* in competence. However, *B. subtilis* does not possess an outer membrane. What, then, would be the role of morphogenetic Com proteins that are similar to those of the gram-negative systems? It seems likely that certain of the *pul* products are needed to assemble a cell surface apparatus which in turn exports pullulanase. The Com proteins may function analogously, assembling components of the DNA transport machine.

At this time it is not possible to rigorously distinguish between a direct and a morphogenetic role for any given *com* protein. For assembly systems in which a final structure has been defined and its components have been described, it is possible to classify a gene product as fulfilling a morphogenetic role if it is required for assembly but is not a component of the structure. Such a classification appears to be possible for the *Pseudomonas* pili; PilB, PilC, and PilD are in this sense morphogenetic proteins, whereas the product of *pilA* is not. This is also the case with certain bacteriophages, in which morphogenetic proteins have been identified that are not themselves part of the completed virion. However, for the *com* system, although it is likely that a multiprotein surface-localized machine is involved in the binding, processing, and transport of DNA, this complex has not been demonstrated and its composition is certainly not understood. These ambiguities can be illustrated by the case of ComG ORF1, a probable ATP-binding protein. It is possible

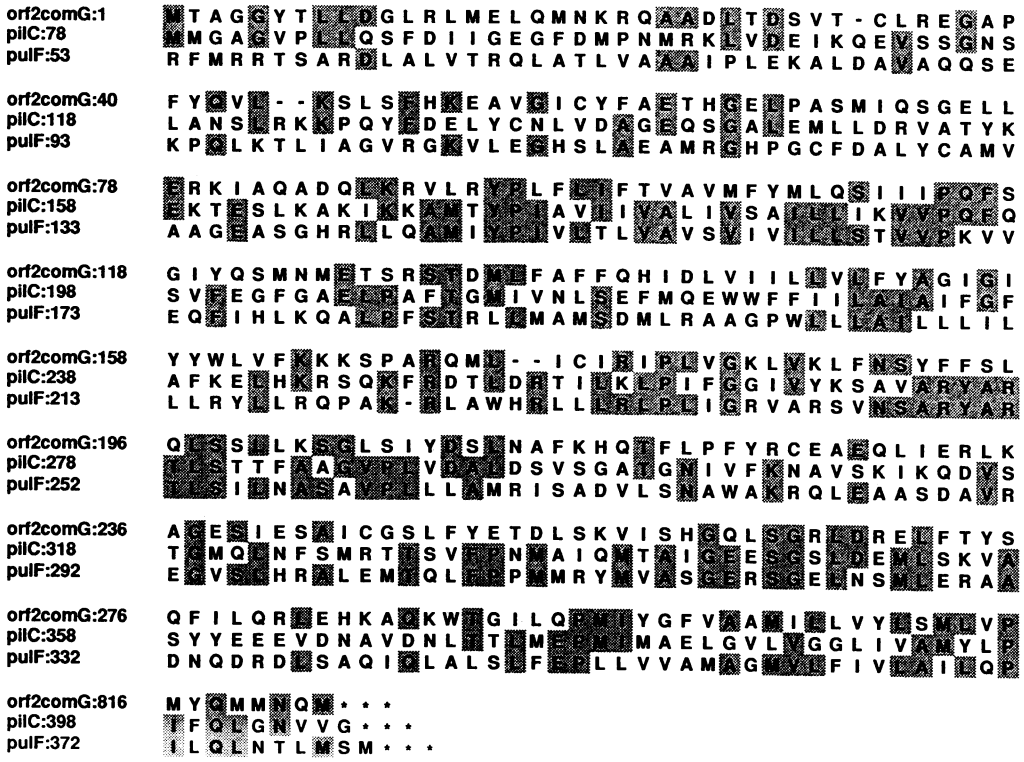

FIG. 3. Similarities of Comg ORF2 (1) to protein members of the pullulanase secretion (43) and pilin assembly (154) systems. Amino acid identities are indicated by shading whenever at least two of them occur at the same position. The residue number of the first amino acid on each line is indicated.

that this protein serves to couple DNA uptake to ATP hydrolysis. Alternatively, it may be required for the transport of another protein component of the competence apparatus, either as an energy coupler or as a molecular chaperone. For instance, PapJ is a likely nucleotide-binding protein that has been proposed to play a role as such a chaperone in the assembly of P pili of *E. coli* (214). Although it is not possible to rigorously classify any of the late competence

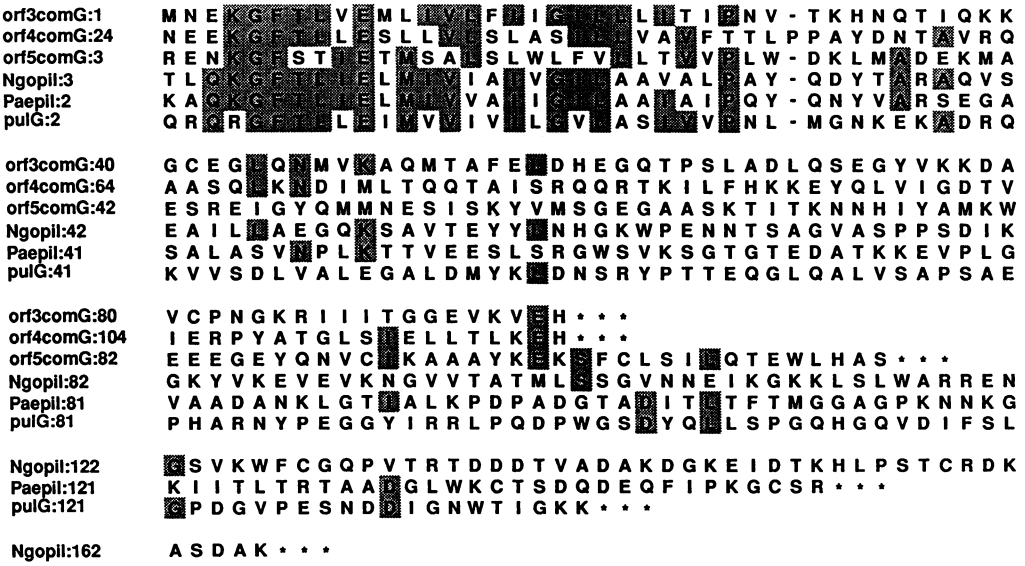

FIG. 4. Similarities of ComG ORF3 to ComG ORF5 (1) to protein members of the pullulanase secretion system (169) and to type IV pilins from *N. gonorrhoeae* (132) and *P. aeruginosa* (182). Amino acid identities are indicated by shading whenever at least three of them occur at the same position, except for the last few lines, for which two identities are indicated. The residue number of the first amino acid on each line is indicated. The type IV pilins are processed by cleavage between the conserved Gly (residues 6 and 7) and Phe (residues 7 and 8).

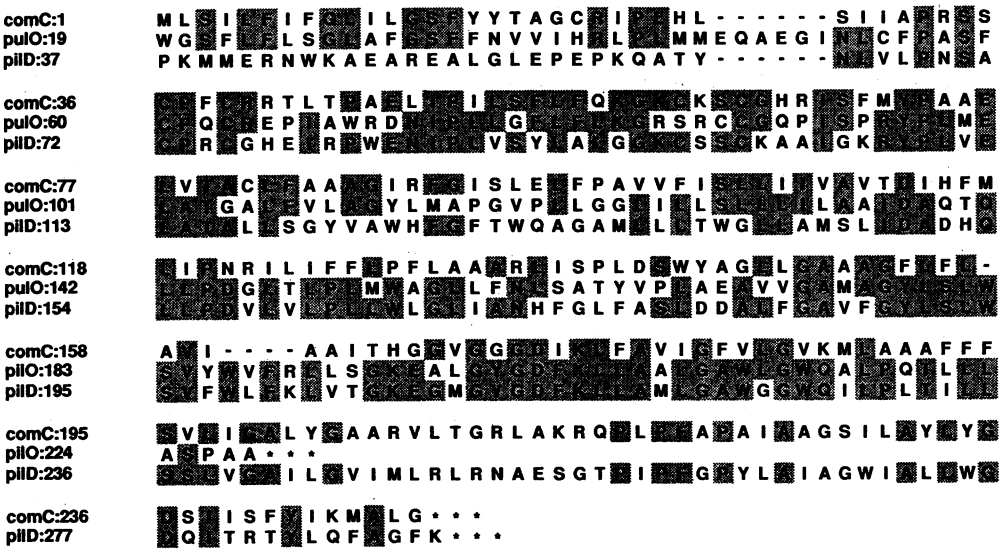

FIG. 5. Similarities of ComC (137) to protein members of the pullulanase secretion (162) and pilin assembly (154) systems. Amino acid identities are indicated by shading whenever at least two of them occur at the same position. The residue number of the first amino acid on each line is indicated.

products as either morphogenetic or directly involved in the uptake of DNA, some reasonable inferences can be presented.

The product of *comE* ORF3 appears to be required specifically for DNA uptake as opposed to binding (see above) and may be directly involved in this process. The ComG ORF3 product, which resembles the structural protein pilin, may play a direct role as part of the competence apparatus.

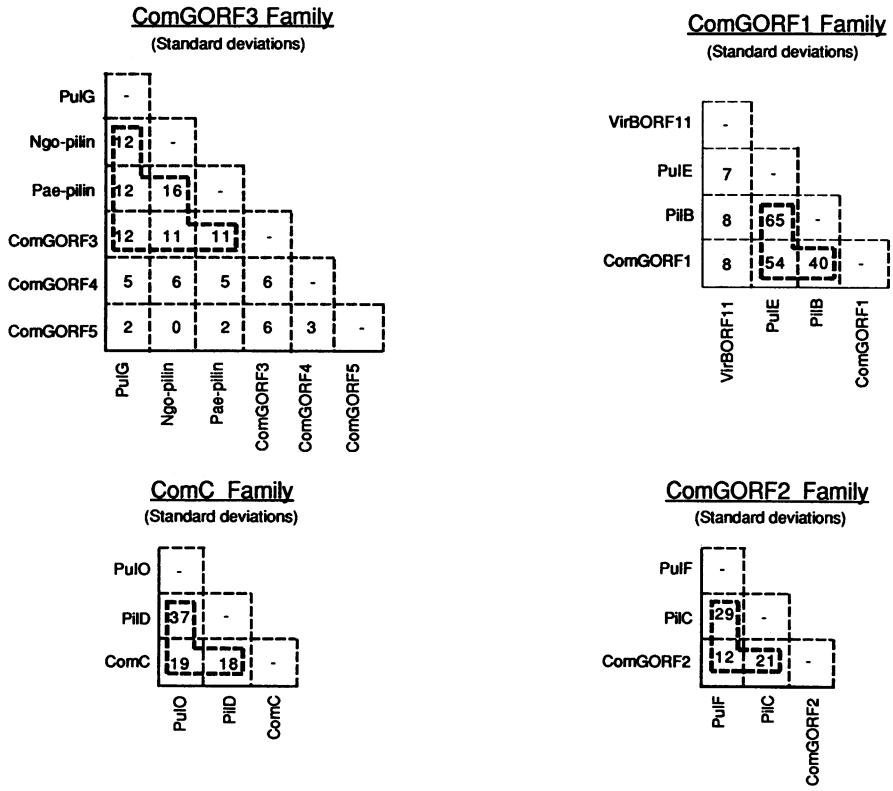

FIG. 6. Numerical representation of similarities of *com* products to protein members of the pullulanase secretion, pilin assembly and synthesis, and *A. tumefaciens* Ti plasmid systems. The numbers are Z-values obtained by using the RDF program (118) with 100 randomizations of the Com amino acid sequences. Z-values in excess of 10 standard deviations are regarded as highly significant, and values of 5 to 10 are regarded as possibly significant.

This is consistent with the evidence reviewed above suggesting a specific structural organization of ComG ORF3 in competent cells (22). It is possible that the products of *comG* ORF4 to ORF7, which, as noted above, resemble ComG ORF3 in some respects, also play structural roles in the competence apparatus. The related N-terminal hydrophobic sequences of ComG ORF3 to ComG ORF5 resemble the conserved N-terminal sequences of the *N*-methylphenylalanine pilins, which have been implicated in the protein-protein interactions involved in the assembly of pili (232). We therefore propose that the five small ComG proteins assemble to form part of a cell surface-associated structure for the binding and uptake of transforming DNA (1, 22). Consistent with this idea is the overlap of the *comG* ORF3 to ORF6 noted above, which suggests that these products may be translationally coupled to ensure their synthesis in appropriate relative amounts. It is noteworthy that the published model for the structure of *Pseudomonas* pili proposes a helical array of pilin subunits with fivefold symmetry and a central pore of 1.2 nm (71). A tubular structure may contribute to the formation of a channel for passage of transforming DNA through the cell wall and membrane. The roles of ComG ORF3 to ComG ORF7 in transformation do not seem to include the formation of pili, since *B. subtilis* does not form these structures.

On the other hand the ComG ORF1 and ComG ORF2 proteins and the ComC protein, which are similar to PilB, PilC and PilD, respectively, may very well be morphogenetic, since the last three products are required for the assembly of pili, but not for the synthesis of pilin or for its transport into either the inner or outer membranes of *Pseudomonas* cells (154). PilD is needed for the proteolytic processing of *Pseudomonas* pilin (154). It will be interesting to determine whether ComG ORF3 to ComG ORF5 are processed by a ComC-dependent mechanism. Another candidate for a role in assembly rather than for direct involvement in the DNA uptake process is the *div-341* (*secA*) locus described above (176, 178, 180, 181). One interpretation of the requirement for this *secA* homolog is that the export of one or more late competence proteins requires the operation of the major (*sec*) secretion pathway, in addition to more specialized gene products.

Much of the preceding discussion is speculative. We desperately need experiments designed to describe the postulated competence machinery and to enumerate its protein components.

Polyhydroxybutyrate

Some interesting data have been interpreted as suggesting a role for polyhydroxybutyrate (PHB) in genetic competence in both naturally competent organisms and in *E. coli* (166–168). In *B. subtilis*, *H. influenzae*, and *Azotobacter vinelandii*, three naturally competent organisms, PHB was detected in the membrane fraction (167). In these systems the amount of PHB was correlated with the level of competence. In *B. subtilis* the competent and noncompetent cell fractions were separated in Renografin and the amount of PHB was found to be three- to fourfold higher in the competent cells. Studies of phase transitions by using a hydrophobic fluorescent probe were interpreted as indicating the presence in the membranes of a labile organized gel structure containing PHB. Even in *E. coli*, when transformability was induced by the use of Ca^{2+} at low temperature, the de novo synthesis of PHB and its incorporation into the membrane occurred, and this was found to correlate with the development of competence (166). A PHB complex isolated from the *E. coli* membrane also contained polyphosphate and calcium and could be incorporated into liposomes (168). Molecular modeling was used to suggest that the PHB forms a channel with a lipophilic exterior and with its interior occupied by a Ca polyphosphate helix. It was proposed that this structure makes up a channel for the permeation of DNA. This work deserves to be extended. Although intriguing, it is based largely on correlations. It would be of interest to determine the transformability of mutants deficient in PHB synthesis.

Competence Gene Products in Other Systems

Lacks et al. have identified a membrane-localized *S. pneumoniae* DNase that is required for the uptake of transforming DNA and its conversion to a single-stranded form (111, 112). This nuclease functions endonucleolytically when in solution, but apparently as an exonuclease in situ. This presumably reflects its topology in the membrane and/or its relationship to other proteins. When isolated from the membrane by gentle detergent extraction, the nuclease was obtained as part of a large complex (173). The *endA* gene, which encodes the nuclease, has been cloned and sequenced (163).

Several transformation-deficient mutations in *H. influenzae* have been described. Particularly interesting is the *rec-2* mutation, which prevents the translocation of DNA from the transformasome to the cytoplasm (14, 130). To generate a more extensive collection of transformation-deficient mutants of *H. influenzae*, a plasmid library of *H. influenzae* DNA was mutagenized by using mini-Tn*10kan* and the disrupted DNA segments were returned to *H. influenzae* by homologous replacement recombination (219). Twenty-four distinct mutations were localized to 10 *Pst*I fragments, and their phenotypic properties were examined. Several of the mutations conferred a loss of DNA binding and may have been due to lesions in either regulatory or late competence genes. Several mutants appeared to take up substantial amounts of DNA, but were deficient in transformation with chromosomal DNA. Most of the latter class exhibited normal or near-normal UV resistance and were probably not *rec* mutants in the classical sense. They are candidates for late *com* mutants.

Finally, as mentioned above, pilin has been implicated in genetic transformation in *Moraxella* and *Neisseria* spp. (20, 184, 198). Perhaps this protein plays a similar structural role in the transformation of both gram-negative and gram-positive bacteria.

REGULATION OF COMPETENCE

Three Modes of Competence Regulation in *B. subtilis*

As noted briefly above, competence in *B. subtilis* is subject to three types of control. In the usual glucose minimal salts-based competence medium, supplemented with amino acids, competence develops postexponentially. The time of transition from exponential growth is defined as T_0. Competence is maximal about 2 h later, at T_2 (3). This reflects the first mode of competence regulation; growth stage dependence. Transformability develops to extremely low levels in complex media. When glucose is substituted by glycerol in the usual competence medium or when the glucose medium is supplemented with glutamine, the expression of competence is also low (3). These observations indicate a second aspect of control: dependence on medium.

Finally, only about 10% of the cells in the culture achieve competence, and these cells are metabolically and morphologically distinct (47, 87, 91, 129, 152, 188). Thus, an event occurs which results in the development of at least two cell types, with expression of competence and late competence genes (3) in one of them.

B. subtilis SOS-Like System

E. coli responds to the stress of DNA damage by inducing the expression of a series of genes, including several devoted to repair of this damage (230). The signal for induction is thought to involve single-stranded DNA which interacts with the RecA protein (in *E. coli*), resulting in cleavage of the LexA repressor and the consequent derepression of the SOS regulon. A similar pathway appears to exist in *B. subtilis*; the RecE gene product is the homolog of RecA (45, 122). In addition to this mode of regulation, however, the *B. subtilis* SOS-like system is at least partially derepressed as cells reach competence (121, 245). This was shown by using transcriptional fusions of *lacZ* to a set of genes that can be induced by DNA damage (*din* genes). These genes are also induced in the competent state, and specifically in the competent-cell subpopulation. The competence-specific induction of the *din* genes, like the damage induction, requires a functional RecE product. Also as part of the competence response, the *recE* gene itself is strongly derepressed (124). However, the competence induction pathway for *recE* expression, unlike the DNA damage-inducible pathway, does not depend on the presence of a functional RecE gene product. Also, unlike the DNA damage-induced pathway, the competence-dependent induction of the SOS-like response depends on the *spo0A* and *spo0H* gene products (245), as does competence itself (see below). It thus appears that the induction of the SOS response by DNA damage differs from that during the development of competence and that the latter regulatory mode overlaps with regulation of the competence regulon. It would be of great interest to determine whether this form of SOS-like induction depends on all the competence regulatory genes described below. The implications of the competence-linked induction of DNA repair genes will be discussed further later in this review, as well as a possible mechanism for the competence-linked induction of the SOS system.

Regulatory Genes

Most of the investigations of the control mechanisms involved in competence have been concerned with identifying and characterizing regulatory genes. More than a dozen such genetic elements have been identified, and their relationships have been explored. As a result, a good deal is known about the regulatory pathways involved, although it is clear that much more remains to be understood. What follows is a catalog of these loci and a brief description of the phenotypes associated with each. This initial description will make little attempt to explore the mechanisms of action of these genes or how they interact to make up a regulatory network; those tasks will be undertaken later in this review. The locations of the regulatory loci on the *B. subtilis* genetic map are shown in Fig. 1.

Mutations of each of the regulatory loci have been tested for their effects on expression of other regulatory genes and on the late competence genes. In most cases this has been done by using *lacZ* transcriptional fusions to the target genes, and in some cases it has been done by direct measurement of steady state mRNA concentration as well. These experiments have generally revealed that the mutations have profound effects on transcription of the late genes, but that they have no detectable effects on one another. Exceptions to this generalization will be noted below.

Competence Regulatory Loci

spo0H. Mutants with null mutations in the *spo0H* locus are asporogenic and are about 20-fold depressed in transformation efficiency (3, 179). These mutants exhibit decreased expression of all of the known late *com* genes, as shown by using transcriptional fusions to *lacZ* (3). The *spo0H* locus encodes σ^H, an alternative σ factor that is required to read essential sporulation genes (62). The Com phenotype of *spo0H* null mutants implies that at least one essential *com* gene is transcribed by $E\sigma^H$; the identity of the $E\sigma^H$ target(s) is unknown.

abrB. Mutants with null mutations in the *abrB* locus are depressed about 20- to 50-fold in competence and fail to express late competence genes at the wild-type level (62). *abrB* encodes a so-called transition state regulator that prevents the inappropriate expression of various genes that are normally expressed postexponentially and acts as a DNA-binding protein (157, 206). The positive effect of *abrB* on the expression of *com* genes may be due to a positive effect of AbrB on expression of at least one essential gene in the competence pathway or to repression of a negatively acting competence control element. The identity of this target gene is not known.

spo0A. Mutants with null mutations in the *spo0A* locus exhibit a highly pleiotropic phenotype, reflecting the central role played the the Spo0A protein in a variety of forms of postexponential expression. *spo0A* null mutants are depressed 500- to 1,000-fold in competence, are severely depressed in the expression of late competence genes (3, 179, 201), and are completely deficient in sporulation (97). One of the functions of Spo0A is to down-regulate the expression of *abrB* on the level of transcription; *spo0A* mutants overexpress AbrB (157). A *spo0A abrB* double null mutant exhibits the same level of competence as does an *abrB* null mutant alone (3). It thus appears that the sole role for Spo0A in the competence system is to prevent the overexpression of AbrB, which must be capable of playing a negative as well as a positive role in the development of competence. This conclusion has been buttressed by further evidence, to be presented below in the section on *mec* mutants.

The amino acid sequence of the N-terminal domain of Spo0A is similar to that of a class of proteins known as response regulators (68, 204). These are in most cases transcriptionally active molecules that can be phosphorylated by cognate proteins collectively referred to as histidine kinases. Phosphorylation of the conserved N-terminal domain of a response regulator is thought to cause a conformational change that alters the activity of the C-terminal, DNA-binding domain. Phosphorylation of Spo0A increases its binding affinity for a sequence upstream from the *abrB* promoter about 20-fold (220), presumably facilitating its action as a repressor. One protein with homology to the histidine kinases that can phosphorylate Spo0A (via a phosphorylation cascade) is the SpoIIJ protein (8, 24, 155).

This and other evidence clearly indicate that phosphorylation of Spo0A is a critical event in the initiation of sporulation. Is phosphorylation a signaling event for compe-

tence? Null mutations of *spoIIJ* have no obvious effect on competence (234, 239). It has recently been shown that spoIIJ acts by phosphorylating the Spo0F protein and that the latter transfers the phosphate to Spo0B, which in turn phosphorylates Spo0A (24). Null mutations in the *spo0F* or *spo0B* gene, like those in *spoIIJ*, have no measurable effects on competence (3). It is tempting, therefore, to conclude that signaling for the development of competence does not proceed via the phosphorylation of Spo0A. Since, as noted above, AbrB plays both positive and negative roles in competence, perhaps the affinity of the unphosphorylated Spo0A protein for the upstream *abrB* site is sufficient to prevent excessive expression of *abrB*, but not so great as to drive expression below the required level. However, it is likely that other histidine kinases can phosphorylate Spo0A, by what is often referred to as crosstalk. Perhaps one or more of these are required to activate Spo0A for purposes of competence. It is also possible that the increase in *spo0A* expression that occurs at T_0 is involved in signaling the onset of competence. However, this increase is glucose repressed (244) and may not occur in competence medium. Further light can be shed on the role of Spo0A by the study of *spo0A* mutations. Green and Youngman (79) have altered the conserved aspartate residue of Spo0A that is thought to be the site of phosphorylation. This change completely abolishes sporulation but results in an intermediate level of competence, about 10- to 100-fold higher than that of a *spo0A* null mutant. This would imply at least that phosphorylation of Spo0A is not absolutely required for the development of competence. The partial competence deficiency may be due to a potentiating effect of phosphorylation or to interference of the mutation with some aspect of Spo0A function other than phosphorylation. Green and Youngman have combined this mutation with a second alteration in Spo0A: a deletion in the protein that appears to render sporulation independent of Spo0A phosphorylation. This combination of mutations restores a nearly wild-type level of competence. This finding is consistent with the first notion stated above, i.e., that phosphorylation of Spo0A by a histidine kinase other than SpoIIJ plays a potentiating role in competence. Perhaps nonphosphorylated Spo0A can partially control *abrB* expression, but not sufficiently to permit full expression of competence. Phosphorylation of Spo0A (by a SpoIIJ-independent pathway) would result in an optimal concentration of AbrB and hence in the maximal expression of competence. This conclusion is tentative, and other interpretations of the data are possible. We must conclude that the role of Spo0A in competence signaling remains uncertain.

spo0K. The only known mutation in the *spo0K* locus, until recently, was *spo0K141* (159). This mutation causes an oligosporogenic phenotype and decreases competence several hundred-fold (179). It is required for the expression of the late competence genes (171). Recently, two laboratories have completed the sequence of *spo0K* and shown it to consist of five ORFs, all of which strongly resemble the similarly organized oligopeptide permease (*opp*) operons of *E. coli* and *Salmonella typhimurium* (156, 174). This permease is related to several of the bacterial periplasmic permease systems (6). The promoter-proximal ORF encodes a probable peptide-binding protein which is periplasmically located in the gram-negative bacteria and is probably covalently anchored to the exterior face of the gram-positive cell membrane by fatty acid acylation (156). The second and third ORFs encode integral membrane proteins, and the fourth and fifth encode hydrophilic proteins, which most likely possess ATPase activity and probably directly couple ATP hydrolysis to transport (6). At least the fifth ORF of *spo0K* (ORFE) is required for competence (174), but its disruption has only a minimal effect on sporulation and on transport of at least one peptide (156, 174).

sin. Null mutations in the *sin* locus depress competence several hundred-fold (75) and depress the expression of late competence genes (85). *sin* was originally identified as an ORF that depresses sporulation when overexpressed (75). It is known to encode a DNA-binding protein which exerts negative effects on the expression of several genes involved in sporulation and also on the expression of various degradative enzymes that normally exhibit increased postexponential expression in *B. subtilis* (76, 126). It is not known whether Sin acts positively or negatively during the development of competence; like the AbrB protein, it may activate expression of an essential competence gene, or it may repress expression of a repressor. *sin* is apparently transcribed by an $E\sigma^A$ promoter and is expressed throughout growth at an approximately constant rate (74).

degU. Mutants with null mutations in the *degU* locus are nearly completely deficient in transformation (213) and do not express the known late competence genes (171). *degU* was originally identified because it is needed for the expression of certain degradative enzymes, mostly extracellular, that are normally synthesized at an enhanced rate postexponentially (12, 95, 106, 107). Upstream from *degU* is the *degS* gene, which is also required for degradative enzyme synthesis, but not for competence (95, 106, 144, 145, 171). The DegS and DegU products are histidine kinase and response regulator proteins, respectively, as suggested originally from their amino acid sequences and confirmed by biochemical tests on the purified proteins (38, 95, 106, 147).

In addition to null mutations in these genes, point mutations that result in hyperproduction of the degradative enzymes have been described and characterized on the sequence level (144, 202). These so-called $degS^h$ and $degU^h$ mutations profoundly interfere with the development of competence (144, 202) and with the expression of the known late competence genes (171). They have usually been interpreted as alterations in the proteins that result in enhanced activation (presumably via phosphoryl group transfer) of DegU by DegS. Inactivation of *degS* restores competence to a $degU^h$ mutant (144). These observations have led to the suggestion that the nonactivated (presumably unphosphorylated) form of DegU is specifically required for competence, whereas the activated form is needed for the expression of degradative enzymes. Strong support for this hypothesis is derived from the *degU146* mutation, which converts a conserved aspartate residue, which is the likely target for phosphorylation, to an asparagine residue (144). In addition, preliminary experiments with crude extracts suggest that the *degU146* allele encodes a protein that is no longer phosphorylated by the DegS protein kinase (38). This mutation decreases degradative enzyme production, but has little if any effect on competence. It is worth nothing that *degS* and *degU* mutations are quite pleiotropic. For instance, the $degS^h$ and $degU^h$ mutants also exhibit sporulation in the presence of glucose (the CRS phenotype) and a lack of motility (12, 107).

***comA* and *comP*.** Null mutations in either the *comA* or *comP* loci reduce competence about 400-fold and interfere with late *com* gene expression (85, 150, 237, 239). *comA* was independently discovered as a locus needed for the expression of the peptide antibiotic surfactin and was named *srfB* (150). *comA* and *comP* are adjacent genes (Fig. 7). On the

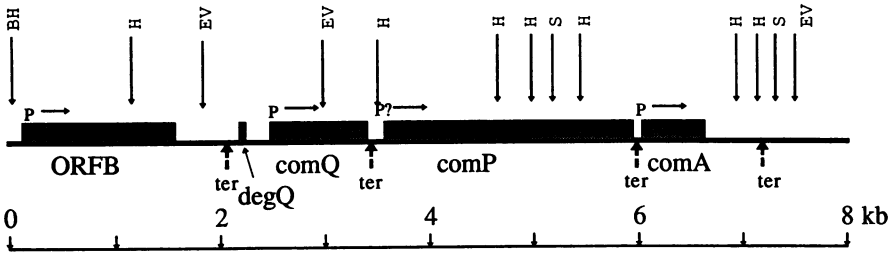

FIG. 7. Genetic and physical map of the *comA-comP* region of the *B. subtilis* chromosome. The *ORFB* determinant was previously called *comB* but is now known not to be a competence gene (see the text). The DNA sequence of the entire 8-kb region has been determined. References are given in the text. Abbreviations: BH, *Bam*HI; H, *Hin*dIII; EV, *Eco*RV; S, *Sst*I. The positions of promoters (P) and terminators (ter) are given. The *degQ* determinant encodes a 46-amino-acid peptide and is not drawn to scale.

basis of their predicted amino acid sequences, they appear to encode response regulator and histidine kinase proteins, respectively. More recently, ComA has been purified from an overproducing *E. coli* clone and shown to accept phosphate from the *E. coli* NRII (NtrB) protein, a histidine kinase, in vitro (215). In addition to the resemblance of the ComA N-terminal domain to other response regulators, the C-terminal domain of ComA exhibits similarity to the C-terminal sequences of several known DNA-binding proteins and to DegU (Fig. 8). These similarities, plus the effects of *comA* and *degU* mutations on expression of a number of genes, imply that the two response regulators function as transcription factors. Their targets are unknown. The predicted ComP amino acid sequence exhibits eight regions of pronounced hydrophobicity near the N terminus, and a specific model has been proposed for the organization of ComP as a polytopic membrane protein (239). Thus, although the signal to which ComP responds is not known (see below), this signal is presumably extracellular or mediated via other membrane components.

lacZ fusions to *comA* have shown that the expression of this gene increases about twofold during growth, reaching a maximum after T_0 (239). High-resolution primer extension has revealed a start site upstream from *comA*, near a probable $E\sigma^A$ promoter. However, disruption of *comP* prevents the twofold increase in *comA* expression, suggesting that this increase is due to readthrough from a promoter upstream of *comP*. The predicted readthrough transcript was detected by S1 nuclease analysis. The significance of this is not known, nor is the location of the promoter responsible for transcription of *comP*.

comA and *comP* are also required for the starvation-induced expression of certain other loci, such as *degQ* (145) and *gsiA* (146), neither of which is required for competence.

comQ. Upstream from *comP* (Fig. 7) is an ORF, identified by sequencing, that encodes a predicted protein of 299 amino

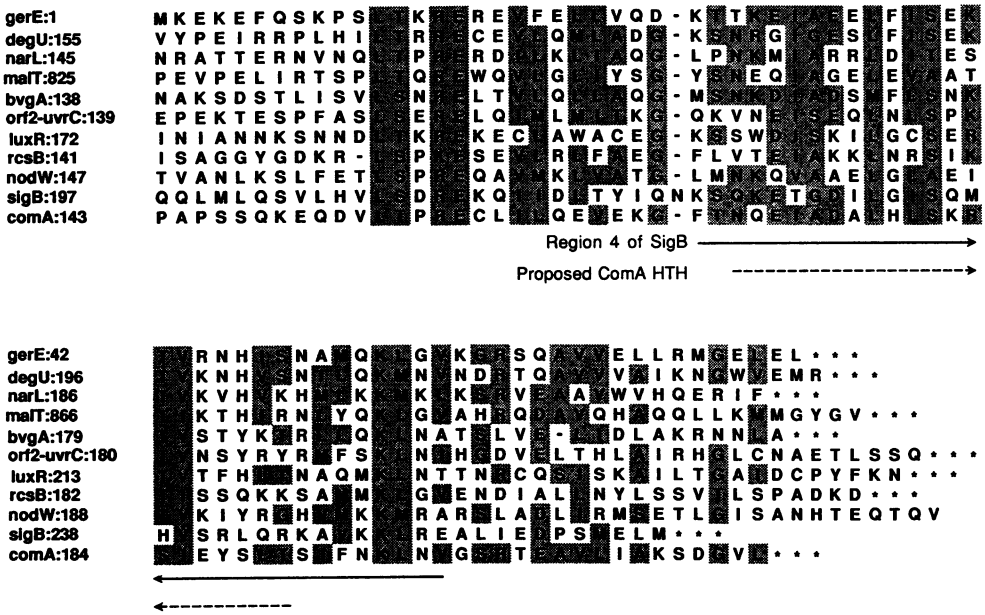

FIG. 8. Similarities of the C-terminal 71-amino-acid residues of ComA (237) to C-terminal moieties of other suspected and known transcription factors. Amino acid similarities are indicated by shading when six or more of them occur at the same position. For this purpose, the amino acids are grouped as follows: M, L, I, V; H, K, R; D, E; Q, N; A, G, S, T; F, W, Y; P; C. The locations of the region 4 helix-turn-helix (93) of SigB (63), based on comparison of sigma factors, and of the proposed helix-turn-helix of ComA are indicated. The sources of the additional protein sequences are GerE (37), DegU (95, 106, 213), NarL (153), MalT (34), BvgA (9), ORF2-UvrC (185), LuxR (44), RcsB (205), and NodW (78).

acids (238). Disruption of this ORF by insertion of antibiotic resistance cassettes resulted in competence deficiency, to the same residual level as that seen in *comA* or *comP* null mutants, and also eliminated expression of the known late competence genes. These disruptions were not polar on *comA* fusions to *lacZ*, and they could be fully complemented for competence by a single copy of the ORF integrated at an ectopic chromosomal site. This ORF has been named *comQ* and appears to be a new competence locus.

High-resolution primer extension analysis identified a transcriptional start site for *comQ* appropriately positioned downstream from a sequence suggestive of a vegetative $E\sigma^A$ promoter (238). Downstream from *comQ* and between it and *comP* is a palindrome that resembles a rho factor-independent transcriptional terminator, and S1 nuclease mapping confirmed the existence of a termination site near this element. As judged by primer extension analysis and by study of a translational fusion of *comQ* to *lacZ*, the expression of this gene in competence medium was relatively high 2 h before T_0 (T_{-2}) and declined to an unmeasurable level after about T_{-1}.

comB (ORFB). *comB* was identified by insertion of Tn*917lacZ* to yield the *comB138* mutation, which was linked to *comA*. The mutant strain carrying this lesion was about 400-fold reduced in competence and failed to express the late competence genes (85, 237). In every respect its phenotype was identical to those of the *comA* and *comP* null mutants. Sequencing revealed that the transposon had inserted in an ORF with a predicted amino acid sequence that did not resemble that of any protein in the data base. However, doubts about the validity of these findings arose when it was noted that constructs with antibiotic resistance cassettes inserted in *comB* were normal in competence and that the cloned *comB* ORF failed to complement the *comB138* insertion (234). Subsequent Southern blot hybridization experiments with chromosomal DNA from the *comB138* mutant revealed that the transposon insertion was accompanied by a large chromosomal deletion that removed more than 7 kb of DNA, including the *comP*, *comA*, and *comQ* loci. It appears that the *comB* ORF is not required for competence, and will provisionally rename it *ORFB* (Fig. 7). It is worth nothing that this is the second instance of a large chromosomal deletion associated with a Tn*917* insertion and in the same region of the chromosome (149).

srfA. A Tn*917lacZ* insertion (*csh-293*) was identified that resulted in both competence deficiency (about 400-fold) and a subtle oligosporogenic phenotype (100). This mutation also depressed expression of the late competence genes (171). A locus was independently identified that was needed for competence and was called *comL* (225). Finally, in a search for regulatory genes required for the postexponential expression of the extracellular peptide antibiotic surfactin, a locus known as *srfA* was characterized (149). It has recently been shown that *csh-293*, *comL*, and *srfA* mutations identify the same locus, and the name *srfA* is retained in this review (148). This gene consists of a large operon (greater than 20 kb) of which only the promoter-proximal portion (about 15 kb) appears to be required for the development of competence (148, 224, 225). This portion of the operon encodes two ORFs with predicted amino acid sequences similar to those of tyrocidine synthetases I and II (127, 135, 148, 151, 223, 225, 233). These enzymes are required for the nonribosomal assembly of peptide antibiotics via the stepwise amino acyl activation and polymerization of amino acids (119). It has been pointed out that the involvement of this operon is particularly interesting in view of the possible involvement of a soluble peptide factor in signaling for competence development (148).

srfA is expressed at a low level throughout growth, but its expression increases sharply at T_0 (100, 149). This is therefore an interesting example of a regulatory locus required for the expression of late competence genes, which itself is postexponentially expressed. This kinetic pattern suggests that *srfA* may play an intermediate role in the competence-signaling pathway, a prediction that is supported by further evidence. The increase in *srfA* expression noted in competence medium at T_0 is dependent on expression of *comQ*, *comP*, *comA*, and *spo0K* and is partially dependent on *spo0A* and *spo0H* (89). The partial dependence of *srfA* expression on *spo0A* and *spo0H* has been noted previously (100, 149). *srfA* expression is not measurably dependent on *abrB*, *sin*, or *degU*. This suggests a hierarchial arrangement of gene products in the competence-signaling pathway, a point which will be discussed further below. However, we may ask at this point whether the increase in *srfA* expression noted at T_0 is important for competence. It is possible that it is not and that the dependence of the increased *srfA* expression on several competence regulatory genes has nothing to do with the dependence of competence on these genes. M. Nakano and P. Zuber kindly provided us with a strain in which the expression of *srfA* had been placed under control of the regulatable *Pspac* promoter (151). In this strain competence develops only in the presence of an inducer of P *spac*, such as isopropyl-β-D-thiogalactoside (IPTG). The level of competence achieved under these conditions is lower than in the wild-type strain, presumably because the *Pspac* promoter is weaker than the normal *srfA* promoter (151). It is interesting that this level of competence is expressed throughout growth, when the strain is grown continuously in the presence of IPTG (89). This observation suggests that the growth stage-specific signal may normally act prior to the increased expression of *srfA* at T_0. When the *srfA-Pspac* construct is combined with several of the null mutations in other competence regulatory genes, the level of transformability reached is extremely low in the absence of IPTG, as it is in the original *Pspac-srfA* strain. In the *Pspac-srfA* strains carrying *comA*, *comQ*, or *spo0K* mutations, growth in the presence of inducer restores competence to the level achieved in the construct carrying the *srfA-Pspac* construct alone in the presence of IPTG (89). When *srfA* is combined with *degU* or *abrB* mutations, the addition of IPTG does not elevate the level of competence. A *Pspac-srfA sin* strain was not tested. In other words, mutations in *comA*, *comQ*, and *spo0K* seem to be bypassed for their effects on competence by expression of *srfA* under *Pspac* control. The *spo0K* mutation used for these experiments was a disruption in the fifth ORF (ORFE) of the *spo0K* operon. Nakano and Zuber have also used their *Pspac-srfA* strain to demonstrate the bypass of *comA* (151).

In the *Pspac-srfA* strains, competence is expressed throughout growth in the presence of IPTG (89). These observations present a consistent picture in which some of the regulatory gene products (*comA*, *comP*, *comQ*, and *spo0K [ORFE]*) are needed only to turn on *srfA* expression in response to environmental signals, whereas others (*degU*, *abrB*, and *sin*) are not and act either later than or in parallel to *srfA* for competence signaling.

Although the expression of *srfA* is not dependent on *degU*, it is markedly reduced in a strain carrying the *degU32* allele (89). This result suggests that the inhibitory effect of the DegU32 mutant protein is exerted upstream of *srfA* and provides evidence against the notion that the low compe-

tence imparted by this *degU*^{h} mutation is due simply to a depletion of unphosphorylated DegU.

comK. The *comK* locus was identified by Van Sinderen et al. (225). It has been shown to increase in its expression at T_0 and to be required for the expression of the late competence genes (225). In addition, *comK* expression is dependent on that of *comA* and *srfA* (*comL*). Van Sinderen and Venema (224) also found that although it was impossible to obtain high expression of ComK on a multicopy plasmid, moderate overexpression appeared to result in the elevated expression of competence during exponential growth. This suggests that the ComK product may act positively in turning on competence at T_0. The *comK* locus has been cloned and sequenced (224), and its predicted product bears no striking relationship to any known protein.

mec Loci

As noted above, the development of competence and the expression of late competence genes is poor in complex media and maximal in glucose minimal salts-based media. To further our understanding of the regulatory loci involved in the response to nutritional environment, mutants were isolated following ethyl methanesulfonate mutagenesis that were capable of expressing late competence genes (and competence itself) in complex media (61). The mutations responsible for this phenotype were called *mec* (medium-independent expression of competence). They were mapped to at least two loci, *mecA* and *mecB* (Fig. 1) (61). Although mutations in *mecA* and *mecB* permitted the development of competence and the expression of late competence genes in complex media, they exhibited enhanced expression at T_0, as in the case of the *mec*$^+$ strain. Thus, it might appear, as we originally inferred, that the growth stage-related regulation is independent of the *mec* mutations. However, the *mecA* and *mecB* mutants did express the late competence genes prior to T_0 at a rate greater than that of the wild type, so that interpretation of this point is somewhat ambiguous (61).

mecA is closely linked by transduction to the *spo0K* locus (171). Recently we have observed that it is also linked to *comK* and is located between these two loci (170). *mecB* is closely linked by transformation to rifampin, streptolydigin, and lipiarmycin resistance alleles (61), which were previously shown to be in or near the *rpoB* gene (117, 197). The latter encodes the β subunit of RNA polymerase. The close transformational linkage observed is consistent with the hypothesis that the two *mecB* mutations studied are alleles of *rpoB*. Other than these comments, I can say little about the nature of the *mecA* and *mecB* loci and of their respective mutations. For instance, it is not even possible to assert whether the mutations result in gain or loss of function.

The *mecA42*, *mecB31*, and *mecB23* mutations bypass all of the competence regulatory genes described above, with the exception of *spo0A*, which is partially bypassed (105, 171, 172). For example, the *mec* mutations permit the development of competence and the expression of late competence genes even in genetic backgrounds carrying inactivating mutations in the *comA*, *comP*, *degU*, *abrB*, *spo0K*, *srfA*, *comK*, and *sin* genes. Taken at face value, the suppression of all of these regulatory mutations implies that the *mec* products act later than the products of the other regulatory genes during competence development.

The partial suppression of the *spo0A* mutant would imply that the negative effect of AbrB (which, as described above, is overproduced in the *spo0A* background) must be exerted at multiple points, both before and after the point of action of *mec*. Some recent experiments are consistent with this interpretation of the relationship of *mecA* to *spo0A* and *abrB* (170). A *spo0A abrB* double null mutation was prepared in a *mecA42* background which also carried a *lacZ* fusion to *comG*. In this strain, essentially normal levels of *comG*-driven β-galactosidase were observed. This is as predicted, since the *abrB* mutation was able to suppress the effect of the *spo0A* mutation, and *mecA42* bypassed the effect of the *abrB* null mutation.

However, the simple interpretations of *mec* action given above, which appear plausible, must be regarded as tentative since it has not been proven that the *mec* products are normal components of the competence-signaling pathway. It is possible that the mutant *mec* products do not normally play a role but are able to intercede so as to bypass the need for the early regulatory products. For instance, *mecA* may encode a DNA-binding protein with low and fortuitous affinity for the promoter of a competence repressor. Overproduction of MecA (in the *mecA42* mutant) might then relieve negative regulation of competence, effectively bypassing the need for the other regulatory components. In spite of this caveat, I will perpetuate the more enjoyable interpretation first stated above.

The ability of the *mec* mutations to bypass the early gene requirements suggests that these loci are all required for a signaling process that converges at some later point. Since the *mec* mutations allow expression of competence and of late competence genes in complex media, in the absence of glucose, and in the presence of a high glutamine concentration (61) (all treatments that ordinarily repress competence), it appears that the signals involved are at least in part nutritional.

A further inference is that the early regulatory gene products, several of which are known or thought to be regulators of transcription (ComA, DegU, Sin, AbrB, Spo0A, and Spo0H), probably do not act directly on the late gene upstream sequences to modulate their expression. This follows from our preferred interpretation of the *mec* suppression data, since the Mec products appear to act later than the other regulatory products.

Finally, it appears that the role of the *mec* genes is not restricted to the competence system. Mueller and Sonenshein (146) have identified two genes (*gsiA* and *gsiC*) that are induced to express at elevated rates under conditions of glucose starvation. *gsiC* is required for the development of competence and for the full expression of sporulation and therefore constitutes a new, presumably regulatory *com* locus of unknown function. *gsiA*, on the other hand, is not required for competence, but is dependent for its expression on *comA* and *comP*. It is not dependent on *degU*, *degS*, *sin*, *srfA*, or *spo0K*, at least in complex medium. However, the *comA* and *comP* dependence of *gsiA* expression is bypassed by mutations in *mecA*, *mecB*, and *gsiC*. In addition to presenting additional complexities, these observations suggest a wider role for the *mec*, *comA*, and *comP* genes and imply that the set of *mec*-bypassable genes makes up a pathway with more than one branch. Additional support for this notion derives from unpublished results of Msadek and Kunst (143). These investigators have observed that the dependence of expression of levansucrase on *degU* is bypassed by mutations in *mecA* and *mecB*. This implies that DegU may not interact directly with targets near the degradative enzyme promoters and again suggests that the *mec*-bypassable pathway contains several branches.

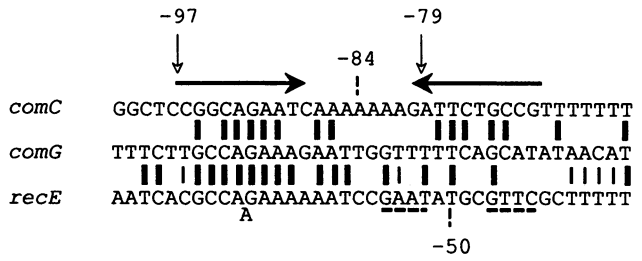

FIG. 9. Comparison of proposed CTF binding sites upstream from *comC*, *comG*, and *recE*. Identities between adjacent sequences are indicated by heavy vertical bars, and additional identities between *comC* and *recE* are indicated by light bars. The arrows indicate the partial dyad symmetry noted in the *comC* sequence. A deletion extending from the left to position −97 has no effect on the expression of *comC*, whereas a deletion extending to −79 reduces *comC* expression (138). The center of dyad symmetry of the *comC* element is at −84, and that of *comG* is at −81 (not shown). The dotted lines indicate a *din* consensus site upstream from the *recE* promoter (29). This site is centered at −50.

What Turns On Late Gene Expression?

The question of what turns on late gene expression is raised acutely by the inference that the early regulatory products do not act directly on the late competence gene promoters. A clue to the answer was provided by observations that DNA fragments derived from the promoter regions of the *comG* and *comC* transcription units were capable of repressing the expression of the known late competence genes when these fragments were present in multicopy (1, 138). This suggested that a limiting transcription factor was being titrated by the promoter fragments, and that the action of this factor was general, since the *comC* fragment repressed *comG* expression and vice versa.

The upstream sequences required for expression of *comC* were explored by deletion analysis (138). It was found that sequences located between −97 and −79 were required for this expression. Centered as −84 was a partial palindrome. Upstream from the *comG* promoter, a similar sequence containing a partial palindrome with a center of symmetry at −81 was noted (Fig. 9). When various *comC* fragments were placed in multicopy in a wild-type *B. subtilis* strain, the intact fragment depressed competence as expected, as did the fragment deleted to −97 from upstream. However, the −79 deletion which removed the sequence found to be essential for *comC* expression did not inhibit transformability. It therefore appears that these sequences define a binding site for a factor required for the expression of at least some of the late competence genes. We have named this factor, defined by these in vivo experiments, competence transcription factor (CTF).

Gel shift experiments have served to identify an activity that has many of the properties expected of CTF (138). This activity binds specifically to a fragment carrying the sequences defined as responsible for CTF binding, on the basis of the in vivo experiments described above. It binds to the fragment that is deleted up to −97, but not to the fragment deleted to −79 (105). The binding activity is detectable only in extracts of the wild-type strain prepared after T_0 and grown in competence medium (138). It is detectable in *mecA42* extracts prepared from complex medium as well. No expression of the binding activity is detectable in mutants with null mutations in *comA*, *degU*, *spo0A*, *srfA*, *spo0H*, *comK*, or *spo0K* (105, 138). The *comA* dependence of the gel shift activity was bypassed by the *mecA42* mutation. This in vitro activity therefore appears to be under competence control and is likely, although not proven, to represent the CTF activity as defined by in vivo experiments.

CTF may be required for expression of the late competence genes that have been tested: *comG*, *comC*, and *comDE*. This is inferred from the in vivo titration experiments (138). The simplest interpretation is that CTF is not only necessary but sufficient for the expression of competence. If this hypothesis is correct, the entire regulatory cascade may be needed only to express or activate CTF. This possibility will be testable when *ctf* is cloned and characterized.

At this time it is appropriate to return to the question of the competence-linked induction of the SOS regulon. Cheo et al. (29) have sequenced the promoter regions of several DNA damage-inducible genes (*din*) from *B. subtilis* and have identified sequences necessary for damage-inducible regulation of the *din* genes. Within these essential regions were conserved sequences that were interpreted as representing recognition elements for the binding of a presumed SOS-specific regulatory protein, possibly analogous to the LexA product of *E. coli*. Several of the *din* gene regulatory regions contained more than one of these putative recognition elements. Upstream from one of the three copies of this element associated with the *recE* gene is a sequence that resembles the putative CTF-binding element (Fig. 9). The SOS recognition sequence is centered at position −50, and the CTF recognition element is centered at −57. This suggests a model for dual regulation of the *recE* determinant in which binding of a LexA-like protein represses expression, and this repression can be relieved in two ways: by DNA damage and RecE-dependent cleavage of LexA, and by binding of CTF and displacement of LexA. The *din* gene promoter regions do not seem to possess CTF-binding sequences. This is consistent with the finding of Love et al. (121) that whereas the induction of *recE* in the competent state does not require a functional RecE product, the competence-specific induction of the *din* genes does require this gene product. Presumably the *din* gene induction is a secondary consequence of *recE* induction and/or activation.

Regulatory Cascade

The information outlined above suggests a pathway of information flow involved in signaling the onset of competence. Various forms of this pathway have been suggested before and differ slightly from one another (52, 60, 171, 225). Hopefully, these differences reflect the evolution of our understanding. They also serve to emphasize the provisional and incomplete nature of much of our knowledge and the correspondingly provisional nature of our schematic attempts to integrate this information.

The various genetic dependencies suggest a hierarchical ordering of the known regulatory gene products (Fig. 10). The dependence of *srfA* expression on several gene products, but not on others, suggests either that some of the regulatory proteins act later than *srfA* (Fig. 10A) or that they lie on a separate branch of the pathway (Fig. 10B). The dependence of *comK* expression on *srfA* and the dependence of late gene expression on *comK* suggest that the *comK* product acts later than *srfA* and before the turn-on of late genes (225). The bypass of *comK* by *mecA42* (105) suggests the order for these elements shown in Fig. 10. The positions of *degU*, *sin*, and *abrB* in Fig. 10A remain ambiguous, since

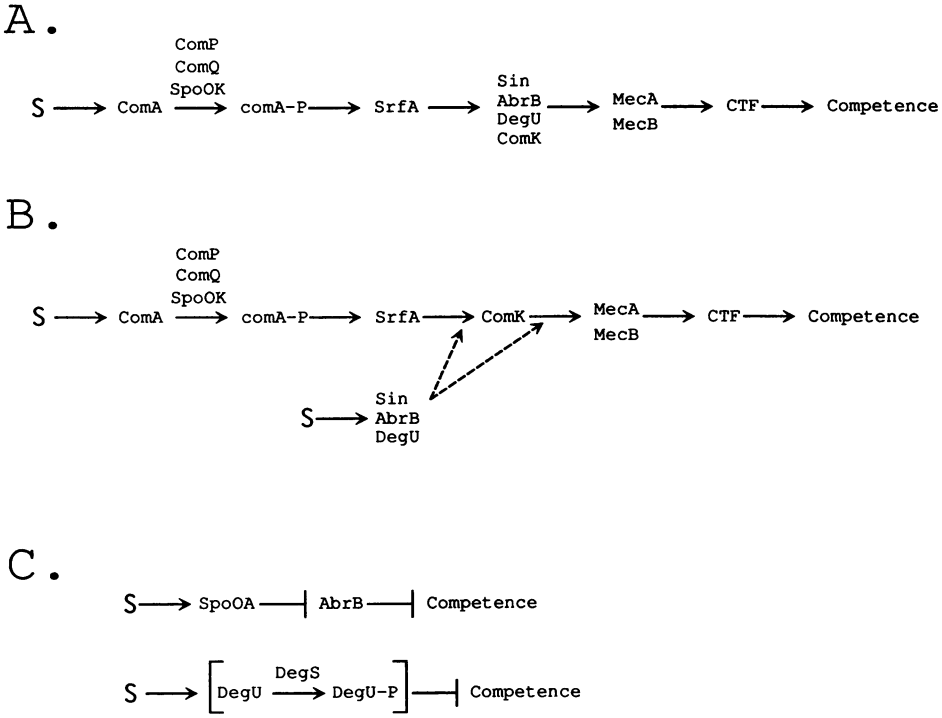

FIG. 10. Regulation of competence; schemes of signal transduction and information flow. The boldface S's indicate likely points at which signals are received by the regulatory apparatus. Arrows and lines terminated by perpendicular bars indicate positive and negative regulation, respectively. (A and B) Alternative arrangements that are consistent with the available data, as outlined in the text. The points of action of Sin, AbrB, and DegU in panel B are ambiguous, as shown. (C) Summary of two possible pathways of negative control. Spo0A is required to down-regulate the expression of *abrB*, which can act negatively as well as positively on competence (see panels A and B). Since activation of DegU appears to act negatively on competence, DegS may be a negative regulator.

the dependence of *comK* on these genes has not been reported.

Also shown in Fig. 10C are two potential negative circuit elements in the competence scheme. As noted above, DegS is not required for the development of competence. However, since the *degS*^{h} and *degU*^{h} mutations confer competence deficiency, it may be that under certain conditions DegS transfers phosphate to DegU, thus acting as a negative regulator. Also as noted above, an excess of AbrB acts negatively on competence, explaining the requirement for the Spo0A protein. It is possible that Spo0A acts at T_0 to down-regulate *abrB* expression and that AbrB is a normal negative regulator of competence. Figure 10 indicates that there are several logical points in the schemes at which signals may impinge on the system. These include the ComA-ComQ-ComP-Spo0K proteins, as well as DegS-DegU and Spo0A, subject to the uncertainties just mentioned.

The position of the *spo0H* product in these schemes is ambiguous. Although the *srfA* mutation *csh-293* was identified on the basis of its *spo0H* dependency (100), we have noted only a two- to threefold effect of *spo0H* deficiency on *srfA* expression in competence medium (89). The bypass of *spo0H* by *mecA42* suggests that σ^{H} is required prior to *mecA*. Similarly ambiguous are the points of negative action of AbrB. As noted above, overproduction of this product seems to be the sole reason for the competence deficiency of *spo0A* mutants. This deficiency is partially bypassed by *mec* mutants, implying multiple sites of this negative AbrB activity, both before and after the point of *mec* action. As noted above, a double *abrB spo0A* mutant is completely bypassed by *mecA42*, since in this strain the AbrB protein is no longer available to act negatively and is no longer required to act positively as a result of the *mecA* mutation.

How Does the Cascade Function?

The formal schemes presented in Fig. 10 tell us little about mechanism. What can be inferred concerning the modes of action of the various regulatory elements? Four known loci are required to turn on *srfA* expression: *compP*, *comA*, *comQ*, and *spo0K*. The observations reported above, that *srfA* under *Pspac* control bypasses the requirements for *comA*, *comQ*, and *spo0KORFE*, suggest that these genes are not required for any purpose other than the activation of *srfA* expression. Thus, they would not be needed for the postulated second branch shown in Fig. 10B, nor do they act after as well as before *srfA*.

As noted above, ComA and ComP are response regulator and histidine kinase members, respectively, of the two-component signal transduction systems. It is likely that ComP serves to activate ComA, both because of the genetic proximity of these two genes and because overexpression of ComA bypasses the competence deficiency of a *comP* mutant (239). This has been observed for other cognate two-component regulators (190, 240). Presumably this activation involves phosphate transfer. Little is known of the role of ComQ. However, the extents of competence deficiency in mutants with null mutations of *comP*, *comA*, or *comQ*, as well as of multiple mutants for these genes, are identical, suggesting that the corresponding gene products may act in

concert. Also, overexpression of ComA bypasses a *comQ* knockout mutant (234) and *comQ* is located immediately upstream from *comP* on the genetic map (Fig. 7). Perhaps ComQ functions as an auxiliary protein facilitating or regulating the transfer of phosphate to ComA or the stability of the phosphorylated protein.

The role of *spo0K* is particularly intriguing. Rudner et al. (174) have shown that at least *spo0K* ORFE is required for competence. We have found that the requirement for this ORF can be partially bypassed by overexpression of ComP or ComQ (234). This appears to be the case for overexpression of ComA as well, although the *spo0KORFE* strain carrying *comA* in multicopy grew poorly and was difficult to work with. As noted above, ComP was predicted to be a membrane protein, possessing eight transmembrane segments in its N-terminal domain (239). The *spo0K* proteins, which closely resemble the gram-negative oligopeptide permease proteins, are almost certainly associated with the membrane. Perhaps the *spo0K* products and ComP "talk" to one another (see reference 174) and the discussion below), transducing a signal via ComA and possibly ComQ.

It has been suggested (174) that the Spo0K-dependent signal transduction system may be sensitive to the concentration of peptide pheromones, such as the soluble sporulation factor of *B. subtilis* (82) or the postulated soluble competence factor (101). It has also been proposed that the *spo0K* system may respond to cell wall-derived peptides, as the homologous *opp* operon is thought to do in *Salmonella* spp. (156). Of course these two hypotheses are not mutually exclusive, since the cell wall-derived peptides may diffuse into the extracellular space and act as soluble factors.

As a result of signaling events involving *comQ*, *comA*, *comP*, and *spo0K*, the expression of *srfA* is activated. This operon contains genes required for the stepwise assembly of peptides (148, 225). It is tempting to propose, as has been done (148, 225), that one of the intermediates in surfactin synthesis, or a derivative of such an intermediate, serves as an extracellular competence factor for *B. subtilis*. Although this is indeed a plausible notion, it should be recalled that the pathway responsible for nonribosomal peptide synthesis involves the formation of aminoacyl adenylates and their coupling to phosphopantetheine arms on a multicomponent synthetic complex (119). There are several opportunities for interaction of these intermediates with other pathways, and perhaps these provide an intracellular signal for competence development. If *srfA* is needed for the synthesis of a peptide competence factor, it might seem less likely that *spo0K* is a sensor for such a peptide, since it acts prior to *srfA*. On the other hand, we can envisage an autocatalytic amplification system, in which a low concentration of the factor is sensed by the *spo0K-comP-comQ-comA* system, resulting in the activation of *srfA* and the increased production of the factor. If this type of "autocrine" signaling cycle operates, the observation (89) that expression of *srfA* under *Pspac* control bypasses the competence requirement for *spo0KORFE* implies that a sensor for the postulated SrfA peptide that does not involve *spo0KORFE* must exist, acting later than *srfA*.

The thoughts and observations presented in the preceding paragraphs suggest the following working hypothesis. A low level of the postulated competence factor may be constitutively produced, since *srfA* is expressed at a low but measurable level throughout growth (89, 100, 149). This implies that as growth proceeds, the extracellular concentration of this factor would increase to a critical level, interacting with *spo0K* products, which in turn transmit a signal to the *comPAQ* proteins, activating ComA by phosphorylation and initiating an increase in *srfA* transcription. This would result in an increase in synthesis of the factor, which in turn would trigger further *srfA* expression by the same mechanism. An autocatalytic switch would thus be operative, responding to cell density. For the postulated extracellular factor to interact with the cell and potentiate the development of competence, a receptor-signal transduction system must also operate downstream of *srfA*. The ability of *srfA* expression under *Pspac* control to bypass a *spo0KORFE* disruption would suggest that the Spo0K ORFE protein is not needed for this downstream step, although perhaps one or more of the Spo0KA to Spo0KD proteins are required for both the upstream and downstream signaling events.

These postulated roles for *spo0K* in competence-specific signal transduction and the proposed interactions of *spo0K* with the ComA-ComP two-component regulators are strikingly reminiscent of the workings of the *E. coli pst* operon during phosphate regulation (164). *pst* and *spo0K* encode related ATP-driven, membrane-associated transport systems. Both operons encode substrate-specific binding proteins, two hydrophobic membrane-associated components, and a less hydrophobic protein that contains an ATP-binding consensus sequence. Both are required for control of global regulons. If our hypothesis is correct, in both cases regulation occurs by interaction with two-component regulators (*comP-comA* and *phoR-phoB*). Both operons include a fifth ORF (*phoU* and *spo0KORFE*), not found in various otherwise similar ATP-driven transport systems, and in both cases this ORF appears to be required for interaction with the two-component regulatory pathways, but not for transport of P_i or of oligopeptides (156). Thus the two global regulation systems may well be similarly organized and may have mechanistic features in common. For instance, the *pho* system does not appear to be regulated by sensing the intracellular level of P_i or the flux of P_i across the membrane, but rather the Pst proteins appear to function as part of a transmembrane signaling device (164). Possibly the products of *spo0K* act similarly downstream of *srfA*, without actually internalizing the postulated competence factor.

Clearly, the development of competence responds to signals that are at least in part nutritional. A typical medium used for the development of competence consists of salts, glucose, amino acids, and yeast extract (7). The yeast extract is dispensable, serving only to increase the growth rate and the yield of competent cells. Substitution of glucose by glycerol in this medium has little or no effect on growth but decreases the yield of transformants and the expression of late competence genes about 10-fold (3). In this medium, as stated above, competence develops postexponentially, although an initial minor wave of competence is almost always detected at about T_{-3} to T_{-2} (3). When amino acids are omitted (except for those required for growth), competence develops to the usual level but is expressed throughout growth, with a further increase at T_0 (170). Addition of amino acids in the form of casein hydrolysate, to final concentrations of either 0.02% or 2%, have the same effect; the expression of competence is repressed during exponential growth, but is induced beginning at T_0. This suggests that repression by amino acids is overridden at the time of translation to stationary phase by some specific postexponential mechanism, other than by exhaustion of amino acids. The regulatory genes required for the exponential expression of competence in the absence of amino acids, as opposed to the postexponential form of expression, are not known.

In a *comP* null mutant, bypassed for competence by overexpression of ComA (239), competence is no longer

dependent on the presence of glucose, nor inhibited by the addition of glutamine. Msadek et al. (145) have observed that *degQ* expression is dependent on *comA* and *comP* (and also on *comQ*). *degQ* is present in the *comA* gene cluster (Fig. 7) but is not required for the development of competence. In fact, the overproduction of DegQ has a slight negative effect on competence (145). The induction of *degQ* expression can be triggered by amino acid starvation, using a *comA*- and *comP*-dependent mechanism. Perhaps one of the signals detected by *comP* and *comA* is the availability of amino acids, although the relationship of the observations of Msadek et al. (145) to ours on the development of competence in the presence and absence of amino acid supplements is not yet clear. These fragmentary observations indicate that the nutritional signals detected by the ComP mechanism are likely to be complex, but may involve at least in part amino acid starvation-dependent and glucose-dependent signals. A major unanswered question involves the relationship of nutritional signaling to the sensing of a competence factor. Perhaps the proposed interaction of *comP* and *spo0K* permits integration of these two types of signals.

It is difficult to even posit a plausible hypothesis concerning the signals and mechanisms responsible for the cell type-specific regulation of competence, without resort to fantasy. Perhaps the heterogeneity of competent cultures is determined by titration of a limiting substrate (for instance a competence pheromone) or by the proportion of cells in a critical window of the cell cycle when a specific signal is received. It is also conceivable that a cell division event occurs with the daughter cells undergoing alternative fates (as in sporulation). The noncompetent daughter might then undergo about three divisions to yield the observed ratio of competent to noncompetent cells. One interesting observation is that ComG ORF1 is required for the separation in Renografin (1). It is not known whether it is also needed for heterogeneity in gene expression of the competent culture, or whether it is required only for the density difference.

Under certain conditions, and possibly during the development of competence, *B. subtilis* extrudes high-molecular-weight chromosomal DNA, apparently without lysis and in the genetic map order (19, 66). This specific extrusion process may conceivably take place from the noncompetent subpopulation. If this proves to be the case, then the competence phenomenon may be viewed as a primitive form of sexual dimorphism. The extrusion process requires further study.

Regulation of Competence in Relation to Other Forms of Postexponential Expression: Signal Transduction Network

When *B. subtilis* reaches the transition from the exponential to the stationary phase, several forms of expression may be activated. In addition to competence and sporulation, the cells may release antibiotics, increase the synthesis of a variety of degradative enzymes, and increase in motility. Only sketchy information exists about the relationship among these various forms of expression. Clearly, sporulation is an ultimate response to stress, since it would seem to preclude the other responses. Sporulation is inhibited by high concentrations of nutrients, particularly glucose. Competence, on the other hand, is stimulated by the presence of glucose. In fact, to a certain extent these seem to be alternative responses when viewed genetically as well as nutritionally. For instance, Sin is required for competence, but its overexpression inhibits sporulation (75). Conversely, ComA is required for competence, but its overexpression inhibits sporulation (237). $degU^h$ and $degS^h$ mutations enable sporulation to occur in the presence of glucose, but prevent the development of competence (144, 171, 202). It is possible that the signaling requirements for competence and sporulation are arranged so that these two responses occur sequentially; perhaps the exhaustion of nutrients that favor the former response lead to the initiation of sporulation. Certainly, competent cells can go on to sporulate, presumably without an intervening period of cell division (42).

These genetic relationships between sporulation and competence serve to introduce an interesting generalization. Few, if any, of the early regulatory genes discussed in this review are specific for competence. In fact, most of the early regulatory genes that act on a given form of postexponential expression seem to affect others as well. This is true of *spo0K*, *comA*, *comP*, *spo0A*, *spo0H*, *abrB*, *sin*, *degU*, *degS*, *srfA*, *mecA*, and *mecB* at least. For *spo0H*, *spo0A*, and *spo0K* this pleiotropy is obvious. As noted above, overexpression of Sin results in asporogeny (75). A loss of *sin* function results in competence deficiency, nonmotility, and overexpression of certain degradative enzymes. $degS^h$ and $degU^h$ mutants, in addition to the phenotypes described above, are nonmotile and, of course, overproduce degradative enzymes. ComP deficiency also affects sporulation (239). *mecB* mutants are oligosporogenic, and *mecA* mutants may exhibit an altered response of sporulation to the presence of glucose (59). Also, as noted above, *mec* mutations affect regulation of the *gsi* (146) and degradative enzyme (143) systems. Thus, for each of these loci, either overexpression or particular alleles affect more than one form of expression. This has lead to the notion of a signal transduction network that regulates the expression of a postexponential super regulon (52). This network is presumably capable of gathering information from various signal molecules and integrating this information to ensure that the appropriate response occurs at the proper time. Particular combinations of signals may therefore lead to competence, competence plus degradative enzyme synthesis, sporulation, etc.

Regulatory Products and Loci in Other Competence Systems

Little is known concerning regulatory products in the *Haemophilus* system. It has been reported that cyclic AMP induces competence when added to growing cultures, suggesting that this molecule may act as a positive transcriptional factor, and that catabolite repression may play a role in competence regulation in this organism (103).

A protein competence factor in *S. pneumoniae* has been identified and partially characterized (218). The *trt* mutation results in the constitutive expression of competence, even when the culture is grown in the presence of trypsin. This has been interpreted as due to bypass of the competence factor requirement by the mutation. The *trt* locus has not been further characterized. Additional so-called conditional mutants have been isolated, which respond to exogenously added competence factor (142). A DNA fragment of 4 to 5 kb containing the conditional *comAB* locus has been isolated and characterized (27, 28, 142). Expression in *E. coli* revealed that this fragment encodes two proteins of 77 and 49 kDa. The sequence (99) of the larger protein (ComA) reveals it to be a member of the family of bacterial ATP-dependent transport proteins (96). The most striking similarity is to a subset of this family that includes proteins needed for the secretion of bacterial toxins; this led Hui and Morrison (99) to suggest that ComA may play a role in the export of the

competence factor. This hypothesis is experimentally testable, and the answers should soon be available.

EVOLUTION OF COMPETENCE

It was pointed out above that the regulation of competence in different systems is quite variable in that competence responds to different signals at various times in the growth of cultures and exhibits no regulation at all in some organisms. This may reflect the needs of each specific organism, which probably has adapted an existing global regulatory system to control competence. It is plausible to regard the evolution of the competence regulatory mechanism in *B. subtilis* as one aspect of the evolution of the postexponential signal transduction network. In this network several motifs reappear: the use of two-component regulators, transcription factors, membrane transducers, extracellular signaling molecules (apparently for sporulation and competence at least), etc. These repeated motifs, together with the amino acid sequence similarities among several of the related proteins, suggest that the members of these categories of proteins have common ancestors.

Many of the observations presented above also suggest a resemblance between the postexponential regulatory network in the *B. subtilis* system and growth factor-dependent signal transduction in eukaryotic systems, leading to developmental and proliferative responses (25, 36). Several similarities can be noted, including a likely role for extracellular peptide signals, membrane receptors, interaction of the receptors with kinases, transfer of phosphate residues to effector molecules, overlapping responses to related signals, and ultimate transcriptional responses. Although the molecules involved in the prokaryotic and eukaryotic systems do not seem to be closely related on the sequence level, the organization of these systems is strikingly similar.

Several hypotheses have been advanced to explain the evolutionary origin of competence mechanisms themselves. These have been conveniently summarized in a recent discussion (115). Two of these hypotheses may be classified as examples of coincidental evolution (115). These propose that competence evolved as a means of feeding on exogenous DNA (203) or as a mechanism for importing templates for the repair of damage to DNA (134). In both cases the ability to be transformed would, at least originally, evolve as a coincidental consequence of the uptake mechanism together with the existence of an efficient pathway for homologous recombination. (The latter might itself have evolved as a consequence of a DNA repair system). On the other hand, transformability may have evolved as a means of DNA transfer encoded by parasitic DNA, such as a phage or plasmidlike element, and the transfer mechanism may then have been captured by the host. Finally, transformational recombination may have evolved as a means of increasing fitness either of individual organisms or of a bacterial population. These hypotheses are not mutually exclusive. Several forms of selective pressure may have operated during the evolution of competence, perhaps at different stages. We can do little but speculate on the plausibility of the various ideas, although, as we shall see, certain experimental findings may be relevant.

The feeding notion is not particularly attractive. For instance, *B. subtilis* secretes a powerful nonspecific nuclease late in growth. This, together with the existence of high-affinity transport mechanisms, would seem to provide the basis for a simple and effective scavenging pathway, making it unlikely that a more complex DNA uptake system capable of transporting macromolecular DNA would evolve for this purpose.

The DNA repair idea has received a fair amount of attention both in connection with the evolution of genetic competence and as a hypothesis for the evolution of sex in general (15). Several interesting experimental studies have been reported that bear on the repair hypothesis during transformation of *B. subtilis* (98, 134, 243). These authors have demonstrated that the transformation frequency for a given marker increases with UV dosage when the cells are irradiated prior to the addition of transforming DNA, but decreases with dosage when the cells are irradiated after transformation. They have concluded that the UV-dependent increase in transformation is not due to the induced expression of the SOS system or to an intrinsic hardiness of the competent subpopulation. They concluded that the increase is due either to an enhancement of recombination by UV irradiation or to transformational repair of otherwise deleterious radiation damage, thus enhancing the survival of the damaged cells. They point out that either conclusion is consistent with the hypothesis that transformation functions in DNA repair. This is obvious in the case of the second alternative. If the first possibility is true, then ". . . cells are inducing sex in precisely the environment predicted by the repair hypothesis" (134). If the enhancement of recombination is targeted to damaged regions, the two explanations for the data would tend to merge, since the targeted recombination should often result in repair. If not, it would be necessary to postulate some sort of generalized induction of recombination by prior UV irradiation which would not necessarily result in repair and would offer little support to the evolutionary repair hypothesis.

Several further observations are also consistent with the repair hypothesis. First, the UV-dependent enhancement of transformation is not observed with plasmid transformation, which proceeds without homologous recombination (134, 243). Second, this enhancement does not depend on the excision repair and inducible SOS systems (243). Third, the enhancement is also observed when UV-damaged donor DNA is used (98). The latter situation presumably mimics the natural case, in which the donor and recipient cells might be exposed to the same conditions and the donor cells would be at least as likely to suffer DNA damage as the recipients. The authors infer from this that the enhanced recombination is indeed targeted (see above), since otherwise no net increase in the yield of transformants would be observed. This is because random (untargeted) recombination would be just as likely to introduce a DNA lesion as to repair one, resulting in no net increase in survivable transformants.

In addition to these experimental findings, an interesting computer-modeling study has explored the plausibility of the hypothesis that the uptake of transforming DNA, even of damaged DNA, would increase the fitness of the recipient population, by providing for the repair of DNA damage (165). This simulation led to the conclusion that transformation can indeed increase the fitness of a population, strengthening the repair hypothesis.

The repair hypothesis is rendered more plausible by the observation that the *B. subtilis* SOS-like system is induced as part of the competence regulon (121, 124). This may reflect the evolution of regulatory mechanisms that result in the activation of a variety of damage repair pathways (recombinational and nonrecombinational) under conditions that are likely to produce such damage. This favors the idea that competence may be, at least in part, a means of accumulating DNA templates for repair. It has been sug-

gested that the switch to oxidative metabolism that occurs in *B. subtilis* at the end of exponential growth may result in an increase of free radical damage to DNA (114). It has also been pointed out that slow-growing cells entering the stationary phase would tend to have fewer chromosomal copies than would exponentially growing cells (98). This absence of redundancy would make the cell less able to repair DNA damage by recombination between sister chromosomes and hence might favor the evolution of other repair pathways, to be expressed precisely when the growth rate decreases. These might include competence as a means of increasing the redundancy of damaged genes.

The notion that transformation may have evolved from a mechanism for the spread of parasitic DNA was considered by Levin (115), who pointed out certain difficulties with this theory. These derive from the fact that the competence mechanism is encoded by chromosomal DNA, and it is difficult to imagine reasonable scenarios for the maintenance and selection of the competence genes. However, it is worth considering a scenario in which the parasitic DNA was originally an autonomously replicating element, such as a bacteriophage. Consider the case of a phage resembling the filamentous viruses of gram-negative organisms (136). These attach to pili and enter the cell by an unknown mechanism, whereupon they replicate by a rolling-circle mechanism, eventually resulting in the net accumulation of mature (single-stranded circular) viral DNA. This DNA interacts with structural and morphogenetic proteins at the cell surface, and assembly, extrusion, and release of the virion follow. Several features of this life cycle are suggestive. The passage of single-stranded DNA across the bacterial cell membrane takes place, as in transformation. Pili are involved in phage attachment and possibly in entry, whereas proteins that resemble pilin are required for transformation. Like the filamentous phages, most of the small *B. subtilis* plasmids replicate by a rolling-circle mechanism. We may propose the existence of an ancient bacteriophage resembling the filamentous phages of gram-negative organisms, capable of growth on *B. subtilis*. This virus would encode the proteins required for its own assembly and replication and would utilize appropriate host proteins (e.g., pilinlike proteins and proteins required for pilus assembly) during its life cycle. With the passage of time, several of these phage functions might have been captured by the host and may have evolved to perform novel functions, unrelated to the needs of the phage ancestor. Perhaps the ability of the bacteriophage assembly and extrusion system to transport DNA provided precursors for the evolution of DNA uptake systems, which contributed to host fitness by providing templates for the repair of DNA damage. The prevalence among gram-positive plasmids of rolling-circle replication systems that closely resemble those of the filamentous phages (83) might be taken as evidence that such phages once existed, and possibly still do. Recently a report of a filamentous phagelike particle in *Clostridium* spp. has appeared (104).

It is likely that the further identification and characterization of competence genes will continue to clarify the relationships of these genes and their gene products to other genes from *Bacillus* spp. as well as from other organisms. These relationships should shed additional light on the evolution of the competence machinery.

FUTURE PROSPECTS

Much has been learned about the competence system of *B. subtilis*, particularly concerning its regulatory aspects, although, clearly, much remains to be discovered. As more regulatory genes are identified, and as more gene products are characterized biochemically, it is reasonable to expect rapid progress in elucidating the mechanisms involved. The active and cooperative interest of many groups in the phenomena of postexponential regulation in *B. subtilis* will continue to aid in this task and to clarify the relationship of competence regulation to that of other postexponential forms of expression. Among the major problems that are now being addressed in several laboratories are identification of the precise signals involved and their target molecules, identification of the targets of the probable transcriptional regulators, and elucidation of the nature and roles of extracellular factors.

The mechanisms of transformation, and in particular of the binding, processing, and uptake of DNA, are even less well understood. Fewer gene products have been identified, and fewer still have been characterized biochemically. The structural dimension will require the description of membrane-associated multicomponent complexes and the dissection of the functional roles of the subunits of these complexes. Although these structural aspects seriously complicate future tasks and present a challenge to the investigator, they also add greatly to the intrinsic biological interest of the competence system.

ACKNOWLEDGMENTS

I express deep appreciation to the many investigators who have communicated unpublished results and have contributed to this review with many stimulating discussions: F. Breidt, Y. Chung, M. K. Dahl, J. Dubnau, N. K. Gaur, G. Grandi, B. Green, A. Grossman, J. Hahn, D. Henner, M. Hobbs, L. Kong, V. Krishnapillai, F. Kunst, T. Leighton, S. P. Livingston, A. Londoño, R. Magnuson, J. S. Mattick. I. Mandic-Mulec, T. Msadek, J. Mueller, M. Nakano, A. Ninfa, A. Pugsley, G. Rapoport, M. Roggiani, Y. Sadaie, D. van Sinderen, I. Smith, L. Sonenshein, G. Venema, Y. Weinrauch, C. B. Whitchurch, P. Youngman, and P. Zuber.

The work from my laboratory was supported by NIH grants AI10311, GM37137, and GM43756.

REFERENCES

1. **Albano, M., R. Breitling, and D. Dubnau.** 1989. Nucleotide sequence and genetic organization of the *Bacillus subtilis comG* operon. J. Bacteriol. **171**:5386–5404.
2. **Albano, M., and D. Dubnau.** 1989. Cloning and characterization of a cluster of linked *Bacillus subtilis* late competence mutants. J. Bacteriol. **171**:5376–5385.
3. **Albano, M., J. Hahn, and D. Dubnau.** 1987. Expression of competence genes in *Bacillus subtilis*. J. Bacteriol. **169**:3110–3117.
4. **Ambudkar, S. V., T. J. Larson, and P. C. Maloney.** 1986. Reconstitution of sugar phosphate transport systems of *Escherichia coli*. J. Biol. Chem. **261**:9083–9086.
5. **Ambudkar, S. V., L. A. Sonna, and P. C. Maloney.** 1986. Variable stoichiometry of phosphate-linked anion exchange in *Streptococcus lactis*: implications for the mechanism of sugar phosphate transport by bacteria. Proc. Natl. Acad. Sci. USA **83**:280–284.
6. **Ames, G. F.-L.** 1990. Energy coupling in bacterial periplasmic permeases. J. Bacteriol. **172**:4133–4137.
7. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–46.
8. **Antoniewski, C., B. Savelli, and P. Stragier.** 1990. The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. J. Bacteriol. **172**:86–93.
9. **Aricó, B., J. F. Miller, C. Roy, S. Stibitz, D. Monack, S. Falkow, R. Gross, and R. Rappuoli.** 1989. Sequences required for expression of *Bordetella pertussis* virulence factors share

homology with prokaryotic signal transduction proteins. Proc. Natl. Acad. Sci. USA **86**:6671–6675.
10. **Arwert, F., and G. Venema.** 1973. Evidence for a non-covalently bonded intermediate in recombining dining transformation in *Bacillus subtilis*, p. 203–214. *In* L. J. Archer (ed.), Bacterial transformation. Academic Press, Inc., New York.
11. **Arwert, F., and G. Venema.** 1973. Transformation in *Bacillus subtilis*. Fate of newly introduced transforming DNA. Mol. Gen. Genet. **123**:185–198.
12. **Ayusawa, D., Y. Yoneda, K. Yamane, and B. Maruo.** 1975. Pleiotropic phenomena in autolytic enzyme(s) content, flagellation and simultaneous hyperproduction of extracellular α-amylase and protease in a *Bacillus subtilis* mutant. J. Bacteriol. **124**:459–469.
13. **Bally, M., G. Ball, A. Baudere, and A. Lazdunski.** 1991. Protein secretion in *Pseudomonas aeruginosa*: the *xcpA* gene encodes an integral inner membrane protein homologous to *Klebsiella pneumoniae* secretion function protein PulO. J. Bacteriol. **173**:479–486.
14. **Barouki, R., and H. O. Smith.** 1985. Reexamination of phenotypic defects in *rec-1* and *rec-2* mutants of *Haemophilus influenzae* Rd. J. Bacteriol. **163**:629–634.
15. **Bernstein, H., F. A. Hopf, and R. E. Michod.** 1988. Is meiotic recombination an adaptation for repairing DNA, producing genetic variation or both? p. 139–160. *In* R. E. Michod and B. R. Levin (ed.), The evolution of sex. Sinauer Associates, Inc., Sunderland, Mass.
16. **Biswas, G. D., T. Sox, and P. F. Sparling.** 1977. Factors affecting genetic transformation of *Neisseria gonorrhoeae*. J. Bacteriol. **129**:983–992.
17. **Bodmer, W., and A. T. Ganesan.** 1964. Biochemical and genetic studies of integration and recombination in Bacillus subtilis transformation. Genetics **50**:717–738.
18. **Bodmer, W. F.** 1966. Integration of deoxyribonuclease-treated DNA in *Bacillus subtilis* transformation. J. Gen. Physiol. **49**:233–258.
19. **Borenstein, S., and E. Ephrati-Elizur.** 1969. Spontaneous release of DNA in sequential genetic order by *B. subtilis*. J. Mol. Biol. **45**:137–152.
20. **Bovre, K., and L. O. Froholm.** 1972. Competence in genetic transformation related to colony type and fimbriation in three species of *Moraxella*. Acta Pathol. Microbiol. Scand. **80**:649–659.
21. **Breidt, F., M. Roggiani, and D. Dubnau.** Unpublished data.
22. **Breitling, R., and D. Dubnau.** 1990. A pilin-like membrane protein is essential for DNA binding by competent *Bacillus subtilis*. J. Bacteriol. **172**:1499–1508.
23. **Bresler, S. E., R. A. Kreneva, and V. V. Kushev.** 1968. Correction of molecular heterozygotes in the course of transformation. Mol. Gen. Genet. **102**:257–268.
24. **Burbulys, D., K. A. Trach, and J. A. Hoch.** 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell **64**:545–552.
25. **Cantley, L. C., K. R. Auger, C. Carpenter, B. Duckworth, A. Graziani, R. Kapeller, and S. Soltoff.** 1991. Oncogenes and signal transduction. Cell **64**:281–302.
26. **Cantor, C. R., and P. R. Schimmel.** 1980. Biophysical chemistry. Part III. The behavior of biological macromolecules. W. H. Freeman and Co., San Francisco.
27. **Chandler, M. S., and D. A. Morrison.** 1987. Competence for genetic transformation in *Streptococcus pneumoniae*: molecular cloning of *com*, a competence control locus. J. Bacteriol. **169**:2005–2011.
28. **Chandler, M. S., and D. A. Morrison.** 1988. Identification of two proteins encoded by *com*, a competence control locus of *Streptococcus pneumoniae*. J. Bacteriol. **170**:3136–3141.
29. **Cheo, D. L., K. W. Bayles, and R. E. Yasbin.** 1991. Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. J. Bacteriol. **173**:1696–1703.
30. **Christie, P. J., J. J. E. Ward, M. P. Gordon, and E. W. Nester.** 1989. A gene required for transfer of T-DNA to plants encodes an ATPase with autophosphorylating activity. Proc. Natl. Acad. Sci. USA **86**:9677–9681.
31. **Ciferri, O., S Barlati, and J. Lederberg.** 1970. Uptake of synthetic polynucleotides by competent cells of *Bacillus subtilis*. J. Bacteriol. **104**:684–688.
32. **Clavé, C., and M.-C. Trombe.** 1987. Is DNA transport driven by the proton electrochemical potential difference in the naturally transfromable bacteria, *Streptococcus pneumoniae*? Bioelectrochem. Bioenerg. **17**:269–276.
33. **Clavé, C., and M.-C. Trombe.** 1989. DNA uptake in competent *Streptococcus pneumoniae* requires ATP and is regulated by cytoplasmic pH. FEMS Microbiol. Lett. **65**:113–118.
34. **Cole, S. T., and O. Raibaud.** 1986. The nucleotide sequence of the *malT* gene encoding the positive regulator of the *Escherichia coli* maltose regulon. Gene **42**:201–208.
35. **Contente, S., and D. Dubnau.** 1979. Marker rescue transformation by linear plasmid DNA in *Bacillus subtilis*. Plasmid **2**:555–571.
36. **Cross, M, and T. M. Dexter.** 1991. Growth factors in development, transformation, and tumorigenesis. Cell **64**:271–280.
37. **Cutting, S., and J. Mandelstam.** 1986. The nucleotide sequence and the transcription during sporulation of the *gerE* gene of *Bacillus subtilis*. J. Gen. Microbiol. **132**:3013–3024.
38. **Dahl, M. K., T. Msadek, F. Kunst, and G. Rapoport.** 1991. Mutational analysis of the *Bacillus subtilis* DegU regulator and its phosphorylation by the DegS protein kinase. J. Bacteriol. **173**:2539–2547.
39. **Danner, D. B., R. A. Deitch, K. Sisco, and H. O. Smith.** 1980. An eleven-base-pair sequence determines the specificity of DNA uptake in *Haemophilus*. Gene **11**:311–318.
40. **Danner, D. B., H. O. Smith, and S. Narang.** 1982. Construction of DNA recognition sites active in *Haemophilus* transformation. Proc. Natl. Acad. Sci. USA **79**:2393–2397.
41. **Davidoff-Abelson, R., and D. Dubnau.** 1973. Kinetic analysis of the products of donor deoxyribonucleate in transformed cells of *Bacillus subtilis*. J. Bacteriol. **116**:154–162.
42. **de Lencastre, H., and P. J. Piggot.** 1979. Identification of different sites of expression for *spo* loci by transformation of *Bacillus subtilis*. J. Gen. Microbiol. **114**:377–389.
43. **d'Enfert, C., I. Reyss, C. Wandersman, and A. P. Pugsley.** 1989. Protein secretion by gram negative bacteria. Characterization of two membrane proteins required for pullulanase secretion by *Escherichia coli* K-12. J. Biol. Chem. **264**:17462–17468.
44. **Devine, J. H., G. S. Shadel, and T. O. Baldwin.** 1989. Identification of the operator of the *lux* regulon from the *Vibrio fischeri* strain ATCC7744. Proc. Natl. Acad. Sci. USA **86**:5688–5692.
45. **de Vos, W. M., S. C. deVries, and G. Venema.** 1983. Cloning and expression of the *Escherichia coli recA* gene in *Bacillus subtilis*. Gene **25**:301–308.
46. **de Vos, W. M., G. Venema, U. Canosi, and T. A. Trautner.** 1981. Plasmid transformation in *Bacillus subtilis*: fate of plasmid DNA. Mol. Gen. Genet. **181**:424–433.
47. **Dooley, D. C., C. T. Hadden, and E. W. Nester.** 1971. Macromolecular synthesis in *Bacillus subtilis* during development of the competent state. J. Bacteriol. **108**:668–679.
48. **Dubnau, D.** Unpublished data.
49. **Dubnau, D.** 1976. Genetic transformation of *Bacillus subtilis*: a review with emphasis on the recombination mechanism, p. 14–27. *In* D. Schlessinger (ed.), Microbiology—1976. American Society for Microbiology, Washington, D.C.
50. **Dubnau, D.** 1982. Genetic transformation in *Bacillus subtilis*, p. 148–178. *In* D. Dubnau (ed.), The molecular biology of the bacilli, vol. I. *Bacillus subtilis*. Academic Press, Inc., New York.
51. **Dubnau, D.** 1989. The competence regulon of *Bacillus subtilis*, p. 147–166. *In* I. Smith, R. A. Slepecky, and P. Setlow (ed.), Regulation of procaryotic development. American Society for Microbiology, Washington, D.C.
52. **Dubnau, D.** 1991. The regulation of genetic competence in *Bacillus subtilis*. Mol. Microbiol. **5**:11–18.
53. **Dubnau, D., and C. Cirigliano.** 1972. Fate of transforming deoxyribonucleic acid after uptake by competent *Bacillus subtilis*: size and distribution of the integrated donor sequences. J. Bacteriol. **111**:488–494.

54. **Dubnau, D., and C. Cirigliano.** 1972. Fate of transforming DNA following up by competent *Bacillus subtilis*. IV. The endwise attachment and uptake of transforming DNA. J. Mol. Biol. **64**:31–46.
55. **Dubnau, D., and C. Cirigliano.** 1972. Fate of transforming DNA following uptake by competent *Bacillus subtilis*. III. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. J. Mol. Biol. **64**:9–29.
56. **Dubnau, D., and C. Cirigliano.** 1973. Fate of transforming DNA following uptake by competent *Bacillus subtilis*. VI. Non-covalent association of donor and recipient DNA. Mol. Gen. Genet. **120**:101–106.
57. **Dubnau, D., S. Contente, and T. J. Gryczan.** 1980. On the use of plasmids for the study of genetic recombination in *Bacillus subtilis*, p. 365–386. *In* S. Zadrazil and J. Sponar (ed.), DNA-recombination interactions and repair. Pergamon Press, Oxford.
58. **Dubnau, D., and R. Davidoff-Abelson.** 1971. Fate of transforming DNA following uptake by competent *Bacillus subtilis*. I. Formation and properties of the donor-recipient complex. J. Mol. Biol. **56**:209–221.
59. **Dubnau, D., R. Davidoff-Abelson, B. Scher, and C. Cirigliano.** 1973. Fate of transforming deoxyribonucleic acid after uptake by competent *Bacillus subtilis*: phenotypic characterization of radiation-sensitive recombination-deficient mutants. J. Bacteriol. **114**:273–286.
60. **Dubnau, D., J. Hahn, L. Kong, M. Roggiani, and Y. Weinrauch.** 1991. Genetic competence as a post-exponential global response. Semin. Dev. Biol., in press.
61. **Dubnau, D., M. Roggiani, and J. Hahn.** 1990. Growth medium-independent genetic competence mutants of *Bacillus subtilis*. J. Bacteriol. **172**:4048–4055.
62. **Dubnau, E., J. Weir, G. Nair, Luke Carter III, Charles Moran, Jr., and I. Smith.** 1988. *Bacillus* sporulation gene *spo0H* codes for σ^{30} (σ^{H}). J. Bacteriol. **170**:1054–1062.
63. **Duncan, M. L., S. S. Kalman, S. M. Thomas, and C. W. Price.** 1987. Gene encoding the 37,000-dalton minor sigma factor of *Bacillus subtilis* RNA polymerase: isolation, nucleotide sequence, chromosomal locus, and cryptic function. J. Bacteriol. **169**:771–778.
64. **Ehrlich, S. D.** 1977. Replication and expression of plasmids from *Staphylococcus aureus* in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **74**:1680–1682.
65. **Eisenstadt, E., R. Lange, and K. Willecke.** 1975. Competent *Bacillus subtilis* cultures synthesize a denatured DNA binding activity. Proc. Natl. Acad. Sci. USA **72**:323–327.
66. **Ephrati-Elizur, E.** 1968. Spontaneous transformation in *B. subtilis*. Genet. Res. **11**:83–96.
67. **Fani, R., G. Mastromei, M. Polsinelli, and G. Venema.** 1984. Isolation and characterization of *Bacillus subtilis* mutants altered in competence. J. Bacteriol. **157**:153–157.
68. **Ferrari, F. A., K. Trach, D. LeCoq, J. Spence, E. Ferrari, and J. A. Hoch.** 1985. Characterization of the *spo0A* locus and its deduced product. Proc. Natl. Acad. Sci. USA **82**:2647–2651.
69. **Filloux, A., M. Bally, G. Ball, M. Akrim, J. Tommassen, and A. Lazdunski.** 1990. Protein secretion in gram-negative bacteria: transport across the outer membrane involves common mechanisms in different bacteria. EMBO J. **9**:4323–4329.
70. **Foldes, J., and T. A. Trautner.** 1964. Infectious DNA from a newly isolated *B. subtilis* phage. Z. Vererbungsi. **95**:57–65.
71. **Folkhard, W., D. A. Marvin, T. H. Watts, and W. Paranchych.** 1981. Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. J. Mol. Biol. **149**:79–93.
72. **Fornilli, S. L., and M. S. Fox.** 1977. Electron microscope visualization of the products of *Bacillus subtilis* transformation. J. Mol. Biol. **113**:181–191.
73. **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.
74. **Gaur, N. K., K. Cabane, and I. Smith.** 1988. Structure and expression of the *Bacillus subtilis sin* operon. J. Bacteriol. **170**:1046–1053.
75. **Gaur, N. K., E. Dubnau, and I. Smith.** 1986. Characterization of a cloned *Bacillus subtilis* gene which inhibits sporulation in multiple copies. J. Bacteriol. **168**:860–869.
76. **Gaur, N. K., J. Oppenheim, and I. Smith.** 1991. The *Bacillus subtilis sin* gene, a regulator of alternate development processes, codes for a DNA-binding protein. J. Bacteriol. **173**: 678–686.
77. **Goodgal, S. H.** 1982. DNA uptake in *Haemophilus* transformation. Annu. Rev. Genet. **16**:169–192.
78. **Gottfert, M., P. Grob, and H. Hennecke.** 1990. Porposed regulatory pathway encoded by *nodV* and *nodW* genes, determinants of host specificity in *Bradyhrizobium japoncium*. Proc. Natl. Acad. Sci. USA **87**:2680–2684.
79. **Green, B., and P. Youngman. (University of Pennsylvania.)** 1991. Personal communication.
80. **Green, D. M.** 1964. Infectivity of DNA isolated from *Bacillus subtilis* bacteriophage, SP82. J. Mol. Biol. **10**:438–451.
81. **Grinius, L.** 1982. Energetics of gene transfer into bacteria. Sov. Sci. Rev. Sect. D **3**:115–165.
82. **Grossman, A. D., and R. Losick.** 1988. Extracellular control of spore formation in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **85**:4369–4373.
83. **Gruss, A., and S. D. Ehrlich.** 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. **53**:231–241.
84. **Gryczan, T., S. Contente, and D. Dubnau.** 1980. Molecular cloning of heterologous chromosomal DNA by recombination between a plasmid vector and a homologous resident plasmid in *Bacillus subtilis*. Mol. Gen. Genet. **177**:459–467.
85. **Guillen, N., Y. Weinrauch, and D. Dubnau.** 1989. Cloning and characterization of the regulatory *Bacillus subtilis* competence genes, *comA* and *comB*. J. Bacteriol. **171**:5354–5361.
86. **Gurney, T., Jr., and M. S. Fox.** 1968. Physical and genetic hybrids formed in bacterial transformation. J. Mol. Biol. **32**:83–100.
87. **Hadden, C., and E. W. Nester.** 1968. Purification of competent cells in the *Bacillus subtilis* transformation system. J. Bacteriol. **95**:876–885.
88. **Hahn, J., M. Albano, and D. Dubnau.** 1987. Isolation and characterization of Tn*917lac*-generated competence mutants of *Bacillus subtilis*. J. Bacteriol. **169**:3104–3109.
89. **Hahn, J., and D. Dubnau.** Unpublished data.
90. **Hahn, J., Y. Kozlov, and D. Dubnau.** Unpublished data.
91. **Haseltine-Cahn, F., and M. S. Fox.** 1968. Fractionation of transformable bacteria from competent cultures of *Bacillus subtilis* on Renografin gradients. J. Bacteriol. **95**:867–875.
92. **He, S. Y., M. Lindberg, A. K. Chatterjee, and A. Collmer.** 1991. Cloned *Erwinia chrysanthemi out* genes enable *Escherichia coli* to selectively secrete a diverse family of heterologous proteins to its milieu. Proc. Natl. Acad. Sci. USA **88**:1079–1083.
93. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57**:839–872.
94. **Henner, D. (Genetech Corp.).** 1991. Personal communication.
95. **Henner, D. J., M. Yang, and E. Ferrari.** 1988. Localization of *Bacillus subtilis sacU*(Hy) mutations to two linked genes with similarities to the conserved procaryotic family of two-component signaling systems. J. Bacteriol. **170**:5102–5109.
96. **Higgins, C. F., I. D. Hiles, G. P. C. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, and A. W. Bell.** 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. Nature (London) **323**:448–450.
97. **Hoch, J. A., and J. Spizizen.** 1969. Genetic control of some early events in sporulation of *Bacillus subtilis* 168, p. 112–120. *In* L. L. Campbell (ed.), Spores IV. American Society for Microbiology, Washington, D.C.
98. **Hoelzer, M. A., and R. E. Michod.** 1991. DNA repair and the evolution of transformation in *Bacillus subtilis*. III. Sex with damaged DNA. Genetics, in press.
99. **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 a bacterial ATP-dependent transport protein family. J. Bacteriol. **173**:372–381.
100. **Jaacks, K. J., J. Healy, R. Losick, and A. D. Grossman.** 1989. Identification and characterization of genes controlled by the sporulation-regulatory gene *spo0H* in *Bacillus subtilis*. J. Bacteriol. **171**:4121–4129.
101. **Joenje, H., M. Gruber, and G. Venema.** 1972. Stimulation of the development of competence by culture fluids in *Bacillus subtilis* transformation. Biochim. Biophys. Acta **262**:189–199.
102. **Joenje, J., and G. Venema.** 1975. Different nuclease activities in competent and noncompetent cells of *Bacillus subtilis*. J. Bacteriol. **122**:25–33.
103. **Kahn, M. E., and H. O. Smith.** 1984. Transformation in Haemophilus: a problem in membrane biology. J. Membr. Biol. **81**:89–103.
104. **Kim, A. Y., and H. P. Blaschek.** 1991. Isolation and characterization of a filamentous viruslike particle from *Clostridium acetobutylicum* NCIB 6444. J. Bacteriol. **173**:530–535.
105. **Kong, L., and D. Dubnau.** Unpublished data.
106. **Kunst, F., M. Debarbouille, T. Msadek, M. Young, C. Mauel, D. Karamata, A. Klier, G. Rapoport, and R. Dedonder.** 1988. Deduced polypeptides encouraged by the *Bacillus subtilis sacU* locus share homology with two-component sensor-regulator systems. J. Bacteriol. **170**:5093–5101.
107. **Kunst, F., M. Pascal, J. Lepesant-Kejzlarova, J.-A. Lepesant, A. Billault, and R. Dedonder.** 1974. Pleiotropic mutations affecting sporulation conditions and the synthesis of extracellular enzymes in *Bacillus subtilis* 168. Biochimie **56**:1481–1489.
108. **Lacks, S.** 1962. Molecular fate of DNA in genetic transformation of *Pneumococcus*. J. Mol. Biol. **5**:119–131.
109. **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.
110. **Lacks, S., B. Greenberg, and K. Carlson.** 1967. Fate of donor DNA in pneumococcal transformation. J. Mol. Biol. **29**:327–347.
111. **Lacks, S., B. Greenberg, and M. Neuberger.** 1975. Identification of a deoxyribonuclease implicated in genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **123**:222–232.
112. **Lacks, S., and M. Neuberger.** 1975. Membrane location of a deoxyribonuclease implicated in the genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **124**:1321–1329.
113. **Lacks, S. A., and B. Greenberg.** 1973. Competence for deoxyribonucleic acid uptake and deoxyribonuclease action external to cells in the genetic transformation of *Diplococcus pneumoniae*. J. Bacteriol. **144**:152–163.
114. **Leighton, T. (University of California.)** 1991. Personal communication.
115. **Levin, B. R.** 1988. The evolution of sex in bacteria, p. 194–211. *In* R. E. Michod and B. R. Levin (ed.), The evolution of sex. Sinauer Associates, Inc., Sunderland, Mass.
116. **Levine, J. S., and N. Strauss.** 1965. Lag period characterizing the entry of transforming deoxyribonucleic acid into *Bacillus subtilis*. J. Bacteriol. **89**:281–287.
117. **Linn, T., R. Losick, and A. L. Sonenshein.** 1975. Rifampin resistance mutation of *Bacillus subtilis* altering the electrophoretic mobility of the beta subunit of ribonucleic acid polymerase. J. Bacteriol. **122**:1387–1390.
118. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
119. **Lipmann, F.** 1980. Bacterial production of antibiotic polypeptides by thiol-linked synthesis on protein templates. Adv. Microb. Physiol. **21**:227–266.
120. **Lopez, A., C. Clavé, R. Capeyrou, V. Lafontan, and M.-C. Trombe.** 1989. Ionic and energetic changes at competence in the naturally transfromable bacterium *Streptococcus pneumoniae*. J. Gen. Microbiol. **135**:2189–2197.
121. **Love, P. E., M. J. Lyle, and R. E. Yasbin.** 1985. DNA-damage-inducible (*din*) loci are transcriptionally activated in competent *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **82**:6201–6205.
122. **Love, P. E., and R. E. Yasbin.** 1986. Induction of the *Bacillus subtilis* SOS-like response by *Escherichia coli* RecA protein. Proc. Natl. Acad. Sci. USA **83**:5204–5208.
123. **Loveday, K. S.** 1978. DNA synthesis in competent *Bacillus subtilis* cells. J. Bacteriol. **135**:1158–1161.
124. **Lovett, C. M., Jr., P. E. Love, and R. E. Yasbin.** 1989. Competence-specific induction of the *Bacillus subtilis* RecA protein analog: evidence for dual regulation of a recombination protein. J. Bacteriol. **171**:2318–2322.
125. **Maloney, P. C.** 1987. Coupling to an energized membrane: role of ion-motive gradients in the transduction of metabolic energy, p. 222–243. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
126. **Mandic-Mulec, I., N. K. Gaur, and I. Smith. (Public Health Research Institute.)** 1991. Personal communication.
127. **Marahiel, M. A., M. Krause, and H. Sharpeid.** 1985. Cloning of the tyrocidine synthetase I gene from *Bacillus brevis* and its expression in *Escherichia coli*. Mol. Gen. Genet. **201**:231–236.
128. **Mastromei, G., C. Barberio, S. Pistolesi, and M. Poisinelli.** 1989. Isolation of transformation-deficient mutants and mapping of competence genes. Genet. Res. (Cambridge) **54**:1–5.
129. **McCarthy, C., and E. W. Nester.** 1967. Macromolecular synthesis in newly transformed cells of *Bacillus subtilis*. J. Bacteriol. **94**:131–140.
130. **McCarthy, D.** 1989. Cloning of the *rec-2* locus of *Haemophilus influenzae*. Gene **75**:135–143.
131. **Mejean, V., and J. Claverys.** 1988. Polarity of DNA entry in transformation of *Streptococcus pneumoniae*. Mol. Gen. Genet. **213**:444–448.
132. **Meyer, T. F., E. Billyard, R. Haas, S. Storzbach, and M. So.** 1984. Pilus genes of *Neisseria gonorrhoeae*: chromosomal organization and DNA sequence. Proc. Natl. Acad. Sci. USA **81**:6110–6114.
133. **Michel, B., B. Niaudet, and S. D. Ehrlich.** 1983. Intermolecular recombination during transformation of *Bacillus subtilis* competent cells by monomeric and dimeric plasmids. Plasmid **10**:1–10.
134. **Michod, R. E., M. F. Wojciechowski, and M. A. Hoelzer.** 1988. DNA repair and the evolution of transformation in the bacterium *Bacillus subtilis*. Genetics **118**:31–39.
135. **Mittenhuber, G., R. Weckerman, and M. A. Marahiel.** 1989. Gene cluster containing the genes for tyrocidine synthetases 1 and 2 from *Bacillus brevis*: evidence for an operon. J. Bacteriol. **171**:4881–4887.
136. **Model, P., and M. Russel.** 1988. Filamentous bacteriophage, p. 375–456. *In* R. Calendar (ed.), The bacteriophages. Plenum Press, New York.
137. **Mohan, S., J. Aghion, N. Guillen, and D. Dubnau.** 1989. Molecular cloning and characterization of *comC*, a late competence gene of *Bacillus subtilis*. J. Bacteriol. **171**:6043–6051.
138. **Mohan, S., and D. Dubnau.** 1990. Transcriptional regulation of *comC*: evidence for a competence-specific factor in *Bacillus subtilis*. J. Bacteriol. **172**:4064–4071.
139. **Morrison, D. A., and M. F. Baker.** 1979. Competence for genetic transformation in *Pneumococcus* depends on synthesis of a small set of proteins. Nature (London) **282**:215–217.
140. **Morrison, D. A., and W. R. Guild.** 1973. Breakage prior to entry of donor DNA in pneumococcus transformation. Biochim. Biophys. Acta **299**:545–556.
141. **Morrison, D. A., and W. R. Guild.** 1973. Structure of deoxyribonucleic acid on the cell surface during uptake by *Pneumococcus*. J. Bacteriol. **115**:1055–1062.
142. **Morrison, D. A., M. C. Trombe, M. K. Hayden, G. A. Waszak, and J. D. Chen.** 1984. Isolation of transformation-deficient *Streptococcus pneumoniae* mutants defective in control of competence, using insertion duplication mutagenesis with the erythromycin resistance determinant of pAMB1. J. Bacteriol. **159**:870–876.
143. **Msadek, T., and F. Kunst. (Pasteur Institute.)** 1991. Personal communication.
144. **Msadek, T., F. Kunst, D. Henner, A. Klier, G. Rapoport, and R. Dedonder.** 1990. The signal transduction pathway controlling the synthesis of a class of degradative enzymes in *Bacillus*

subtilis: expression of the regulatory genes and analysis of mutations in *degS* and *degU*. J. Bacteriol. **172**:824–834.
145. **Msadek, T., F. Kunst, A. Klier, and G. Rapoport.** 1991. DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *Bacillus subtilis* pleiotropic regulatory gene *degQ*. J. Bacteriol. **173**:2366–2377.
146. **Mueller, J. P., and A. L. Sonenshein. (Tufts University.)** 1991. Personal communication.
147. **Mukai, K., M. Kawata, and T. Tanaka.** 1990. Isolation and phosphorylation of the *Bacillus subtilis degS* and *degU* gene products. J. Biol. Chem. **265**:20000–20006.
148. **Nakano, M. M., R. Magnusson, A. Myers, J. Curry, A. D. Grossman, and P. Zuber.** 1991. *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. J. Bacteriol. **173**:1770–1778.
149. **Nakano, M. M., M. A. Marahiel, and P. Zuber.** 1988. Identification of a genetic locus required for the biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. J. Bacteriol. **170**:5662–5668.
150. **Nakano, M. M., and P. Zuber.** 1989. Cloning and characterization of *srfB*, a regulatory gene involved in surfactin production and competence in *Bacillus subtilis*. J. Bacteriol. **171**: 5347–5353.
151. **Nakano, M. M., and P. Zuber. (Louisiana University.)** 1991. Personal communication.
152. **Nester, E. W., and B. A. D. Stocker.** 1963. Biosynthetic latency in early stages of deoxyribonucleic acid transformation in *Bacillus subtilis*. J. Bacteriol. **85**:785–796.
153. **Nohno, T., S. Noji, S. Taniguchi, and T. Saito.** 1989. The *narX* and *narL* genes encoding the nitrate-sensing regulators of *Escherichia coli* are homologous to a family of prokaryotic two-component regulatory genes. Nucleic Acids Res. **17**:2947–2957.
154. **Nunn, D., S. Bergman, and S. Lory.** 1990. Products of three accessory genes, *pilB*, *pilC*, and *pilD*. J. Bacteriol. **172**:2911–2919.
155. **Perego, M., S. P. Cole, D. Burbulys, K. Trach, and J. A. Hoch.** 1989. Characterization of the gene for a protein kinase which phosphorylates the sporulation-regulatory proteins Spo0A and Spo0F of *Bacillus subtilis*. J. Bacteriol. **171**:6187–6196.
156. **Perego, M., C. F. Higgins, S. R. Pearce, M. P. Gallagher, and J. A. Hoch.** 1991. The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. Mol. Microbiol. **5**:173–185.
157. **Perego, M., G. B. Spiegelman, and J. A. Hoch.** 1988. Structure of the gene for the transition state regulator *abrB*: regulator synthesis is controlled by the *spo0A* sporulation gene in *Bacillus subtilis*. Mol. Microbiol. **2**:689–699.
158. **Piechowska, M., and M. S. Fox.** 1971. Fate of transforming deoxyribonucleate in *Bacillus subtilis*. J. Bacteriol. **108**:680–689.
159. **Piggot, P., and J. G. Coote.** 1976. Genetic aspects of bacterial endospore formation. Bacteriol. Rev. **40**:908–962.
160. **Poindexter, J., and D. Dubnau.** Unpublished data.
161. **Pugsley, A., C. d'Enfert, I. Reyss, and M. G. Kornacker.** 1990. Genetics of extracellular protein secretion by gram-negative bacteria. Annu. Rev. Genet. **24**:67–90.
162. **Pugsley, A. P., and I. Reyss.** 1990. Five genes at the 3′ end of the *Klebsiella pneumoniae pulC* operon are required for pullulanase secretion. Mol. Microbiol. **4**:365–379.
163. **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.
164. **Rao, N. N., and A. Torriani.** 1990. Molecular aspects of phosphate transport in *Escherichia coli*. Mol. Microbiol. **4**:1083–1090.
165. **Redfield, R. J.** 1988. Evolution of bacterial transformation: Is sex with dead cells ever better than no sex at all? Genetics **119**:213–221.
166. **Reusch, R. N., T. W. Hiske, and H. L. Sadoff.** 1986. Poly-β-hydroxybutyrate membrane structure and its relationship to genetic transformability in *Escherichia coli*. J. Bacteriol. **168**: 553–562.
167. **Reusch, R. N., and H. L. Sadoff.** 1983. D-(−)-Poly-β-hydroxybutyrate in membranes of genetically competent bacteria. J. Bacteriol. **156**:778–788.
168. **Reusch, R. N., and H. L. Sadoff.** 1988. Putative structure and functions of a poly-β-hydroxybutyrate/calcium phosphate channel in bacterial plasma membranes. Proc. Natl. Acad. Sci. USA **85**:4176–4180.
169. **Reyss, I., and A. P. Pugsley.** 1990. Five additional genes in the *pulC-O* operon of the gram-negative bacterium *Klebsiella oxytoca* UNF5023 which are required for pullulanase secretion. Mol. Gen. Genet. **222**:176–184.
170. **Roggiani, M., and D. Dubnau.** Unpublished data.
171. **Roggiani, M., J. Hahn, and D. Dubnau.** 1990. Suppression of early competence mutations in *Bacillus subtilis* by *mec* mutations. J. Bacteriol. **172**:4056–4063.
172. **Roggiani, M., L. Kong, and D. Dubnau.** Unpublished data.
173. **Rosenthal, A. L., and S. D. Lacks.** 1980. Complex structure of the membrane nuclease of *Streptococcus pneumoniae* revealed by two dimensional electrophoresis. J. Mol. Biol. **141**:133–146.
174. **Rudner, D. Z., J. R. LeDeaux, K. Ireton, and A. D. Grossman.** 1991. The *spo0K* locus of *Bacillus subtilis* is homologous to the oligopeptide permease locus and is required for sporulation and competence. J. Bacteriol. **173**:1388–1398.
175. **Russel, M., and P. Model.** 1982. Filamentous phage pre-coat is an integral membrane protein: analysis by a new method of membrane preparation. Cell **28**:177–184.
176. **Sadaie, Y.** 1989. Molecular cloning of a *Bacillus subtilis* gene involved in cell division, sporulation, and exoenzyme secretion. Jpn. J. Genet. **64**:111–119.
177. **Sadaie, Y. (National Institute of Genetics.)** 1991. Personal communication.
178. **Sadaie, Y., and T. Kada.** 1983. Effect of spetum-initiation mutations on sporulation and competent cell formation in *Bacillus subtilis*. Mol. Gen. Genet. **190**:176–178.
179. **Sadaie, Y., and T. Kada.** 1983. Formation of competent *Bacillus subtilis* cells. J. Bacteriol. **153**:813–821.
180. **Sadaie, Y., and T. Kada.** 1985. *Bacillus subtilis* gene involved in cell division, sporulation, and exoenzyme secretion. J. Bacteriol. **163**:648–653.
181. **Sadaie, Y., H. Takamatsu, K. Nakamura, and K. Yamane.** 1991. Sequencing reveals similarity of the wild type *div*$^+$ gene of *Bacillus subtilis* to the *Escherichia coli secA* gene. Gene **98**:101–105.
182. **Sastry, P. A., B. L. Pasloske, W. Paranchych, J. R. Pearlstone, and L. B. Smillie.** 1985. Comparative studies of the amino acid and nucleotide sequences of pilin derived from *Pseudomonas aeruginosa* PAK and PAO. J. Bacteriol. **164**:571–577.
183. **Schatz, P., and J. Beckwith.** 1990. Genetic analysis of protein export in *Escherichia coli*. Annu. Rev. Genet. **24**:215–248.
184. **Seifert, H. S., R. S. Ajioka, C. Marchal, P. F. Sparling, and M. So.** 1988. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. Nature (London) **336**:392–395.
185. **Sharma, S., T. F. Stark, W. G. Beattie, and R. E. Moses.** 1986. Multiple control elements for the *uvrC* gene unit of *Escherichia coli*. Nucleic Acids. Res. **14**:2301–2318.
186. **Shirasu, K., P. Morel, and C. I. Kado.** 1990. Characterization of the *virB* operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis. Mol. Microbiol. **4**:1153–1163.
187. **Singh, R. N.** 1972. Number of deoxyribonucleic acid uptake sites in competent cells of *Bacillus subtilis*. J. Bacteriol. **110**:266–272.
188. **Singh, R. N., and M. P. Pitale.** 1967. Enrichment of *Bacillus subtilis* transformants by zonal centrifugation. Nature (London) **213**:1262–1263.
189. **Sisco, K. L., and H. O. Smith.** 1979. Sequence-specific DNA uptake in *Haemophilus* transformation. Proc. Natl. Acad. Sci. USA **76**:972–976.
190. **Slauch, J. M., S. Garrett, D. E. Jackson, and T. J. Silhavy.** 1988. *envZ* functions through *ompR* to control porin gene expression in *Escherichia coli* K-12. J. Bacteriol. **170**:439–441.

191. **Smith, H., D. B. Danner, and R. A. Deich.** 1981. Genetic transformation. Annu. Rev. Biochem. **50**:41–68.
192. **Smith, H., K. Wiersma, S. Bron, and G. Venema.** 1983. Transformation in *Bacillus subtilis*: purification and partial characterization of a membrane-bound DNA-binding protein. J. Bacteriol. **156**:101–108.
193. **Smith, H., K. Wiersma, S. Bron, and G. Venema.** 1984. Transformation in *Bacillus subtilis*: a 75,000-dalton protein complex is involved in binding and entry of donor DNA. J. Bacteriol. **157**:733–738.
194. **Smith, H., K. Wiersma, G. Venema, and S. Bron.** 1985. Transformation in *Bacillus subtilis*: further characterization of a 75,000-dalton protein complex involved in binding and entry of donor DNA. J. Bacteriol. **164**:201–206.
195. **Soltyk, A., D. Shugar, and M. Piechowska.** 1975. Heterologous deoxyribonucleic acid uptake and complexing with cellular constituents in competent *Bacillus subtilis*. J. Bacteriol. **124**: 1429–1438.
196. **Somma, S., and M. Polsinelli.** 1970. Quantitative autoradiographic study of competence and deoxyribonucleic acid incorporation in *Bacillus subtilis*. J. Bacteriol. **101**:851–855.
197. **Sonenshein, A. L., H. B. Alexander, D. M. Rothstein, and S. H. Fisher.** 1977. Lipiarmycin-resistant ribonucleic acid polymerase mutants of *Bacillus subtilis*. J. Bacteriol. **132**:73–79.
198. **Sparling, P. F.** 1966. Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. J. Bacteriol. **92**:1364–1371.
199. **Spatz, C. H., and T. A. Trautner.** 1970. One way to do experiments on gene conversion? Transfection with heteroduplex SPP1 DNA. Mol. Gen. Genet. **109**:84–106.
200. **Spizizen, J.** 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. USA **44**:1072–1078.
201. **Spizizen, J.** 1965. Analysis of asporogenic mutants in *Bacillus subtilis* by genetic transformation, p. 125–137. *In* L. L. Campbell, and H. O. Halvorsen (ed.), Spores III. American Society for Microbiology, Washington, D.C.
202. **Steinmetz, M., F. Kunst, and R. Dedonder.** 1976. Mapping of mutations affecting synthesis of exocellular enzymes in *Bacillus subtilis*. Mol. Gen. Genet. **148**:281–285.
203. **Stewart, G. J., and C. A. Carlson.** 1986. The biology of natural transformation. Ann. Rev. Microbiol. **40**:211–235.
204. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and the regulation of adaptive responses in bacteria. Microbiol. Rev. **53**:450–490.
205. **Stout, V., and S. Gottesman.** 1990. RcsB and RscC: a two-component regulator of capsule synthesis in *Escherichia coli*. J. Bacteriol. **172**:659–669.
206. **Strauch, M. A., G. B. Spiegelman, M. Perego, W. C. Johnson, D. Burbulys, and J. A. Hoch.** 1989. The transition state transcription regulator *abrB* of *Bacillus subtilis* is a DNA binding protein. EMBO J. **8**:1615–1621.
207. **Strauss, N.** 1965. Configuration of transforming deoxyribonucleic acid during entry into *Bacillus subtilis*. J. Bacteriol. **89**:288–293.
208. **Strauss, N.** 1966. Further evidence concerning the configuration of transforming deoxyribonucleic acid during entry into *Bacillus subtilis*. J. Bacteriol. **91**:702–708.
209. **Strauss, N.** 1970. Early energy-dependent step in the entry of transforming deoxyribonucleic acid. J. Bacteriol. **101**:35–37.
210. **Strauss, N.** 1970. Transformation of *Bacillus subtilis* using hybrid DNA molecules constructed by annealing resolved complementary strands. Genetics **66**:583–593.
211. **Strom, M. S., and S. Lory.** 1991. Amino acid substitutes in pilin of *Pseudomonas aeruginosa*. Effect on leader peptide cleavage, amino terminal methylation, and pilus assembly. J. Biol. Chem. **266**:1656–1664.
212. **Strom, M. S., D. Nunn, and S. Lory.** 1991. Multiple roles of the pilus biogenesis protein PilD: involvement of PilD in excretion of enzymes from *Pseudomonas aeruginosa*. J. Bacteriol. **173**: 1175–1180.
213. **Tanaka, T., and M. Kawata.** 1988. Cloning and characterization of *Bacillus subtilis iep*, which has positive and negative effects on production of extracellular proteases. J. Bacteriol. **170**:3593–3600.
214. **Tennent, J. M., F. Lindberg, and S. Normark.** 1990. Integrity of *Escherichia coli* P pili during biogenesis: properties and role of PapJ. Mol. Microbiol. **4**:747–758.
215. **Tieng, E., M. Roggiani, A. Ninfa, and D. Dubnau.** Unpublished data.
216. **Tomasz, A.** 1966. Model for the mechanism controlling the competent state in *Pneumococcus* cultures. J. Bacteriol. **91**: 1050–1061.
217. **Tomasz, A., and R. D. Hotchkiss.** 1964. Regulation of the transformability of pneumococcal cell cultures by macromolecular cell products. Proc. Natl. Acad. Sci. USA **51**:480–486.
218. **Tomasz, A., and J. L. Mosser.** 1966. On the nature of the pneumococcal activator substance. Proc. Natl. Acad. Sci. USA **55**:58–66.
219. **Tombs, J.-F., G. J. Barcak, M. S. Chandler, R. J. Redfield, and H. O. Smith.** 1989. Transposon mutagenesis, characterization, and cloning of transformation genes of *Haemophilus influenzae* Rd. J. Bacteriol. **171**:3796–3802.
220. **Trach, K., D. Burbulys, J. Wu, R. Jonas, J. Day, P. Kallio, M. Strauch, M. Perego, G. Spiegelman, C. Fogher, and J. Hoch.** Control of the initiation of sporulation, p. 13. Abstr. Int. Conf. *Bacillus subtilis* Genome.
221. **Vagner, V., J.-P. Claverys, S. D. Ehrlich, and V. Mejean.** 1990. Direction of DNA entry in competent cells of *Bacillus subtilis*. Mol. Microbiol. **4**:1785–1788.
222. **van Nieuwenhoven, M. H., K. J. Hellingwerf, G. Venema, and W. N. Konings.** 1982. Role of proton motive force in genetic transformation of *Bacillus subtilis*. J. Bacteriol. **151**:771–776.
223. **van Sinderen, D., and G. Grandi. (University of Groningen.)** 1991. Personal communication.
224. **van Sinderen, D., and G. Venema. (University of Groningen.)** 1991. Personal communication.
225. **van Sinderen, D., S. Withoff, H. Boels, and G. Venema.** 1990. Isolation and characterization of *comL*, a transcription unit involved in competence development in *Bacillus subtilis*. Mol. Gen. Genet. **224**:396–404.
226. **Vermuelen, C. A., and G. Venema.** 1974. Electron microscope and autoradiographic study of ultrastructural aspects of competence and deoxyribonucleic acid absorption in *Bacillus subtilis*: ultrastructure of competent and noncompetent cells and cellular changes during development of competence. J. Bacteriol. **118**:334–341.
227. **Vijayakumar, M. N., and D. A. Morrison.** 1986. Localization of competence-induced proteins in *Streptococcus pneumoniae*. J. Bacteriol. **165**:689–695.
228. **Vosman, B., G. Kuiken, and G. Venema.** 1988. Transformation in *Bacillus subtilis*: involvement of the 17-kilodalton DNA entry nuclease and the competence-specific 18-kilodalton protein. J. Bacteriol. **170**:3703–3710.
229. **Walker, G. C.** 1985. Inducible DNA repair systems. Annu. Rev. Biochem. **54**:425–457.
230. **Walker, G. C.** 1987. The SOS response of *Escherichia coli*, p. 1346–1357. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
231. **Ward, J. E., D. E. Akiyoshi, D. Regier, A. Datta, M. P. Gordon, and E. W. Nester.** 1988. Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. J. Biol. Chem. **263**:5804–5814. (Erratum, **265**:4768, 1990.)
232. **Watts, T. H., C. M. Kay, and W. Paranchych.** 1983. Spectral properties of three quaternary arrangement of *Pseudomonas* pilin. Biochemistry **22**:3640–3646.
233. **Weckerman, R., R. Furbass, and M. A. Marahiel.** 1988. Complete nucleotide sequence of the *tycA* gene coding the tyrocidine synthetase I from *Bacillus brevis*. Nucleic Acids Res. **16**:11841.
234. **Weinrauch, Y., and D. Dubnau.** Unpublished data.
235. **Weinrauch, Y., and D. Dubnau.** 1983. Plasmid marker rescue transformation in *Bacillus subtilis*. J. Bacteriol. **154**:1077–1087.

236. **Weinrauch, Y., and D. Dubnau.** 1987. Plasmid marker rescue transformation proceeds by breakage-reunion in *Bacillus subtilis*. J. Bacteriol. **169**:1205–1211.
237. **Weinrauch, Y., N. Guillen, and D. Dubnau.** 1989. Sequence and transcription mapping of *comB* and of *comA*, a *Bacillus subtilis* competence gene that exhibits similarity to members of the two-component procaryotic regulatory system. J. Bacteriol. **171**:5362–5375.
238. **Weinrauch, Y., T. Msadek, F. Kunst, and D. Dubnau.** Unpublished data.
239. **Weinrauch, Y., R. Penchev, E. Dubnau, I. Smith, and D. Dubnau.** 1990. A *Bacillus subtilis* regulatory gene product for genetic competence and sporulation resembles sensor protein members of the bacterial two-component signal-transduction systems. Genes Dev. **4**:860–872.
240. **Weston, L. A., and R. J. Kadner.** 1987. Identification of UhP polypeptides and evidence for their role in exogenous induction of the sugar phosphate transport system of *Escherichia coli* K-12. J. Bacteriol. **169**:3546–3555.
241. **Whitchurch, C. B., M. Hobbs, S. P. Livingston, V. Krishnapillai, and J. S. Mattick.** 1991. Characterization of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. Gene, in press.
242. **Wilson, G. A., and K. Bott.** 1968. Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system. J. Bacteriol. **95**:1439–1449.
243. **Wojciechowski, M. F., M. A. Hoelzer, and R. E. Michod.** 1989. DNA repair and the evolution of transformation in *Bacillus subtilis*. II. Role of inducible repair. Genetics **121**:411–422.
244. **Yamashita, S., F. Kawamura, H. Yoshikawa, H. Takahashi, Y. Kobayashi, and H. Saito.** 1989. Dissection of the expression signals of the *spo0A* gene of *Bacillus subtilis*: glucose represses sporulation-specific expression. J. Gen. Microbiol. **135**:1335–1345.
245. **Yasbin, R., J. Jackson, P. Love, and R. Marrero.** 1988. Dual regulation of the *recE* gene, p. 109–114. *In* A. T. Ganesan and J. A. Hoch (ed.), Genetics and biotechnology of bacilli, vol. 2. Academic Press, Inc., San Diego, Calif.
246. **Youngman, P., J. B. Perkins, and R. Losick.** 1984. A novel method for the rapid cloning in *Escherichia coli* of *Bacillus subtilis* chromosomal DNA adjacent to Tn917 insertions. Mol. Gen. Genet. **195**:424–433.
247. **Youngman, P., J. B. Perkins, and K. Sandman.** 1985. Use of Tn*917*-mediated transcriptional gene fusions to *lacZ* and *cat-86* for the identification and study of *spo* genes in *Bacillus subtilis*, p. 47–54. *In* J. A. Hoch and P. Setlow (ed.), Molecular biology of microbial differentiation. American Society for Microbiology, Washington, D.C.
248. **Youngman, P., P. Zuber, J. B. Perkins, K. Sandman, M. Igo, and R. Losick.** 1985. New ways to study developmental genes in spore-forming bacteria. Science **228**:285–291.