CLONING OF, AND STUDIES ON, GENES CODING FOR SUBTILISINS CARLSBERG AND BPN'

by

Myra Freda Jacobs B.Sc. (Hons.)

This thesis is presented for the degree of

Doctor of Philosophy

Department of Biological Sciences

University of Warwick

England

October 1986

TEXT
CUT OFF IN THE
ORIGINAL

PAGE NUMBERS CUT OFF IN THE ORIGINAL

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

LIST OF TABLES

ACKNOWLEGEMENTS

I would first of all like to thank Niels Fiil, who on Novo's behalf, arranged support for this study, and made it possible for me to complete this thesis.

In connection with the move to Denmark in 1983 that this entailed, I wish to thank Ulla Rinus for practical help, and Lasse Christiansen and Helle Wøldike for patient coaching in Danish vowels. My less than perfect diction is not due to want of effort on their part.

I am indebted to Jan-Ingmar Flock for having introduced me to Bacillus during the time that we worked together in the former Searle Research Laboratories, High Wycombe, and for having borne the difficult task of supervising me from distant Sweden with such enthusiasm. My thanks also go to Nick Mann for helpful criticism regarding this thesis and for smoothing out administrative difficulties.

I wish to thank Kirsten Lundsgaard Jensen, Gitte Bak Poulsen and Thea Gartner for much appreciated occasional technical help, Lise Damkjær and Karen Randrup especially for expert assistance repectively in Western Blotting and cDNA synthesis, and Northern Blotting. I wish to extend my thanks to the whole of Genetisk Laboratorium for creating a pleasant working environment.

I am especially grateful to Tove Christensen and Sven Hastrup for their unfailing willingness to discuss results, and for their sensible suggestions. Helle Wøldike has my thanks too for moderating my wilder speculations with similarly sound advice.

Many thanks go to Ulla Rinus and Marit Petersen for secretarial assistance in preparing the thesis.

Lastly, I wish to thank my husband Michael Fuchs for his forbearance, for doing more than his share of the household chores for so long, and for emerging from this thesis with still only the vaguest notion of what a gene is.

MYRA JACOBS
October, 1986.

DECLARATION

I declare that this thesis has been composed by myself, and has not been used in any previous application for a degree.

The results presented are an account of my own work, carried out under the supervision of Dr. J-I. Flock in High Wycombe, England, and Bagsværd, Denmark.

Where experiments were carried out in collaboration with other workers, their contribution is acknowledged in the text.

Myra Jacobs

October 1986

ABBREVIATIONS

2^{o}	secondary
3D	three-dimensional
AA	amino acid
Ap	ampicillin
APase	alkaline phosphatase
APR, *apr*	alkaline serine protease, and gene encoding it, respectively
APS	ammonium persulphate
BGSC	Bacillus Genetic Stock Centre, Ohio State University, USA
BPB	bromophenol blue
BSA	bovine serum albumin
BAEE	benzoylarginine ethyl ester
c.	circa, approx.
cAMP	cyclic AMP
CAP	catabolite activator protein
CAT	chloramphenicol transacetylase
cDNA	complementary DNA
Cm	chloramphenicol
Cm^{200}	chloramphenicol 200 ug/ml
CRP	same as CAP
DFP	diisopropylphosphofluoride
DIP	diisopropylphosphoric ester
DTT	dithiothreitol
E	RNA polymerase core enzyme = $(\alpha_2\beta\beta')$
EDTA	ethylenediamine tetraacetate
EGTA	ethyleneglycol-bis(β-amino-ethyl ether) N,N'-tetra-acetic acid
ELISA	enzyme-linked immune-specific assay
Emy	erythromycin
FAGLA	furylacryloyl-glycyl-L-leucine amide
glnA	glutamine synthetase gene
gltA	glutamate synthase gene
h	hours
IMP	inosine monophosphate
IPTG	isopropyl-β-D-thiogalactoside
IS	Insertion Sequence
ISP	Intracellular Serine Protease
kD, kDa	kilodaltons
kb, kbp	kilobases, kilobase pairs
Km	kanamycin
K_m/k_{cat}	Michaelis/catalytic rate constant
LB	L Broth
mins.	minutes
MOPS	3(N-morpholino)propanesulfonic acid
MUX	4-methyl-umbelliferyl-β-D-xylopyranoside
NCIB	National Collection of Industrial Bacteria, Torry Research Station, Aberdeen, Scotland

NMP,NDP,NTP	nucleoside mono-, di-, triphosphate
NPR, *npr*	neutral metalloprotease, and gene encoding it, respectively
nt	nucleotide
ORF	open reading frame
pfu	plaque forming units
PIPES	piperazine-N,N'-bis(2-ethane-sulfonic acid)
PMSF	phenylmethylsulfonyl fluoride
PNPX	p-nitro-phenyl-β-D-xyloside
Prt^+	exoprotease-producing phenotype
PVP	polyvinyl pyrrolidone
RBS	ribosome binding site
SD	Shine-Dalgarno region
Spo^+	sporulation-positive phenotype
T_o or To	time marking the cessation of exponential growth
T1, T2 etc	1, 2 hours after T_o
TAME	p-tosyl-L-arginine methyl ester
TCA	trichloroacetic acid
Tet	tetracycline
TRIS	tris(hydroxymethyl)aminomethane
tt, ts	test tube, temperature sensitive
UV	ultraviolet (light)
wt	wild-type
X, x	times
XCFF	xylene cyanol dye
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
Xyl^+	xylosidase-positive phenotype
xynB	gene encoding xylosidase

SUMMARY

The alkaline serine proteases (subtilisins) secreted by the *Bacillus* spp. can be categorized into two groups on the basis of immunological cross-reactions. A gene coding for the Carlsberg type was cloned from *B. licheniformis* (III), and one for the BPN'/NOVO type, from *B. subtilis* (IV).

The subtilisin BPN' gene was partially characterized . The sequence and expression data obtained correspond well with those observed by others. *In vitro* substitution of part of the coding sequence with the gene encoding chloramphenicol transacetylase was followed by targeted integration into the *Bacillus* chromosome (IV.6). This resulted in a mutant strain having only c. 30% exoproteolytic activity compared with wild-type, due specifically to decreased serine protease secretion. Immunologically, and by gel electrophoresis, no subtilisin could be detected. It was shown that the deletion was readily transferred to closely related strains by selection for chloramphenicol resistance. In combination with other mutations, the subtilisin deletion could form the basis for a better *Bacillus* host for the synthesis of foreign proteins.

The subtilisin Carlsberg gene was cloned and characterized. The sequence data obtained (II.6) allowed, for the first time, a comparison of the entire precursor coding sequence with those available for enzymes of the BPN' type (Appendix 1). Overproduction of the gene product in *B. subtilis* was achieved, both from the endogenous promoter (III.7.5), and from two heterologous promoters (III.7.2.2). It was demonstrated by various techniques that the extra exoprotease production was due specifically to synthesis of the Carlsberg enzyme. Results with the heterologous promoters pointed to a significant delay before the appearance of subtilisin in the medium, probably due to post-transcriptional factors (III.7.3). Using mRNA extracted from *B. licheniformis*, the transcriptional terminator was located precisely (III.8.3). Transcriptional initiation site determination was attempted by two different procedures. Two sites were approximated by the S1 mapping technique (III.8.2). In contrast, cDNA synthesis and sequencing pointed to a single start site (Appendix 3). The relation of these sites to identifiable elements in the sequences up- and downstream from the gene is discussed.

The Carlsberg gene was cloned only with difficulty, unlike the homologous gene for subtilisin BPN'. Possible reasons for this, and alternative cloning strategies, are discussed (III.9; V).

Chapter I: INTRODUCTION

I.1 General Aspects

The cloning of subtilisin genes was undertaken because of the many fields of interest that would be open to exploration once the genes were to hand. The topics discussed in this Introduction give some idea of the range of different aspects of physiological , biochemical and biotechnological interest that a study of the subtilisin gene offers.

Quite apart from the potential to exploit the genes for purposes of enzyme production or the synthesis of foreign products, elucidation of the complex regulation of the subtilisin genes will continue to fascinate molecular biologists for some time to come.

I.1.1 Industrial Importance of the Bacilli

The genus Bacillus, which comprises a heterogenous group of Gram-positive, aerobic, spore forming, soil organisms, have been widely used in the biotechnology industry. Almost all species secrete a wide variety of antibiotics and hydrolytic enzymes. The latter have properties that reflect the natural habitats, which include some environments characterized by extremes of pH and temperature.

In addition there are other aspects which have made the bacilli attractive to use in the large scale demanded by industry: most species are non-pathogenic for man and animals, the genetics of Bacillus subtilis 168 have been intensively studied and are fairly well understood, and the organisms are relatively easy to cultivate on simple substrates.

A comprehensive review of the extracellular enzymes produced by Bacillus spp. has been made (Priest, 1977), and the commercial applications of exoenzymes have also been discussed in detail (Fogarty et al, 1974a&b; Debabov, 1982). The insecticidal delta endotoxins produced by B. thuringiensis have

differential specificities for various genera, and may well find wider commercial application in the future (Haider et al, 1986; Herrnstadt et al, 1986).

The brewing and detergent industries in particular have created an enormous market for the proteases and amylases of Bacillus spp. Until recently micro-organisms for industrial application have been selected by the methods of classical microbiology viz. screening of natural isolates for desired properties, followed by extensive mutagenesis to increase yields to cost-effective levels.

I.1.2 Molecular Genetics and Industrial Microbiology

Recent advances in molecular genetics have provided more sophisticated routes to produce an enzyme with useful properties in high yield. A few of the new possibilities are listed below.

1) Cloning of structural genes and regulatory loci offers a means of improving yield by raising gene copy number, though possibly not to the levels above those obtainable by a rigorous traditional programme of mutagenesis and selection.

2) The properties of an enzyme may be changed by several routes apart from chemical modification (which was reviewed recently by Kaiser et al, 1985). Given the 3-dimensional structure of an enzyme, the technique of site-specific mutagenesis can be applied to selected regions of the corresponding cloned gene and specific codons altered ("Protein Engineering"). It has already been demonstrated that the catalytic characteristics of an enzyme can be improved in this way (Wilkinson et al, 1984).

The technique has also been applied to a subtilisin, one of a class of alkaline serine proteases secreted by most Bacillus spp. Among the subtilisins, the Carlsberg type has for some years been the enzyme produced in larger amounts than any

other worldwide (Aunstrup et al, 1979). Methionine-222, which is responsible for the enzyme activity's susceptibility to oxidation (Stauffer & Etson, 1969), was replaced with cysteine by site-specific mutagenesis to produce a more resistant and industrially useful enzyme (Estell et al, 1985). Other changes have resulted in modified subtilisin enzymes with increased specific activity ie. K_m/k_{cat} (Wells, personal communication).

3) Genes for related enzymes may be combined *in vivo* by intramolecular combination on a plasmid (Michel et al, 1982), or *in vitro* to yield chimeric enzymes having a combination of the properties of the parental gene products and possibly altered substrate specificity. By application of the *in vivo* route in *E. coli*, workers at Genentech have derived α-amylases having specific activity and thermostability intermediate between the 2 parental enzymes,which had either a high activity (ex *B. licheniformis*) or good thermostability (ex *B. stearothermophilus*) (Gray et al, 1986). This approach has also been applied to the subtilisins. Such recombination has the advantage over site-specific mutagenesis in that no detailed knowledge of the 3D structure is required, and hybrids can rapidly be generated. The screening of the properties of the hybrid enzymes is however very labour-intensive.

4) The ability of the bacilli to secrete large amounts of products into the medium, from which they can readily be recovered, has made them apparently attractive hosts for the expression of heterologous genes. So-called expression-secretion vectors, where endogenous signals for transcription, translation and export of the protein product are fused to the heterologous gene have been developed - both as episomes and as chromosomally integrated systems and are discussed below in section I.10.

I.1.3 The Process of Sporulation in the Bacilli

In response to nutrient deprivation, the bacilli enter a developmental program that culminates in the formation of an endospore. The vegetative cells cease normal cell division, and partition into the sporangium and forespore. This matures and is ultimately liberated by lysis from the mother cell as a heat-resistant spore. The sporulation process has been the subject of intensive investigation and has served as a primitive model for cell differentiation and regulated expression.

An electron microscopic study of the process in B. subtilis identified 6 cytological stages (Ryter, 1965) and was followed by an investigation of mutants, which can be isolated with a blockage at all of the stages described (Ryter et al, 1966).

The morphological changes observed are the result of extensive alterations in the biochemistry of the organism, and some of these events have been correlated temporally with the various stages in the development of the spore (Warren, 1968).

At least 30 loci affecting sporulation have been identified and mapped (Hoch, 1976; Piggot & Coote, 1976). For the purposes of the discussion that follows below of transcriptional regulation in B. subtilis, of proteases produced by the Bacillus spp., and the regulation and physiological role of the exoprotease subtilisin in particular, the spoO genes are the most relevant. SpoO mutants are arrested at stage 0, the earliest stage in endospore development: they fail to accumulate chromosomal material at one pole of the cell and do not form a septum i.e. they do not display any of the morphological changes associated with spore formation. A failure to secrete serine protease at the end of exponential growth is a genetic marker for spoO mutants (Pigott & Coote, 1976).

I.2: Diversity of Proteases Produced by Bacillus spp.

I.2.1 Characterization and Relationship to Sporulation.

After the end of exponential growth, B.subtilis and related species secrete several enzymes as a result of de novo synthesis, including proteases, RNAases and amylases (Coleman, 1967).

A temporal relationship was established early on between the maximal rate of exoprotease production and the onset of sporulation for B. licheniformis (Bernlohr, 1964), B. subtilis (Coleman, 1967), and B. cereus (Levinsohn & Aronson, 1967).

Even though several reports in the 1950s indicated that at least 2 proteolytic activities were present in the medium, many later studies assumed that a single protease was secreted (eg. Coleman 1967). However, first one (Rappaport et al, 1965), then 2 different serine protease activities were defined in the B. subtilis culture medium, one having also appreciable esterase activity (Boyer & Carlton, 1968). A neutral zinc-requiring enzyme had also been described (McConn et al, 1964) and it was later established that all three activities are typically found in culture supernatants of some Bacillus spp.(Prestidge et al, 1971; Millet, 1970 ; Hall et al, 1966), albeit with some species differences. B.licheniformis for example, produces little neutral protease; B. cereus produces levels of serine protease so low they have escaped detection in some studies (cf. Sadoff et al, 1970; Millet, 1969).

Mandlestam suggested criteria for categorizing the biochemical events leading to spore formation (Dancer & Mandlestam, 1975b): (1) the main primary "dependent" events, (2) side-effects triggered by events in the primary sequence and (3) changes in vegetative functions that occur due to changes in cultural conditions to which the cells are exposed when sporulation is induced. These latter changes are not involved

in the sporulation process.

With these criteria the relatedness of exoenzyme production to sporulation can be formulated as: whether exoenzymes are necessary for sporulation to occur, or whether certain steps of the sporulation sequence are necessary for triggering exoenzyme synthesis.

In many studies on sporulation-defective mutants , both extracellular amylase and protease production have been measured, as secretion of these is often affected simultaneously. It was soon apparent that the above questions can be answered in the negative for amylase: some spoO mutants have normal α-amylase production, but do not secrete protease, and amyE mutants sporulate normally (Schaeffer, 1969). When grown on starch, B. subtilis produces amylase throughout vegetative growth and not just post-exponentially (Welker & Campbell, 1967b). Thymidine starvation, which inhibits sporulation, inhibits exoprotease production as well as RNAase, but leaves alpha-amylase synthesis unaffected (Dancer & Mandelstam, 1975b). Amylase falls thus into Mandelstam's 3rd category.

The situation has however been far from clear-cut for the exoproteases, especially the serine protease.

The study of early-blocked sporulation mutants, defective in many characters including exoenzyme production, suggested that exoprotease production is an obligatory step in the sporulation process (Ballassa, 1969; Mandelstam & Waites, 1968; Michel & Millet, 1970).

For many years, all mutants isolated with impaired serine protease activity were also defective in sporulation (Dancer & Mandelstam, 1975b) although not all Spo^- mutants are Prt^-. Reversion to the Spo^+ phenotype was accompanied by the ability to produce exoprotease (Schaeffer,

1969), and, as mentioned above, loss of this ability serves as a genetic marker for spo0 mutations (Piggot & Coote, 1976).

Below, the various proteases that have been characterized in the bacilli are described individually. A recent review of this kind is not available elsewhere. An appreciation of the multiplicity of proteolytic activities in Bacillus is required for an evaluation of I.10.3.4.,(where the use of Bacillus as a host for heterologous gene expression is discussed), as well as my own work presented in Chapter IV.

It has recently been established unequivocally that neither of the major exoproteases is essential for sporulation.

The numerous earlier investigations addressing the question of a "dependent" role for the proteases is thus reevaluated critically. The relevant data for each individual exo- and intracellular protease are summarized under separate headings (I.2.2.4,-7,-9, I.2.3.3).

The accumulated data illustrate the complexity of post-exponential protease expression in Bacillus, and the difficulties in assigning roles to individual activities before the relevant cloned genes are to hand.

As pointed out by Priest (1977) and others, any attempt to correlate studies on Bacillus proteases must contend with the confusion in the literature attributable to several causes, not least the misclassification of strains. The use of different media and individual assays by the various workers in the field also makes correlation more difficult.

I.2.2 Extracellular Proteases

I.2.2.1 Alkaline Serine Proteases / APR (Subtilisins): Characterization and Diversity

One of the first proteinases to be isolated in pure form

from the Bacillus spp. was subtilisin Carlsberg, produced by what was erroneously classified as an asporogenous B. subtilis strain (Guntelberg, 1956; Guntelberg & Ottesen, 1956). It was characterized as having a pH optimum in the alkaline range and was inhibited by DFP.

It subsequently became clear that protease of this type is produced by both B. licheniformis and B. pumilus strains but not by B. subtilis spp. (Keay & Moser, 1969), after the end of exponential growth (Bernlohr, 1964). Closely related proteases, (subtilisins BPN', NOVO and subtilisin Amylosacchariticus) that can, however, be distinguished from the Carlsberg-type enzyme both immunologically and by virtue of having one extra amino acid residue (making a total of 275), are secreted by B. amyloliquefaciens and B. subtilis spp. The frequently used strain B. subtilis NRRLB 3411, which produces a BPN' type enzyme (Keay and Moser, 1969) is most probably a strain of B. amyloliquefaciens (Vasantha et al, 1984). The amino acid sequences of several of the enzymes mentioned above were determined by Smith and coworkers: subtilisin Carlsberg (Smith, et al 1968); subtilisin BPN' (Markland & Smith, 1967), subtilisin NOVO (Olaitan et al, 1968) (The latter is identical to BPN' both being secreted by B. amyloliquefaciens spp.(Welker and Campbell,1967a&b), and the names merely reflecting different commercial sources); and subtilisin Amylosacchariticus (Kurihara et al, 1972) which is secreted by a variety of B. subtilis.

More recently the primary amino acid sequence of subtilisin DY, which is very similar to subtilisin Carlsberg was published (Nedkov et al, 1985), as well as that of "alkaline mesentericopeptidase", which is secreted by a B. mesentericus strain, and shows striking homology especially to subtilisin amylosacchariticus (Svendsen et al, 1986). A thermostable alkaline proteinase "thermitase" made by Thermoactinomyces vulgaris has also been sequenced and this cysteine-protease has undoubted homology to the subtilisins

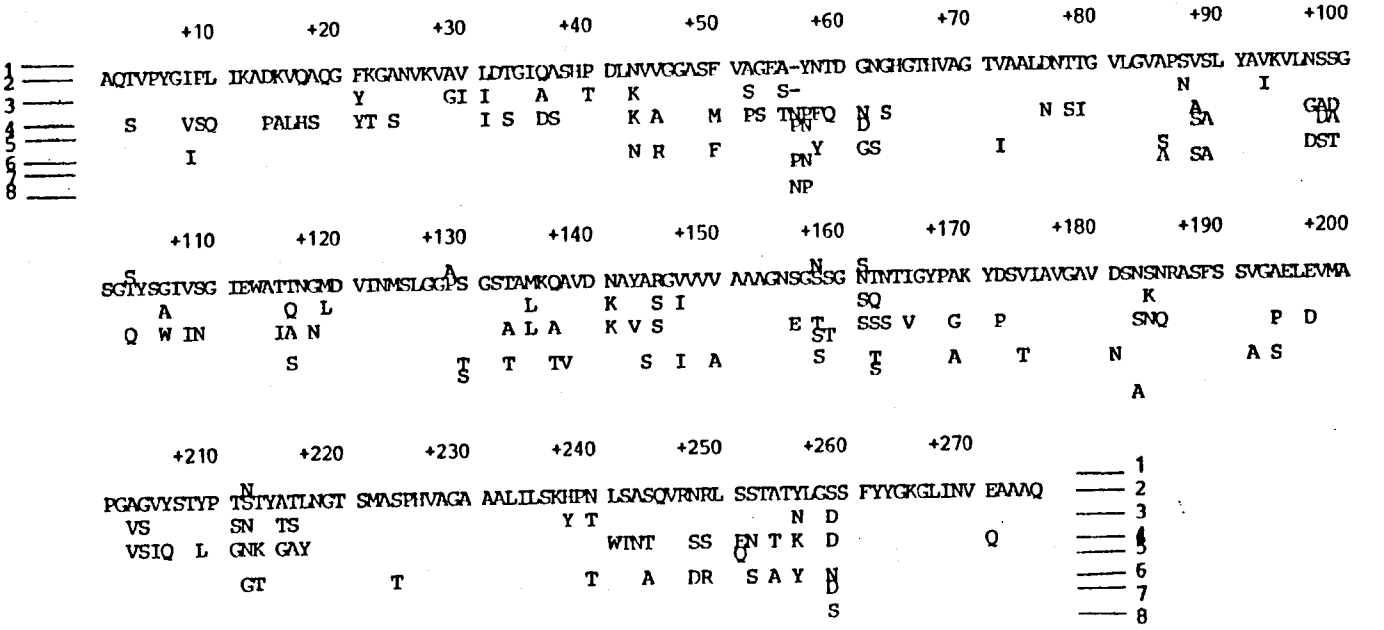
(Meloun et al, 1985).

I.2.2.2 apr Gene: 3-Dimensional Protein Structure

Genes coding for a number of subtilisins from reliably classified strains have been cloned: B. amyloliquefaciens: Wells et al, 1983; Vasantha et al, 1984; B. subtilis: Stahl and Ferrari, 1984; Wong et al, 1984; B. licheniformis: Jacobs et al, 1985. The aprE gene has been mapped and lies between glyB and metD ("sprE" Henner & Hoch, 1980; Stahl & Ferrari, 1984). In Figure 1 the amino acid sequence determined from purified enzymes and those deduced from the coding sequences of cloned genes are compared. From Figure 1 the relationships between the sequences are obvious and the provenance of the various subtilisin types is at last unambiguous.

Cloning revealed that the primary translation products of subtilisin genes have an N-terminal extension comprising a "signal sequence" and a zymogen pro-region, which are discussed in more detail in sections I.8.2 and I.9.1.

The crystal structure of subtilisin BPN' was determined by Wright et al (1969) and that of subtilisin NOVO by Drenth et al (1972). With one exception, the 84 residues of the total 275 present in BPN' that differ in subtilisin Carlsberg, occur on the molecule's exterior. Based on this fact, the 3-dimensional structure of subtilisin Carlsberg was assumed to be the same as that of BPN'. This was only recently demonstrated to in fact be the case (McPhalen et al, 1985a; Bode et al, 1986). The latest study revealed a Ca^{2+} ion within the Leu75-Thr79 loop, which is liganded to several nearby AAs. Similarly, two Ca^{2+} ions have been placed in the BPN' structure, where the second is liganded to Glu195, Asp-197 and Tyr171 (Genex, personal communication).The crystal structure of BPN' complexed to specific inhibitors has also been determined (Mitsui et al, 1977; McPhalen et al, 1985b). Within the primary sequence there are a number of repetitions, suggesting evolution from a subunit (Smith et al, 1968), but



1)Subtilisin Carlsberg AA sequence (Smith et al,1968); differences from sequence 2).
2)Deduced subtilisin Carlsberg sequence (Jacobs et al,1985).
3)Subtilisin DY AA sequence (Nedkov et al,1985); differences from sequence 2).
4)Deduced sequence of subtilisin of B. amyloliquefaciens (Wells et al,1983); differences from sequence 2).
5)Subtilisin BPN' AA sequence (Markland & Smith,1967); differences from sequence 4).
6)Deduced sequence of subtilisin of B. subtilis 168 (Stahl & Ferrari,1984); differences from sequence 4).
7)Subtilisin amylosacchariticus AA sequence (Kurihara et al,1972); differences from sequence 6).
8)Mesentericopeptidase (Svendson et al,1986) * ;differences from sequence 6).

* positions 56 &57 in AA sequences (5) and (7) were erroneously reversed in the original determinations quoted

FIGURE 1: COMPARISON OF SUBTILISIN SEQUENCES

this was not supported by an analysis of the coding sequence (Jacobs et al, 1985).

Comparisons of protein structure, physiochemical characteristics, substrate specificities and active site studies of the subtilisins have been made (Ottesen & Svendsen, 1970; Markland and Smith, 1971). The chemical modification of the subtilisins and resulting effects on substrate binding have also been reviewed (Svendsen, 1976).

I.2.2.3 Active Site Studies

The involvement of Serine-221 in the active site of subtilisin was first suggested by the observation that the enzyme is inactivated by DFP. The active serine was identified after partial acid hydrolysis of subtilisin NOVO and isolation of the ^{32}P-DIP-labelled peptide (Sanger & Shaw, 1960), and is located in a stretch of highly conserved residues (See Figure 1). As one would expect, PMSF also inhibits the enzyme strongly by reaction with Ser221 (Ottesen & Svendsen, 1970). Histidine-64 has similarly been shown to be involved at the active site by covalent attachment to a labelled inhibitor benzyloxy-carbonyl-phenylalanine bromomethyl ketone (ZPBK) (Markland et al, 1968).

A catalytic triad that is now known as the "charge relay system" has been found in all serine proteases investigated to date, including the bacterial subtilisins and the mammalian pancreatic and plasma proteases (Fersht, 1985). Apart from the Ser221 and His64 mentioned above, Asp32 has been implicated by analogy with the pancreatic enzymes (Blow et al, 1969). These three residues lie in a steric arrangement in the 3-D crystal structure consistent with the proposed catalytic mechanism, with their side groups within hydrogen-bonding distance of one another (Ser and His, His and Asp resp.), as demanded by the charge relay hypothesis. Modification of the AAs involved in the catalytic triad by site-specific mutagenesis leads to inactive gene products that

cannot be exported from the cell (Powers et al,1986).

I.2.2.4 Relationship of APR to Sporulation

Cloning of the genes for the subtilisins was a priority for many groups for several reasons, primarily (1) the possibility of altering the enzyme non-randomly both for mechanistic studies and for commercial reasons (Thomas et al, 1985 ; Wells & Powers,1986; Newswatch of 18/11/85, 6/1/86, Fox, 1985) and (2) a cloned gene offered the possibility of resolving unambiguously whether or not the serine protease is of consequence in the sporulation process, which had been the subject of controversy for many years.

Despite the existence of species of spore-formers which do not seem to secrete a serine protease, such as B. megaterium (Millet & Aubert, 1969), and which secrete very little if any exoprotease at all, such as B. brevis (Kurotsu et al, 1982), several lines of investigation helped to perpetuate the belief in subtilisin's indispensible role in sporulation.These were (a) ambiguous evidence from the pleiotropic sporulation mutants mentioned above, (b) in vivo serine protease inhibitor studies described in I.2.3.3.1 (Dancer & Mandelstam, 1975a), and (c) ts mutants inadequately characterized as regards reversion and complementation (Leighton et al, 1973).

As early as 1972, a structural gene mutation had been obtained for subtilisin, but was not mapped. Shoer and Rappaport (1972) isolated a B. subtilis WI mutant with reduced extracellular proteolytic activity that sporulated normally. A mutant protein smaller than wild-type subtilisin was produced. Labelling of the mutant protein with ^{32}P-DFP indicated that only 1 in 53 molecules was active, suggesting either a low level of suppression of a chain-terminating mutation, or else contamination by another proteinase.

Mandelstam & Waites (1968) described a B. subtilis mutant E22 which failed to secrete extracellular protease. This

strain did not show the rapid intracellular protein turnover that proceeds sporulation in wild-type cells, and it did not sporulate. They concluded that these phenomena were attributable to loss of the serine exoprotease. Leighton et al (1973) concluded similarly that the temperature sensitive serine protease secreted by mutant ts-5 was responsible for the blockage showed by the strain at stage 0 in sporulation when grown at the nonpermissive temperature. Another mutant, ts19, was studied by Millet, and had a phenotype similar to ts-5. At the permissive temperature, sporulation could proceed, but the serine protease was more thermolabile than wt. Reversion and transformation experiments succeeded in separating 2 independent ts mutations: Spo^+ transformants, which nonetheless retain the impaired exoprotease, were obtained (Millet et al, 1976). Indeed, mapping of the cloned subtilisin gene has subsequently shown it to lie at a position quite distinct from the ts-5 allele. Doi has suggested that this mutation need not necessarily also comprise 2 independent lesions, but may encode a product that renders subtilisin more thermolabile (Wong et al, 1984). Both Mandelstam's and Leighton's groups were persuaded that the protease acted on the β-subunit of RNA polymerase in modifying the vegetative form to the "sporulation-specific" : such a modification had been reported (Losick et al, 1970). It had also been achieved in vitro using purified intracellular serine protease, but not with serine exoprotease (Millet et al, 1972). Modification of the β-subunit was however later shown to be an artefact of isolation (Linn et al, 1973).

The controversy was finally resolved by the use of the cloned subtilisin structural gene to generate a mutant in which the wild-type allele has been replaced with a gene having an in vitro-derived deletion of 684 bp in the coding sequence. This strain, BG2019, sporulates as well as the wild-type does (Stahl & Ferrari, 1984). Kawamura & Doi (1984) have obtained similar results with a smaller deletion. It can thus be con-

cluded that subtilisin has but a "scavenger" function, i.e. generates usable nutrients from macromolecules in the environment when more accessible nutrients have been depleted.

I.2.2.5 Neutral metalloprotease/NPR: Characterization and Diversity

Metal-chelator (eg EDTA, O-phenanthroline) -sensitive proteinases active at neutral pH are widely distributed in micro-organisms, including the bacilli (Morihara, 1974). They are not inhibited by -SH agents.

The metalloproteases are not as widely used industrially as are the serine proteases. The enzyme obtained from B.amyloliquefaciens is the most widely used, and commercial preparations enriched for the neutral protease are obtained by adjustment of fermentation medium and conditions, and also by mutation of the host to reduce serine protease formation.

Some Bacillus species secrete only the metalloprotease eg.B.megaterium (Millet et al, 1969), B. polymyxa, B.thermoproteolyticus (Keay & Wildi, 1970) and B. thuringiensis (Li & Yousten, 1976).

First crystallized from B. amyloliquefaciens in 1951 (Fukomoto & Negoro, 1951), the enzyme was further characterized by Tsuru and coworkers (McConn et al, 1964).

The neutral enzyme has also been isolated from B. subtilis var. amylosaccharíticus (Tsuru et al, 1966a & b), B. megaterium ("megapeptidase", Millet & Aubert, 1969) B. thuringiensis (Li & Yousten, 1976) and B. thermoproteolyticus ("thermolysin", Endo, 1962; Matsubara, 1970).

The enzymes from these various sources have many common properties. The thermolysin is, however, distinguished by its exceptional heat-stability (retains 90% activity after 30 mins. at 70°C, and 50% after incubation at 80°C for an hour).

The other neutral proteases exhibit poor stability with respect to heat, purification and freeze-thawing.

1 g-atom Zn^{2+} is present per 38-42 kd molecular weight, and is essential for activity, while Ca^{2+} is required for structural stability and prevents autolysis.

The enzymes display limited activity on small peptide substrates, and have no amidase or esterolytic activity, unlike the subtilisins. Hydrophobic or bulky amino acid residues are preferred as the amino acid N-terminal to the peptide bond to be cleaved (Matsubara & Feder, 1971; Morihara, 1974).

The "megapeptidase" is produced throughout exponential growth in minimal medium; catabolite repression in complex medium leads to secretion only during sporulation (Millet et al,1969).

On the basis of immunological studies, amino acid composition, comparative stability, and activity on large substrates, it was concluded that the commercially available NPRs of B. subtilis NRRLB3411 (probably a B. amyloliquefaciens strain) and B. subtilis var. amylosaccharíticus were indistinguishable, but were different from thermolysin. Megapeptidase was immunologically related to thermolysin (Keay, 1969; Keay & Wildi, 1970).

I.2.2.6 npr Gene

Two ts mutants were isolated where the substrate specificity and pH-optima of the altered enzymes were different from wild-type, suggesting a mutation in the structural gene nprE. This was mapped to a position between pyrA and metC (Uehara et al, 1979).

There is a tightly linked regulatory locus nprR1. The nprR2 allele, isolated first in the Japanese strain B. natto, causes specific overproduction of neutral protease (Uehara et

al,1974). NPR production is also regulated by loci which also affect other enzymes eg hpr (Prestidge et al, 1971;.Higerd et al,1972), sco and catA (Dod & Balassa, 1978), pap-9, sacU, amyB (Steinmetz et al, 1976; Yoneda & Maruo, 1975). Recently the neutral protease genes of B. subtilis (Yang et al, 1984), B. amyloliquefaciens (Vasantha et al, 1984), and the gene encoding the thermostable NPR of B.stearothermophilus (Fujii et al, 1983) have been cloned. Sequencing and in vitro transcription-translation (B. subtilis and amyloliquefaciens genes) show that, as is the case with alkaline protease, NPR is synthesized in a large pre-pro-form. The primary translation product includes a typical N-terminal "signal peptide" of 27-28 AAs, followed by a pro-region of approximately 194 AAS, which is also processed and does not form part of the extracellular enzyme of 300 AA residues. Both the precursor and the mature enzyme migrate anomalously on SDS gels: (viz. precursor deduced M_r: 55,000, apparent M_r: 67,000, mature deduced M_r: 33,000, apparent M_r: 45,000) (Yang et al, 1984). Perfect identity of the deduced AA sequence from B.amyloliquefaciens (Vasantha et al, 1984) and the 39 N-terminal AAs sequenced from the NPR of B. subtilis NRRLB3411 (Levy et al, 1975) makes it likely that the latter strain has been misclassified, as has been mentioned (I.2.2.1). This would account also for the fact that the B. subtilis AA sequence deduced from the gene matches the same published AA sequence rather less well (Yang et al, 1984). The sequence of the npr genes of B. subtilis and B. amyloliquefaciens show a high degree of homology in the coding area but not in the immediately adjacent upstream regions.

I.2.2.7 Relationship of NPR to Sporulation

Early-blocked pleiotropic sporulation mutants fail to secrete NPR (Piggot & Coote, 1976). However, mutants deficient in NPR, but able to sporulate normally were obtained for B. subtilis (Hageman and Carlton, 1973; Michel & Millet, 1970; Mandelstam and Waites, 1968), B. cereus (Aronson et al, 1971) and B. megaterium (Millet & Aubert, 1969; Millet et al, 1969).

Ferrari and coworkers (Yang et al, 1984) have used the cloned genes to create a double chromosomal deletion mutant BG2054 lacking both APR and NPR. There was no detectable effect on growth, morphology or sporulation, thus reconfirming unambiguously that NPR is of no relevance to the sporulation process and calling into question the conclusions of other workers (Jolliffe et al, 1980) on the relationship of extracellular protease and cell morphology (Yang et al, 1984). With a combination of the uncharacterized lesion nprE18 and an apr deletion mutant Doi could similarly not detect any compromise of the cell's ability to sporulate (Kawamura & Doi, 1984).

The physiological role of the extracellular neutral and alkaline proteases is probably to function only as scavengers.

I.2.2.8 Esterase/Bacillopeptidase F:
Characterization and Diversity

The enzyme was first identified by Boyer & Carlton (1968), who called it "acidic protease". It was also characterized by Millet (Millet, 1970; Mamas & Millet, 1975) as "enzyme II", Spizizen (Prestidge et al, 1971) as "enzyme C", and Hageman and Carlton (1973) as "bacillopeptidase F".

Several authors mention that the enzyme is seen in at least 2 forms, (one apparently an aggregate) by electrophoretic or column separation. The enzyme is inhibited by PMSF, and exhibits low activity on large protein substrates relative to the alkaline serine protease, accounting for c.20% of serine proteolytic activity (Boyer & Carlton, 1968). It hydrolyses some synthetic esters not hydrolysed by either the neutral metalloprotease or subtilisin, eg.benzoylarginine-, and benzoyltyrosine ethyl esters.

Data on the heat stability and pH optima relative to subtilisin of this enzyme conflict somewhat (cf. Boyer and Carl-

ton, Prestidge et al).

The evidence for an intracellular location of active bacillopeptidase F should be viewed with caution. While Prestidge & coworkers (1971) were able to measure extracellular protease using synthetic substrates shown to be almost specific for the different enzymes viz. TAME, FAGLA and BAEE for subtilisin, the neutral metalloprotease and the esterase resp., intracellular proteolysis could not be similarly ascertained because the assays rely on absorption measurements in the UV region. Thus p-nitro-phenylacetate, which is not a substrate for NPR or APR, was used to estimate the esterase intracellularly. But other esterases could of course have contributed to the activity measured in this study.

In an immunological study Hageman and Carlton (1973) observed crossreaction of anti-bacillopeptidase F with cell extracts of wild-type B. subtilis 168 and with a mutant L4, which did not produce detectable extracellular NPR and secretes a low level of APR. This was not achieved with extracts of sporulation mutant S-87, which produces only a very low level of NPR extracellularly. In zymograms of the wild-type extract, a band of activity corresponding in position to mature neutral protease was seen (but with hindsight this is surely due to contamination: neutral protease is synthesized as a much larger prepro- zymogen). No zymogram band corresponding to authentic bacillopeptidase F was present in either wt or the L4 strain, which was puzzling at the time, since cross-hybridization had been achieved. A band of active protease not co-electrophoresing with either subtilisin, neutral protease or bacillopeptidase F markers was common to the wild-type and L4, but missing from the low protease secreting sporulation mutant S-87. This band was interpreted as being either a modified bacillopeptidase F (though with hindsight such a precursor is unlikely to display much protease activity) or else a hitherto uncharacterized intracellular protease, which is required for sporulation: the latter inter-

pretation has been favoured in the literature (Hageman and Carlton, 1973). It was, however, not ascertained directly whether or not the band of activity common to 168 and L-4 also contained the immunoreactive material. Subsequent accounts make it likely that ISP-I was the enzyme involved in this study (see I.2.3.2).

Recently, bacillopeptidase F has been isolated under improved conditions (Roitsch & Hageman, 1983). It was shown that some previously used procedures could lead to proteolytic breakdown, and that upon storage at high pH, extra enzymatically active fast-moving electrophoretic forms appeared. These possibly account for the multiplicity of electrophoretic forms seen by others (Mamas & Millet, 1975).

When isolation and purification take place at pH6 in the presence of PMSF, which inhibits autodigestion, 2 high molecular weight, immunologically indistinguishable forms are found. Sedimentation gives M_r values of 33,000 and 50,000 respectively for these 2 fractions. Both have bound carbohydrate, and phosphoserine residues. Both amino acid composition and size estimation of the smaller component agree well with earlier data for the esterase (Boyer and Carlton, 1968; Mäntsälä & Zalkin, 1980).However, the larger component has an amino acid composition that is strikingly different from the smaller, and totally unlike that of the subtilisin-like intracellular serine protease ISP-I (see I.2.3.2). It is only just possible from the AA composition that the smaller form is derived from the larger by the loss of a basic segment, such as the pro-regions of NPR and APR, for the ratio of acidic to basic residues in the case of the 33,000 species is 5.5, and just 2.8 for the larger fraction.

Bacillo-peptidase F and ISP-II (Srivastava & Aronson, 1981; see I.2.3.4) are not immunologically related (Roitsch & Hageman, 1983).

I.2.2.9 Relationship of Esterase to Sporulation

The esterase is the least studied of the proteases discussed here, and no direct studies have been carried out to investigate its relationship to the sporulation process. Its production is however impaired in early-blocked sporulation mutants, along with subtilisin (Michel and Millet, 1970; Hageman and Carlton, 1973 :the S-87 mutant mentioned above).

I.2.2.10 Miscellaneous

B. licheniformis secretes a non-metallo-, non-serine type protease, inhibited by neither EDTA nor PMSF. This "component C" is produced to approximately 10% of the amount of subtilisin ("Alcalase"), and has only been characterized further in the industrial context, having been an undesirable, highly allergenic contaminant of commercial enzyme preparations (Verbruggen, 1975).

I.2.3 Intracellular Proteases and Esterases

I.2.3.1 Characterization and Diversity

An enormous effort has been expended in trying to ascertain a role for serine exoproteases, or their intracellular counterparts, in sporulation.

A number of early studies on extracellular protease production deal also with protease activity found intracellularly (eg Hageman & Carlton, 1973; Prestidge et al, 1971; Mandelstam & Waites, 1968). It is important to recognize that it was only in 1975 that the signal hypothesis became widely held (Blobel & Dobberstein, 1975). In early work it was assumed that observed protease or esterase activity were due to the intracellularly located counterparts of the already characterized exoenzymes, as differences demonstrated intracellularly were reflected in the products found in the medium. Subsequently we have learnt that the secreted forms of neutral and serine proteases are synthesized as longer, probably inactive or poorly active precursors, and that there

is an ever-growing repertoire of characterized bona fide intracellular proteases, distinguishable from the secreted species by several criteria.

Two acetyl esterase serine protease activities were described in B. subtilis lysed spheroplasts, using beta-naphtholacetate as the substrate (Higerd & Spizizen, 1973). "Esterase A" could be isolated from both vegetative and sporulating cells, and had an estimated M_r of 31,000 or 160,000 as determined by gel filtration or SDS-gel electrophoresis respectively. "Esterase B" had an M_r of 51,000 by gel filtration and was specific to sporulating cells. Three out of four pleiotropic, asporogenous mutants had reduced levels of esterase B, but the 2 mutants previously mentioned viz L-4 and S-87 (see I.2.2.8), had wild-type levels of both esterases A and B. These esterases have low activity with azocasein as the substrate.

At least 4 different intracellular protease fractions appearing after T_0 (where T_0 is defined as the end of exponential growth) were distinguished in B. subtilis ATCC6051 (Hiroishi and Kadota, 1976), on the basis of substrate preference and inhibition profiles. "Fraction C" was clearly equivalent to the "ISP-II" described by others (Srivastava and Aronson, 1981; see I.2.3.4), and to "esterase B" (Higerd & Spizizen, 1973; see above). "Fraction B" was evidently the same as the subtilisin-like intracellular serine protease ISP-I (Reysset & Millet, 1972) which is discussed in greater detail in I.2.3.2. In addition there was an EDTA-inhibited esterase (fraction A), and the remaining fraction (fractions D+E) had both esterolytic and proteolytic activity that were inhibited by DFP. These fractions cannot be correlated with other reports.

Hageman and Carlton (1973) described only one intracellular proteolytic activity, as mentioned in I.2.2.8, which was common to wild-type and the NPR-negative L-4.

I.2.3.2 ISP-I

This intracellular serine protease, found only in sporulating cells, and first described by Millet, differs in biochemical properties from the secreted subtilisin: it is inactivated by PMSF, but also by EDTA, since there is an absolute requirement for Ca^{2+}, and the enzyme has a much lower pI than does subtilisin (viz. pH4). Azocoll is a good substrate for the enzyme, which shows a narrower specificity than subtilisin towards synthetic esters (Reysett & Millet, 1972) and also towards large protein substrates. For example, the ISPs of B.amyloliquefaciens (Markaryan et al, 1981) and B. brevis (Kurotsu et al, 1982) cleave only some of the bonds that are susceptible to the exoenzyme secreted by the former strain.

Enzymes with similar properties have been isolated from sporulating B. megaterium (Millet, 1971), B. thuringiensis (Lecadet et al, 1977), B. cereus (Cheng & Aronson, 1977) and B. brevis (Kurotsu et al, 1982).

Various roles have been proposed for ISP-I, based on its in vitro activities and the study of mutants deficient in it. They include sporulation-specific protein turnover, and processing (Cheng & Aronson, 1977), modification of RNA polymerase (Millet et al, 1972; Lecadet et al, 1977) and possibly antibiotic degradation which confers protection on the producer strain. Such is probably the case with B. brevis, whose ISP can degrade gramicidin S, whereas the antibiotic inhibits the ISPs of other species in vitro and is often used in affinity chromatography to purify ISP (Kurotsu et al, 1982; Strongin et al, 1979a). In the case of B. licheniformis, its own serine exoprotease is inhibited by the antibiotic bacitracin it produces, but its ISP is not (Vitkovic & Sadoff, 1977).

I.2.3.3 Relationship of ISP to Sporulation

I.2.3.3.1 Indirect Evidence

There was some evidence suggesting that intracellular serine protease activity may be indispensible for sporulation.

This came first from inhibitor studies where PMSF, administered between T_0 and T_3, inhibited both extracellular subtilisin and sporulation (Dancer & Mandelstam, 1975a). It was known that mutants lacking extracellular protease failed to show the high intracellular protein turnover typically seen (Monro, 1961) under conditions that induce sporulation in the wild-type (Mandelstam & Waites, 1968). The ts-5 subtilisin mutation was believed to be responsible for the sporulation defect in that strain (Leighton et al, 1973). As it was difficult to imagine an extracellular role in sporulation for the protease it was concluded that the inhibitors in above-mentioned study were acting on an intracellular form of the enzyme.

However, apart from the possible presence of other PMSF-inhibitable proteases there is uncertainty about how specific the effects of PMSF are in such in vivo inhibition experiments. Serine protease inhibitors were shown to have a "non-specific" effect on energy generation in intracellular proteolysis studies in E. coli (Schechter et al, 1973). These workers correlated an apparently lower rate of protein break-down than that normally seen under carbon source deprivation with dramatically lowered ATP levels in these cells.

Work with sporulation-deficient ISP-I ts mutants was rather more convincing than the inhibitor studies.

I.2.3.3.2 ISP-I ts Mutants and Spore Protease

The ISP-I of strain B. subtilis SMY-512 has an apparent M_r of c. 64,000, or 32,000 under native or denaturing conditions respectively. At the non-permissive temperature, mutant ts-15 failed to sporulate and displayed negligible protein turnover, but normal exoprotease production. A cryptic protein of M_r 32,000, that cross-hybridized with anti-ISP-I was extracted under both native and denaturing conditions. The implication was that this mutant ISP could not form the active

dimer complex. Studies of revertants of ts15 which had regained all wild-type characters, and transformation experiments with mutant DNA, pointed to a single gene being involved. This contrasted with previous studies where apparently similar mutants had failed to give revertants (Kerjan et al, 1979).

A study of B. cereus germination mutants, one of which, upon screening showed reduced post-exponential ISP activity, was less satisfactory (Cheng & Aronson, 1977). The purified enzyme of the G8 mutant was more heat-labile than wild-type, and the mutant was temperature-sensitive for spore-production (though only in the absence of available amino-acids), producing spores deficient in coat protein. Nonetheless, wild-type levels of high M_r components that cross-hybridized with anti-coat protein were present. G8 however lacked the 12,000 component (of a size with major coat protein found in mature spores) seen in the wild-type strain. Revertants, some of which retained some ts characters, were obtained at a frequency that tended to imply a point mutation in one gene (Cheng & Aronson, 1977). It is difficult to accept that the complex phenotypes displayed by G8 could be due to a single protease gene mutation. The abnormal coat protein processing might well be due to the ISP-I lesion, but this was not conclusively shown as revertants (chosen on the basis of spore-lysozyme resistance) did not conform to the wild-type processing pattern in pulse-chase experiments.

Moreover, the protease that is active during germination has been characterized, and is quite different from ISP-I: Spore protease is synthesized in the form of an inactive precursor (P_{46}) at the end of exponential growth, and is processed several hours later to another inactive form P_{41}. It is activated by processing to P_{40} only in the first few minutes of germination, and is absolutely specific for the small basic proteins that comprise more than 85% of the spore protein, whose degradation it initiates (Loshan et al, 1982). Mutants having redu-

ced levels of spore protease exhibit a lower rate of proteolysis during germination, but no other phenotypic defects - they display normal outgrowth (Postemsky et al, 1978). So even if one assumes that the mutant G8 ISP is responsible for aberrant processing of the spore protease precursors, it could not account for the abnormal outgrowth of the germination mutant G8.

I.2.3.3.3 ISP-I Gene; Effects of Gene Deletion & Overproduction of ISP-I

There was evidence from N-terminal AA sequencing of ISP from various sources: B. amyloliquefaciens (Stepanov et al, 1977; Strongin et al, 1978a); B. subtilis (Strongin et al, 1979a & b); B. licheniformis (Strongin et al, 1979b), that suggested that these proteins were structurally related to subtilisin. Comparison of the ISP and exoprotease of B. subtilis strain A-50 suggested that there were, however, 2 separate genes coding for intra- and extracellular serine proteases, as the respective partial sequences were similar, but not identical (Strongin et al, 1978b). (Note that in the above studies, subtilisin Carlsberg was erroneously attributed to B. subtilis and strain A-50 turned out to be a strain of B. amyloliquefaciens (Markaryan et al, 1981)).

This has recently been confirmed by the cloning of the ISP-I gene by Beppu and co-workers. In B. subtilis, ISP was overproduced when the cloned gene was present on a high copy number plasmid. The AA sequence deduced from the gene is shown in Fig. 2 (Koide et al, 1986; and personal communication), where it is compared with that of subtilisin Carlsberg. There are several blocks of identical sequence, with limited homology at the extreme ends of the protein. No "signal sequence" is found. Only typical -10 regions for sigma-43 or -29 (I.3.4 & I.3.7) are situated upstream from the transcriptional start site.

It is evident that the N-terminal sequences obtained by Strongin and coworkers derive either from a processed or a degraded form of the polypeptide that is encoded by the ISP-I gene (that

```
ISP     +10         +20         +30         +40         +50         +60         +70         +80
MNGEIRLIPY VTNEQIMDVN ELPEGIKVIK APEMWAKGVK GKNIKVAVLD TGCDTSHPDL KNQIIGGKNF SDDDGGKEDA
PDVAYVEEDH VAHALAQTVP YGIPL  AD  ---VQ Q F   A V         IQA      - -VV  --- -ASFVAG AY
CARLSBERG           +5          +15         +22         +32         +42         +47         +56

        +90         +100        +110        +120        +130        +140        +150        +160
ISDYNGHGTH VAGTIAANDS NGGIAGVAPE ASLLIVKVLG GENGSGQYEW IINGINYAVE QKVDIISMSL GGPSDVPELE
NT G            V  L N TT VL    S V  YA    - NSS   T SG  VS  EW TT NGM V N        GSTAMK
        +66         +76         +86         +95         +105        +115        +125        +135

        +170        +180        +190        +200        +210        +220        +230        +240
EAVKNAVKNG VLVVCAAGNE GDGDERTEEL SYPAAYNEVI AVGSVSVARE LSEFSNANKE IDLVAPGENI LSTLPNKKYG
Q  D  YAR   V  A    S -GSSGN NTI G   K DS      A DSNSN RAS  SVGA  LEVM   AGV Y  Y TST A
        +145        +155        +164        +174        +184        +194        +204        +214

        +250        +260        +270        +280        +290        +300        +310        +320
KLTGTSMAAP HVSGALALIK SYEEESFQRK LSESEVFAQL IRRTLPLDIA KTLAGNGFLY LTAPDELAEK AEQSHLLTL
T N     S    A  A   L  KHPNLSASQ VRNRLSSTAT YLGSSFYYGK GLINVEAAAQ
        +224        +234        +244        +254        +264        +274
```

Comparison of the deduced amino-acid sequence of the Intracellular Serine Protease (upper sequence, Beppu) and that of subtilisin Carlsberg (lower sequence, this thesis).

Only differences from ISP are indicated for subtilisin Carlsberg. A number of gaps were introduced to improve the alignment. These are indicated by (-).
The deduced Carlsberg sequence is shown from AA-15, and the first AA of the secreted enzyme is underlined.

FIGURE 2: COMPARISON of ISP-I with SUBTILISIN CARLSBERG

is, nevertheless, active). Isolated ISP-I's N-terminal AA lies at +20, or in the case of a minor species, at +17 (cf. Strongin et al, 1978a with the deduced AA sequence deduced from the gene in Figure 2). The significance of the N-terminal heterogeneity and diversity is not known. Strain differences, *in vitro* artefacts, or differential processing may all contribute.

The effect of expression of the cloned gene in a high copy number plasmid has been tested in the *S-87* mutant described earlier, which had only 20% of wt ISP activity. *S-87* (pYK3) overproduced both ISP, and surprisingly, exoprotease as well, but the mutant sporulation phenotype remained unchanged. In the wild-type strain however, the effect of carrying pYK3 was that sporulation was inhibited. The effect of supplying the gene in a single copy was thus ascertained in the mutant. Though normal levels of ISP-I were produced, the strain still did not sporulate. On the other hand, interruption of ISP-I coding sequences (by deletion or insertion) led to significant impairment of sporulation by the wild-type. This was observed in synthetic, but not rich medium, and suggests a role for ISP-I in normal sporulation, but this role is non-essential.

Interestingly, just upstream of the ISP-I ORF, a second open reading frame running in the opposite direction, and potentially encoding a product of M_r 22,000, is found. Removal of this ORF from the clone leads to lysis of the *E. coli* host unless the isolated ISP-I gene is on a low copy number plasmid, or the ISP-I gene orientation is reversed. This suggests a role as a repressor or a specific inhibitor of ISP-I for the 22 kDa protein.

If this proves to be the case, it would fit very well with a size of 15-20,000 estimated for the specific protein inhibitor of ISP-I isolated from both growing and sporulating *B. subtilis* cells by Millet (Millet, 1977), and which was also present in the *ts-15* ISP-I mutant described by Kerjan (see I.2.3.3.2) at both permissive and non-permissive temperatures. The inhibitor

may have a regulatory function in modulating enzyme activity. Comparison of ISP-I antigen, and activity levels, indicate a post-translational activation step, and since S-87 has wild-type antigen levels, it may be defective in production of the active form, possibly due to higher levels of inhibitor, or to altered (tighter) binding to same (Burnett et al, 1986).

I.2.3.4 ISP-II

ISP-II can be distinguished unequivocably from ISP-I by its pH optimum (pH7), stability, size (47,000), substrate specificity and inhibition profile (Srivastava & Aronson, 1981; Hiroishi & Kadota, 1976). While both are produced post-exponentially, they peak at different times (ISP-II earlier), which presumably reflects different physiological roles. ISP-II comprises only about 10-20% of the Azocoll-hydrolysing activity of extracts of sporulating cells; it is relatively more active against arginyl esters, and is not inhibited by EDTA. A B. subtilis ts mutant of ISP-II has been isolated (Srivastava & Aronson, 1981), which has all the phenotypic alterations previously described for the B. cereus G8 mutant, and there attributed to the alteration in ISP-I (Cheng & Aronson, 1977).

I.2.3.5 Membrane-bound Proteases

Membrane-bound forms of a variety of secreted enzymes of the bacilli have been described, and may act as intermediates in secretion, or as a pool of enzyme for regulated release eg. α-amylase (Nagata et al, 1974), alkaline phosphatase (Yamane & Maruo, 1978), lipase (Kennedy & Lennarz, 1979) and β-lactamase (Yamamoto & Lampen, 1976). Membrane vesicles from late log cells of B. subtilis YY88 (pap-9), which overproduces proteases 20-fold, yielded a variety of proteases upon extraction with Triton-X 100 (Mäntsälä & Zalkin, 1980). More than 90% of these comprised APR and NPR indistinguishable from the secreted enzymes by mobility on SDS gels, immunoprecipitation and inhibition profiles. Both forms first appear at the same time

as their extracellular counterparts, although levels of the membrane-bound proteases peaked several hours earlier. Other recent studies show that the preprosubtilisin form is membrane-associated prior to the release of the extracellular form (Power et al, 1986).

The remaining activity could be separated into at least 4 previously uncharacterized "M proteases", which were all weakly inhibited by PMSF and had high esterase activity relative to APR and NPR.

As the total membrane-bound proteolytic activity amounted to only 1-2% of that secreted by YY88 it is likely that the enzymes involved have previously escaped detection in wild-type strains due to their low levels.

I.3: Transcriptional Control of Gene Expression in the Bacilli

I.3.1 Sigma Factors

It now seems that prokaryotes may have the capacity to control gene expression at the level of higher order chromosomal structure (Lilley, 1986), as is the case for eukaryotes. However, in the genus Bacillus especially, it has been demonstrated that to a large degree, differential gene expression is exercised at the level of transcription by modification of the promoter specificity of RNA polymerase, among other more general mechanisms.

Early investigations showed that sporulation-associated morphogenesis was accompanied by biochemical changes, and these could be correlated with the appearance of new mRNA species (Doi & Igarashi, 1964; DiCioccio & Strauss, 1973), and with a distinct, temporally defined pattern of protein synthesis (Linn & Losick, 1976).

Such a program of changes is analogous to the events that occur in bacteriophage infection: after infection of E. coli by T4 'phage, different classes of viral RNA are synthesized in a de-

fined temporal sequence. All except "pre-early" RNAs depend on 'phage-specific protein synthesis; the "switching-on" of later stages thus required at least one virus-encoded component. The subunit composition of RNA polymerase was found to change in the course of 'phage development, and a new 'phage-encoded "sigma" subunit was shown to be responsible for the altered template specificity of the enzyme i.e. initiation of transcription from viral rather than host "promoters" (viz. specific sequences that are recognized by RNA polymerase just upstream from the transcriptional initiation site) (Travers 1969, 1970; Travers & Burgess, 1969).

Coincident with these developments in the understanding of T4 biology it was found that the template specificity of RNA polymerase becomes altered during spore formation in B. subtilis (Losick & Sonenshein, 1969). Rifampicin-resistant beta subunit mutants were obtained, where the lesion led to loss of sporulation ability (Sonenshein & Losick, 1970), and differences in the subunit structure of RNA polymerase from vegetative and sporulating cells were demonstrated (Losick et al, 1970). These observations supported the postulate that, analogous to the T4 model, the developmental program leading to the formation of the spore is also controlled by polymerase modification.

Though it later became clear that the beta subunit alterations were artefacts (Linn et al, 1973), studies on B. subtilis 'phage SPO1 rapidly produced evidence of 'phage-specified sigma factors for which distinct promoter recognition sequences could be identified. Thus host holoenzyme transcribes early genes, and SPO1 gene 28, 33 & 34-encoded sigma-like factors direct transcription in the host from specifically middle and late genes respectively (Pero et al, 1975; Tjian & Pero, 1976; Duffy & Geiduschek, 1977; Talkington & Pero, 1978). Distinct consensus sequences have been defined for the three classes of SPO1 genes (Talkington & Pero, 1979; Lee et al, 1980; Lee & Pero, 1981) (See Table 1).

Subsequently, additional sigma factors were isolated from uninfected cells:

In 1979 a novel sigma, σ^{37}, was isolated from the RNA polymerase of early sporulating cells, and conferred on the core enzyme ("E"), as σ^{43} (formerly called sigma-55) failed to do, the ability to transcribe 2 cryptic cloned genes (viz. 0.4kb/spo VG, ctc/spoVC) whose expression is associated with sporulation (Haldenwang & Losick, 1979, 1980). Another sigma factor, σ^{28}, was found to occur in non-infected, vegetatively growing cells soon after, and also conferred a distinct template specificity on core enzyme (Wiggs et al, 1981), thus establishing RNA polymerase modification as a general means of altering template specificity and thereby controlling transcription.

In all, 7 sigma-like factors have to date been characterized, and are designated by a number corresponding to molecular weight, or in the case of 'phage factors, by the number of the specifying gene. These are tabulated, together with their promoter "consensus" sequences, in Table 1.

Also tabulated (Table 2) are other RNA polymerase-associated polypeptides, whose role is as yet unclear, as well as a scheme indicating at which stages of growth the different holoenzymes can be isolated from uninfected cells and the relative abundance of some. Of the above-mentioned core-associated polypeptides, delta has specifically been described as a transcription inhibitor and appears to make $E\sigma^{43}$ more selective by destabilizing enzyme-DNA complexes (Doi, 1982). It exerts little effect on the activity of strong promoters. It should be noted that the presence of a particular holoenzyme at a certain stage of growth does not always imply activity at that stage.

Just how sigma factors exert their effects is not known; cross-linking studies with E. coli sigma^{70} and the lac UV5 promoter have shown direct interaction with both of the conserved re-

TABLE 1: CONSENSUS SEQUENCES FOR BACILLUS SUBTILIS PROMOTERS

(non-transcribed strand shown)

Sigma factor	-35 region	-10 region	no. sequences compared	reference
$\sigma^{55} = \sigma^{43}$	TTGACA 989865	TATAAT 698969	9	Moran et al (1982)
σ^{gp28}	T-AGGAGA--A	TTT-TTT	6	Lee & Pero (1981)
$\sigma^{gp33-34}$	CGTTAGA	GATATT	5	Johnson et al (1983)
σ^{28}	CTAAA 22222	CCGATAT 2222222	2	Gilman et al (1981)
σ^{29}	TTAATAAAT 444647444	CATATT-T 566744 6	7*	Moran (1985) personal communication
σ^{37}	AGG-TTT-A 442 342 3	GG-ATT-TTT 33 444 333	4*	modified from Wong &Doi(1984)
σ^{32}	AAATC 22222	TA-TG-TT-TA 22 22 22 22	2*	Johnson et al (1983)

* These sequences include ctc and spoVG

Numbers below consensus nts refer to the number of cases where the nt is found at that position.

TABLE 2: COMPOSITION OF RNA POLYMERASE AT DIFFERENT STAGES OF GROWTH IN BACILLUS SUBTILIS

Modified from Doi (1982)

CORE

$E = \alpha_2 \beta \beta'$

Core-associated subunits

veg	spo
δ	p^{23}
ω^1	p^{34}
ω^2	p^{85}

Modified from Doi (1982)

Forms of RNA polymerase isolated from B. subtilis cells at different stages of growth and sporulation

E	$E\sigma^{55}$	$E\sigma^{37}$	$E\sigma^{28}$	$E\sigma^{32}$	$E\sigma^{29}$	$E\delta$	EP^{34}	EP^{23}	
+	+	+	+	+	-	+	-	-	vegetative cells
+	+	+	+	+	-	+	-	-	stationary spoOA
+	+	+	-	+	-	+	-	-	stage 0
+	+	-	-	-	+	-	+	-	stage II
+	+	-	-	-	+	-	+	+	stage III
	5000	500		30					relative abundance*

* Johnson et al(1983)

NOTE: The presence of a particular form does not necessarily imply activity at that period of growth (see text concerning sigma-28).

gions that constitute many promoter sites (Simpson, 1979; Chenchick et al, 1981).

Although the number of cases of transcription by low-abundance sigma-containing RNA polymerase that have been studied to date is still limited, some generalisations can be made. The properties of genes transcribed by RNA polymerase associated with the various sigma factors is given in brief below.

I.3.2 Summary of Functional Properties of Sigma Factors

Major component:

σ^{43}: component of major form of holoenzyme; transcribes genes during vegetative growth.

Minor factors found in vegetative cells:

σ^{37}: involved in post-exponential expression of genes under complex regulation.

σ^{32}: transcribes sporulation genes in vitro, role if any in vegetative growth uncertain, as is case for σ^{37}.

σ^{28}: transcribes a limited number of genes during exponential phase whose gene products seem to be involved in the transition to sporulation phase. Most spoO functions are required for σ^{28} activity.

Abundant sporulation-specific factor:

σ^{29}: persists only for an interval from T_2 to T_4.

More detailed data follows below.

I.3.3 Methodology Used in Transcription Studies

As the methodology is crucial for an evaluation of the data, the most frequently exploited techniques in the studies discussed below are listed here. They include the isolation of ternary initiation complexes of polymerase, DNA template and nascent RNA chain on nitrocellulose filters; in vitro RNA synthesis on cloned native and mutagenized templates with isolated holoenzyme (i.e. core & sigma factor), or enzyme reconstituted from purified components; identification of transcribed regions by

hybridization of labelled transcripts tc immobilized DNA fragments by blotting; identification of transcriptional start sites by 1) using templates truncated at a site within the transcribed region and monitoring the size of the resultant "run-off" RNA, 2) the S1 mapping technique (Berk & Sharp, 1977), 3) dinucleotide transcriptional priming (Minkley & Pribnow, 1973) and 4) RNA sequencing.

It is pertinent to add that the composition of RNA polymerase isolated is much dependent on the method of extraction and purification (Doi, 1982).

I.3.4 Sigma-55/43 Promoters

In the case of promoters recognized by the principal form of B. subtilis holoenzyme ($E\sigma^{55}$, or rather, $E\sigma^{43}$), which transcribes genes active in vegetatively growing cells (Shorenstein & Losick, 1973a & b), and the early genes of phages SP01 and Ø29, the -35 and -10 regions conform closely to the consensus hexanucleotide recognition sequences attributable to E. coli promoters, of which a large number have been characterized. (Reviews E. coli promoters: Rosenberg & Court, 1979; Siebenlist et al, 1980; Reviews B. subtilis promoters: Lee et al, 1980; Moran et al, 1982a; Losick & Pero, 1981).

In E. coli, as well as Gram-positive organisms, initiation takes place preferentially at a purine in position 7 or 8 downstream from the last T in the -10 sequence (Murray & Rabinowitz,1982; Aoyama & Takanami, 1985). There are some indications that certain positions in the consensus are more highly conserved and important for efficient utilization in B. subtilis than is the case in E. coli eg. the "invariant G" in the -35 region. However, at least one strong B. subtilis promoter has been characterized which does not have the invariant G in the putative -35 region (Nakahama et al, 1985).

Promoters that conform to the consensus coming either from the

B. subtilis chromosome eg. veg, tms (Moran et al, 1982 a), from B. subtilis phages eg. SPO1 and Ø29 (Murray & Rabinowitz, 1982) or from E. coli eg lacUV5, tac (Lee et al, 1980, Moran et al, 1982a; Peschke et al, 1985) which are all utilized efficiently by E. coli polymerase, are transcribed to greatly varying extents by $E\sigma^{43}$ from B. subtilis (see also I.7.3).

Thus other features besides the canonical consensus sequence are important in B. subtilis $E\sigma^{43}$ transcription. Indeed, most genes which are transcribed by the major holoenzyme show additional common features.

Apart from the -35 and -10 regions, σ^{43} promoters often, but not invariably, show the pentanucleotide PuPyPuGT at about -15, and have predominantly A,T between -1 and -6 from the transcriptional start (Wang & Doi, 1984; Moran et al, 1982a). The sequence about -15 has been implicated in methylation studies of RNA polymerase-promoter binding complexes to give a strong protective effect (LeGrice & Sonenshein, 1982).

A special category, the early 'phage genes recognized by $E\sigma^{43}$, have additional homology immediately upstream of the -10 and -35 regions (Murray & Rabinowitz, 1982).

"Spacer" length between the -10 and -35 regions is not so critical and may vary from 16 to 18 bp (Peschke et al, 1985).

The functional boundaries of B. subtilis promoters definitely extend upstream from the polymerase recognition/binding sites: an AT-rich box consisting of alternating stretches of As and Ts occurs commonly around -40 to -50 for σ^{43} promoters, especially in the case of highly expressed genes (see Wang & Doi, 1984; Banner et al, 1983). It has been shown to be important for the activity of such promoters in the Gram-positive Streptococcus (Stassi et al, 1982), as well as for other types of Bacillus promoters (Banner et al, 1983). Experiments using E. coli 'phage T7 promoters indicate that the AT-rich region may

also be important for $E\sigma^{43}$ function in B. subtilis (Peschke et al, 1985). This is discussed in more detail in I.3.9.

Quite commonly, more than one copy of the hexanucleotide consensus sequences are found close together:

$E\sigma^{43}$ has been shown by methylation and DNAase I protection to bind to both of the 2 almost identical copies of the promoter consensus found both at the usual distance from the transcriptional initiation site of the veg gene, (which had been firmly established in vitro), and also centred about -50 and -75 (Moran et al, 1982a; LeGrice & Sonenshein, 1982). The upstream site lies between 2 arms of dyad symmetry; a possible additional regulatory signal in vivo. This second copy is transcriptionally active when separated from the first (Peschke et al, 1985). It may serve to sequester RNA polymerase molecules and so improve promoter strength in vivo. Similar tandem promoters are found in front of inter alia rrnO and rrnB (Stewart & Bott, 1983).

At the onset of sporulation there is a marked decrease in σ^{43} activity. However, the transcription of vegetative genes continues and their mRNAs comprise 60-80% of the total in sporulating cells (Doi, 1982). Extracts of sporulating cells contain as much of the sigma-43 factor as do vegetative cells, but it fails to copurify with core enzyme. This loss of σ^{43} activity is blocked by spoOA,-OB,-OC,-OE,-OF,-OH,-OJ, $Rfm^{r}Spo^{-}$ mutations, whilst later sporulation lesions allow it (Losick, 1982).

I.3.5 Sigma-37 and -32 Promoters

Sigma-37 is present in both vegetative cells (where levels are low and seem to be nutritionally regulated) and in early sporulating cells where it is abundant, but becomes barely detectable after the second hour of spore formation (Haldenwang & Losick, 1979, 1980; Losick, 1982). Doi has proposed that the role of $E\sigma^{37}$ is to transcribe genes that are catabolite repressed, encode extracellular enzymes, or are expressed only in stationary phase from a comparison of the limited number of

examples of σ^{37}-controlled genes available to date (Wong et al, 1984).

Genes encoding σ^{37} (rpoF or sigB; Binnie et al, 1986) and σ^{32} have recently been cloned. Sigma-32 is the spoOH or sigC gene product (Leighton, personal communication).

In many cases, genes controlled by these sigma factors are transcribed from two closely-spaced start sites, a feature also seen in E. coli for highly regulated genes (eg gal, Musso et al, 1977; Queen & Rosenberg,1981). The conserved -35 and -10 regions that have been defined for σ^{37} promoters often overlap with an active promoter for another of the already characterized holoenzymes. Such tandem or overlapping promoters could give the possibility of gene expression modulation over a long period of growth and development. Some examples follow.

I.3.5.1 P43 Gene: $E\sigma^{37}$, $E\sigma^{43}$; Identical Start

The cryptic P43 gene is transcribed by both $E\sigma^{37}$ and $E\sigma^{43}$ in vitro. In vivo transcripts from exponential cells and from stationary phase T_3, where expression is much stronger, have the identical start site. By comparison with other sequences, overlapping -10 and -35 regions typical of other σ^{37} and σ^{43} promoters can be identified (Wang and Doi, 1984). Concentration-dependent differential use of $E\sigma^{43}$ and $E\sigma^{37}$ possibly accounts for the increased expression of P43 in stationary phase.

I.3.5.2. ctc Gene: Sigmas-37,-32,-29; Identical Start Site; Differential Interaction

The ctc gene (probably spoVC, whose gene product is required for the transition from stage IV to V in spore formation) is actively transcribed under conditions of nutrient deprivation, but is independent of spoO functions. It is expressed at a low level in vegetative cells. The utilization of the ctc promoter by both $E\sigma^{37}$ and $E\sigma^{29}$ was established (Haldenwang & Losick, 1980; Haldenwang et al, 1981) and analysed in great detail by

Moran and coworkers (Tatti & Moran, 1985; Tatti et al, 1985; Tatti & Moran, 1984). Although an insertion mutation in the gene for sigma-37 did not impair either growth or sporulation, the transcription of *ctc* was adversely affected (Binnie et al, 1986).

In vivo transcription (T_4 RNA) was compared with an *in vitro* $E\sigma^{37}$-specified transcript, and start points were found to differ by only a few bases. At -10 and -35, sequences homologous to other σ^{37} promoters were found, and these have been shown by DNAaseI foot-printing (Moran et al, 1981a) and by methylation protection experiments (Moran et al, 1982b) to lie within the $E\sigma^{37}$ binding site. Site-specific mutagenesis of the bases within the putative conserved region affected the efficiency of promoter utilization most dramatically, whereas base changes at positions outside produced modest effects. It thus seems that both -10 and -35 regions are important for recognition (Tatti & Moran, 1984).

Both $E\sigma^{29}$ and $E\sigma^{32}$ produce an *in vitro* transcript with the same start site as does $E\sigma^{37}$, it was subsequently established. Consensus recognition sequences for these holoenzymes are, however, very different (see Table 1). This apparent paradox has been resolved by studies with promoters altered by site-specific mutagenesis.

It was clear that the 3 holoenzymes interact differentially with nucleotides in the same promoter area: some positions are vital for utilization by one form, unimportant for another.

In accordance thus with expectation, a change at -15, where the G-->A transition respectively reduced and increased homology to the σ^{37} and σ^{29} consensus -10 regions, led to a negative effect on $E\sigma^{37}$ utilization, but stimulated transcription by $E\sigma^{29}$.

Most dramatically, a mutant *ctc* promoter with 4 base substitutions in and about the -10 region all but abolished utilization

by $E\sigma^{37}$, but stimulated $E\sigma^{29}$ activity to at least 5x wild-type levels. Mutations at -16 and -36 however, depressed transcription from both σ^{29} and σ^{37} (Tatti & Moran, 1985). Substitution of the absolutely conserved A at -12 (underlined and aligned for the 3 consensus -10 regions in Table 1), had a predictably negative effect on transcription by all three holoenzymes $E\sigma^{37}$, $E\sigma^{29}$, and $E\sigma^{32}$ (Tatti et al, 1985).

In the course of development, the ctc gene is probably transcribed in tandem by first holoenzyme containing σ^{37} and σ^{32}, then by $E\sigma^{29}$.

I.3.5.3 spoVG Gene: Sigmas-37,-32,-29; Different Start Sites

This gene, also known as 0.4 kb, is transcribed at a low rate in vegetative cells, but is active in early and late sporulation, where its product participates in spore coat maturation (transition to stage VI). Unlike the ctc gene, spoVG expression in vivo, but not in vitro, requires 6 of the 7 spo0 functions tested (Ollington et al, 1981; Zuber & Losick, 1983).

In vitro, the gene is transcribed by $E\sigma^{37}$ (Haldenwang & Losick, 1980), $E\sigma^{32}$ (Johnson et al, 1983) and by $E\sigma^{29}$ (Haldenwang et al, 1981). Transcription by $E\sigma^{37}$ and $E\sigma^{32}$ has been studied in detail and proceeds from individual initiation sites separated by 10 bp (in contrast to their shared start site in ctc) (Johnson et al, 1983). In vivo stationary phase transcription is similarly initiated at 2 closely-spaced sites and these differ from the in vitro start sites by only a few bases (Moran et al, 1981b).

An AT-rich stretch lying between -50 and -70 from the downstream initiation site was shown to be necessary for utilization of the downstream promoter (σ^{32}) in particular, in vitro (Moran et al, 1981b), as well as for the in vivo expression of spoVG. Deletion of these sequences abolished the inhibitory effect on sporulation otherwise exerted by the gene in a high copy plasmid (Banner et al, 1983). $E\sigma^{37}$ uses the upstream ini-

tiation site exclusively, $E\sigma^{32}$ the downstream.

Losick has suggested that competition between $E\sigma^{37}$ and $E\sigma^{32}$ for the promoter may modulate both the level of transcription, as well as possibly the translatability of the mRNAs, which differ in length at the 5' end (Johnson et al, 1983). Differential expression of the cistrons in the gal operon of E. coli is regulated in this manner (Musso et al, 1977; Queen & Rosenberg, 1981).

Thus far, only the 2 genes ctc and spoVG have been identified as being transcribed by $E\sigma^{32}$. Two regions that lie upstream from the downstream (σ^{32}) initiation site in spoVG are suggested as constituting -10 and -35 regions for $E\sigma^{32}$. Typical σ^{37} recognition sequences lie appropriately before the upstream start site. The hexanucleotide regions form an alternating "mosaic" of -35,-10 regions before spoVG. In contrast, in the case of spoVC/ctc, where the same highly conserved sequences occur, but where there is a single initiation site for both holoenzymes, there is considerable overlap of -10 regions, but not -35 regions, due to a different spacing for both promoter sequences from that found in spoVG (Johnson et al, 1983; Tatti & Moran, 1984). Later data from mutagenized ctc promoter studies, where substitution of the "invariant A" at position -12 compromises utilization by $E\sigma^{37}$, $E\sigma^{32}$, and $E\sigma^{29}$, can most likely be extrapolated to the spoVG gene.

I.3.5.4 Subtilisin Gene: $E\sigma^{37}$, Two Initiation Sites

Doi and coworkers have shown that the B. subtilis subtilisin gene is transcribed from the same 2 start sites, separated by some 15bp, both in vivo and in vitro (see Fig. 33), and that this is specified by $E\sigma^{37}$. The gene cannot be transcribed in vitro by $E\sigma^{43}$ or $E\sigma^{29}$ (Goldfarb et al, 1983; Wong & Doi, 1984; Wong et al, 1984).

The downstream initiation site is preceded by a typical σ^{37} promoter; whether utilization of the upstream site is due to

$E\sigma^{37}$ or to a minor contaminant of the holoenzyme preparation is unclear. As sigma-37 is present throughout exponential growth, but subtilisin only expressed post-exponentially even when the gene is on a high copy plasmid, additional control mechanisms must operate: some spoO functions are required (see I.6.5.3), and a region of dyad symmetry reminiscent of operator regions in E. coli (Rosenberg & Court, 1979) includes the σ^{37} consensus.

I.3.5.5 B.licheniformis APase: Sigmas-37,-43: Individual Genes

An interesting example of σ^{37} utilization was found in B.licheniformis, where alkaline phosphatase synthesis is under complex control. Low amounts of a membrane-bound form are synthesized during logarithmic growth, whereas large quantities of a form, which is similar by many criteria, are secreted in stationary phase under conditions of inorganic phosphate limitation. Hulett has cloned 2 copies of the gene, which are situated in tandem in the chromosome, and are transcribed independently by B. subtilis $E\sigma^{43}$ (APase I) or $E\sigma^{37}$ (Apase II) in vitro. It remains to be seen if these observations can be correlated with the distinct patterns of vegetative and stationary phase expression (Hulett et al, 1985).

I.3.6 Sigma-28 Promoters

First isolated as a polymerase form uniquely able to transcribe one T7 promoter, the holoenzyme $E\sigma^{28}$ displays a most stringent template preference. Two clones from a gene bank were found that are actively transcribed by $E\sigma^{28}$ (Wiggs et al, 1981). Analysis of in vitro transcripts revealed that the 2 genes had totally conserved -10 and -35 regions (see Table 1); in addition both had the tetranucleotide TAGA at -1 to +3 (Gilman et al, 1981). A quantitative S1 protection assay has shown that both cloned loci are transcribed in vegetatively growing cells to give about 10 RNA copies per cell, from the same transcription initiation site that had been ascertained by in vitro transcription. $E\sigma^{28}$ transcription is to some extent nutritionally regulated during growth: at glucose levels that inhibit

sporulation (viz. 3%), or with only AAs as carbon source, activity is lower than with metabolizable carbon sources glucose or glycerol. Brief nitrogen or phosphate starvation has no effect. The abundance of specific transcripts decreases during the first hour of sporulation, and they are almost undetectable after the second hour. σ^{28} transcripts are absent from spoOA,-OB,-OE and -OF mutants, but are present in wild-type levels in strains carrying extragenic spoO suppressors (rvt).

These observations suggest that sigma-28-controlled genes may be involved in the triggering of spore development.

The finding that $E\sigma^{28}$ is present in stationary phase spoOA cells even though its transcription products are undetectable, indicates that this spoO function may be required to activate the specific promoters in vivo, or to block an inhibitor of σ^{28} (Gilman & Chamberlin, 1983).

Using in vitro transcripts of total B. subtilis DNA by $E\sigma^{28}$, additional strong promoters have been isolated from a gene bank. As it is known that some promoters though poor in vitro are nevertheless powerful in the cell, there may well be important "weak" promoters that will not be detected by such a screening procedure (Gilman et al, 1984). It is estimated that only some 20-30 strong sigma-28 promoter sites occur on the chromosome and it was proposed that they encode functions that sense nutrient deprivation (eg. membrane receptors) and manage the transition from vegetative to sporulative phase (Gilman & Chamberlain, 1983). One sigma-28-transcribed gene has been shown to encode a protein that would seem to be flagellin (Leighton, personal communication).

I.3.7 Sporulation-Specific Sigma-29

During sporulation the amount of $E\sigma^{43}$ and $E\sigma^{37}$ that can be extracted from B. subtilis declines concomitantly with the appearance of σ^{29}-containing polymerase (Haldenwang et al, 1981). Sigma-29 is the product of the spoIIG (or sigE) gene, and seems

to be synthesized from an inactive precursor, P^{31}. It appears early in sporulation, for which spo0A,-0B,-0E,-0F, and -0H, spoIIA,and -IIE functions are required, but not spo0J. This sigma factor persists for only 2 hours (from T_2 to T_4) and sporulation can proceed only as far as stage II in its absence (Trempy et al, 1985a & b). Proteins of similar size that have antigenic similarity to sigma-29 of B. subtilis are found in sporulating cells of several Bacillus species (Trempy & Haldenwang, 1985).

It is clear from the above that σ^{29} is required at a time later than the period of its maximal synthesis, and there are examples of other such gene products (Trempy et al, 1985). No sigma-like factors have yet been isolated from cells at an advanced stage of sporulation.

Errington et al (1985) pointed out the similarity of the spoIIAC sequence with that of spoIIG, and this was even more striking when 2 sequencing errors were rectified (Stragier, 1986). It is likely that this gene codes for another sigma factor of approx. M_r29,000 in early sporulation and which might have been overlooked. It appears at about the same time as the spoIIG product, is also spo0A dependent and shares the conserved putative core- and DNA-binding regions found in other sigma factors (see Table 3b).

A number of genes, some cryptic, that are transcribed in vitro by $E\sigma^{29}$ have been identified in the cloned replication-origin proximal purA-cysA cluster, and its highly conserved promoter has been defined (Haldenwang et al, 1981; Tatti et al, 1985). The genes include spoVG and ctc, already mentioned (I.3.5.3 and I.3.5.3), the veg gene (also transcribed by $E\sigma^{37}$ & $E\sigma^{43}$) and the L gene.

Site-specific mutagenesis of the ctc gene has demonstrated the importance of the absolutely conserved A at -12 (underlined in

Table 1): promoter utilization by σ^{29} (and also by σ^{32} and σ^{37}) was greatly compromised when this conserved base was replaced with T. The highly conserved CA at -15 and -14 respectively, give more efficient utilization than when either position is mutated to G (Tatti et al, 1985).

Experiments with other mutant promoters has shown that both -10 and -35 regions are important for recognition by $E\sigma^{29}$ (Tatti & Moran, 1985).

Latest results suggest that the consensus -35 is not identical with the most efficiently used sequence (Moran, personal communication).

The spoIID gene has been cloned, and is transcribed exclusively by $E\sigma^{29}$ *in vitro*. *In vivo* transcription is confined to a period immediately following the onset of sporulation. It is absent in those spo mutants that do not produce sigma-29, but occurs in others which do not affect sigma synthesis (Rong et al, 1986).

It seems likely that sigma-29 serves to transcribe genes whose activity is required during the period in which it is present, rather than transcribe a select set of genes. With the exception of the L gene and spoIID, sigma-29-transcribed genes seem also to be activated by other holoenzymes at different times in growth (Ray & Haldenwang, 1986).

I.3.8 Other Sigmas, Transcriptional Regulatory Factors

It seems likely that additional low-abundance sigmas remain to be discovered, and the way in which other factors, such as spo0 functions, interact to influence transcription needs to be elucidated. Cloning of genes coding for both sigma factors and spo0 functions is beginning to throw some light on the hitherto obscure processes involved (see also I.6.5).

The "cascade" model (Losick & Pero, 1981) suggested that the minor sigma factors would sequentially confer changes in promoter specificity on RNA polymerase, and would so control the

sporulative program. Since then it has become clear largely from isolation of a variety of intergenic suppressors of sporulation defects that sporulation is not the inflexible dependent sequence of events previously thought, and a more general regulatory role has been proposed for sigma factors (Price & Doi, 1985).

There are abundant reports in the literature of genes with transcriptional start sites that are not preceded by sequences homologous to a hitherto defined promoter consensus.

For example, the 0.3 kb gene is activated at an intermediate stage in sporulation (T_4, stage III). The gene is apparently not in an operon, as a typical terminator lies at -50. It cannot be transcribed *in vitro* by $E\sigma^{43}$, $E\sigma^{37}$ or $E\sigma^{29}$ (Stephens et al, 1984). Induction of 0.3 kb-lacZ fusions is blocked by spoOB, spoIIA, spoIIE and spoIII. Centered about -5 and -27, this gene has sequences highly homologous to similarly centered sequences upstream of the *B. thuringiensis* crystal protein "upstream" transcriptional start, which is active at roughly the same stage of sporulation. These regions are quite distinct from other characterized promoter sites. It has been demonstrated that crystal protein synthesis is transcriptionally regulated, and that early in sporulation a different transcriptional initiation site (lying 16 bp downstream) from the one mentioned above, is used (Wong et al, 1983). This early site shows modest homology to the σ^{29} consensus at -10 but not at -35. There is evidence from *E. coli* that genes that are positively regulated by an effector in addition to polymerase often lack the -35 region homology, eg. λ PRE. Where the -35 homology *is* present, a basal low level of expression is seen, as one would expect, in the absence of effector eg. λ PRM (Rosenberg & Court, 1979).

Another example is the genes coding for the small, acid-soluble spore proteins (SASPs) of *B. megaterium*, and *B.subtilis*, which are under transcriptional control and are co-ordinately

expressed only during sporulation. These have highly conserved regions upstream from the transcriptional start that do not resemble promoters for other genes viz. an exceedingly AT-rich tract, followed by a -35 of general form T/A T/A G T/A A T/A and a -10 region G G/C A A A C (Fliss & Setlow, 1985; Connors et al, 1986; Fliss et al, 1986).

While *in vitro* transcription (albeit poor) of *spoVG*, as well as *ctc* can be directed by $E\sigma^{37}$, *in vivo* expression of the former requires *spoOA*,-*OB*,-*OE*,-*OF*,-*OH* (Ollington et al, 1981) and *spoOK* (Zuber & Losick, 1983). Their gene products may act directly either as positive regulators, or to inactivate repressors. Since having the *spoVG* promoter region on a multicopy plasmid in a wild-type background inhibits sporulation, it is likely that it titrates a low-abundance positively-acting regulatory protein (Banner et al, 1983). The requirement for *spoOA* and -*OB* can be suppressed either by propagating *spoVG* on a plasmid, or by having the gene in an *abrB* background (an intergenic suppressor). These two options do not however eliminate the need for the *spoOH* gene product (viz. sigma-32)(Losick, personal communication).

Identification of other transcription modulating factors, and elucidation of their roles is being studied by the use of gene fusions, where a gene whose product is easily assayed is put under the control of a regulated gene, eg *spoVG*/*lacZ* (Zuber & Losick, 1983) and *ctc*/*CAT* (Truitt et al, 1985).

Using the latter plasmid-borne *ctc* fusion has led to the isolation of non-copy number chromosomal mutants with altered chloramphenicol resistance. Mutants with increased *ctc* transcription of which some, but not all were Spo^-, were found, and where specifically either early ($T_{1.5}$) or late (T_5) transcription was affected. This method also allows the detection of mutants with *reduced* $E\sigma^{37}$ transcriptional activity. Several of the mutations affecting *ctc* expression map around *sigB*, the gene encoding sigma-37 (Binnie et al,1986).

I.3.9 Transcriptional Regulation by RNA Polymerase Modification: a Barrier to Interspecies Expression?

From the above discussion, it is clear that there is, in B. subtilis, a requirement for a variety of specific factors apart from RNA polymerase core enzyme in order for transcription of some non-inducible genes to occur. These requirements differ from one class of genes to another. Obviously the absence of such factors, or others that can compensate for them and allow promoter recognition, may well pose a barrier to transcription of some Bacillus-derived genes in other species, or indeed in certain Bacillus regulatory mutants.

Expression of heterologous genes in prokaryotic host strains is one of the major concerns of the gene technology industry. The use of some Gram-positive organisms, including Bacillus spp., as hosts appears especially attractive due to their effective protein secretory system. It is therefore relevant to examine available data on interspecies expression from the viewpoint of transcription signals and machinery. For this purpose, the best understood representatives among the Gram-negative and -positive eubacteria respectively, are E. coli and B. subtilis.

While B. subtilis genes are as a rule phenotypically expressed in E. coli (Chi et al, 1978), the converse is rare (Ehrlich, 1978). One example of the latter is the thymidylate synthetase gene. When the heterologous gene became integrated into the B. subtilis chromosome, it complemented a Thy^- phenotype, presumably having come under host control (Rubin et al, 1980).

Heterologous gene expression in B. subtilis has effectively been limited to genes from other Gram-positive bacteria, eg. Staphylococcus aureus (Gryczan & Dubnau, 1978).

As regards promoter recognition sequences, it has been mentioned that the most frequently found promoter type in B. subtilis viz. σ^{43} recognition sequence, is essentially identical to the conserved consensus for the E. coli σ^{70} holoenzyme, but that the former seems to have additional requirements over and above the conserved -10 and -35 regions.

In keeping with these observations, E. coli polymerase can transcribe some B. subtilis chromosomal and phage genes with σ^{43} consensus sequences more efficiently even in vitro than does the homologous enzyme (Moran et al, 1982a; Murray & Rabinowitz, 1982; Lee et al, 1980). Conversely, $E\sigma^{43}$ is able to transcribe from E. coli-derived promoters, such as lacUV5, but only very inefficiently even at low ionic strength, where the enzyme becomes less selective (Shorenstein & Losick 1973b; Wiggs et al, 1979, Lee et al, 1980; Peschke et al, 1985).

An AT-rich stretch at -40 to -50 is a feature of many B. subtilis σ^{43} promoters (Moran et al, 1982a) and is found occasionally for E. coli chromosomal and phage genes. When Gram-negative promoters of this type are assayed, they are often utilized extremely efficiently by the vegetative B. subtilis holoenzyme, with initiation taking place at a site identical to that used in E. coli (Peschke et al,1985).

More surprisingly, it was recently demonstrated that a promoter used exclusively by the minor $E\sigma^{37}$ form in B. subtilis, viz. the subtilisin aprE promoter, can be utilized both in vivo and in vitro by E. coli RNA polymerase, with initiation taking place faithfully at the same two initiation sites used in B. subtilis. It is not known whether a minor contaminating form of E. coli polymerase ($E\sigma^{70}$) is responsible for the activity (Wong & Doi, 1984). Such forms are known to exist - see I.3.11. The failure therefore to obtain phenotypic expression of B. subtilis or other Gram-positive-derived genes in E. coli even where transcription does occur, has convincingly been shown to lie

at the translational level. This barrier to expression is considered separately in I.7.3 and I.7.4.

I.3.10 Homologies Between Major Sigma Factors

From the activities discussed above it might be anticipated that similarities in the E. coli and B. subtilis RNA polymerases (major species) themselves, specifically in the respective promoter selectivity-conferring sigma factors, would be found. Reconstitution experiments with heterologous combinations of sigmas and cores yield active enzyme (Shorenstein & Losick, 1973b). Thus while the peptide maps of σ^{43}, σ^{37}, σ^{29}, delta, P^{34} and P^{23} differed from one another (Haldenwang & Truitt, 1982; Wong & Doi, 1982), those of σ^{43} (B. subtilis) and σ^{70} (E. coli) revealed a limited homology. This was also manifested by immunological cross-reaction of anti-σ^{43} with σ^{70}, but not with the other polypeptides tested (Wong & Doi, 1982).

The genes coding for $\sigma^{55/43}$ and σ^{70} (rpoD) have been cloned and sequenced (Price et al, 1983; Gitt et al, 1985; Burton et al, 1981), and they display a strong homology, despite the difference in size of the gene products. When the deduced AA sequences are aligned at the C-terminus, the homology between σ^{43} AA140-371 and the σ^{70} AA 360-613 is 84%, including conservative changes (Gitt et al, 1985) (see Table 3a&b).

The rpoD genes are moreover found in similarly organized operons of 3 genes, but the regulatory elements of the operons differ (Wang & Doi, 1986).

I.3.11 Homologies Between Minor Sigma Factors

In the E. coli rpoD operon, rpoD is preceded by dnaG (primase) (Burton et al, 1983). In response to heat shock, a promoter within dnaG is induced from which rpoD is transcribed by $E\sigma^{32}$ (Taylor et al, 1984).

This recently discovered minor sigma factor (σ^{32} or HptR) is

TABLE 3: HOMOLOGIES BETWEEN SIGMA FACTORS

(A)

SIGMA	Total AA	Position	% Homology	Position	% Homology	Position	% Homology	Position	% Homology	Ref.
Ec σ^{70}	613	372-613	84*	403-417	100	326-613	43*	375-439	68*	Gitt et al, 1985
Bs σ^{43}	371	140-371		77-90		1-284		58-122		Landick et al, 1984
Ec σ^{32}	284									Stragier et al, 1984
Bs σ^{29}	239									

Ec E. coli

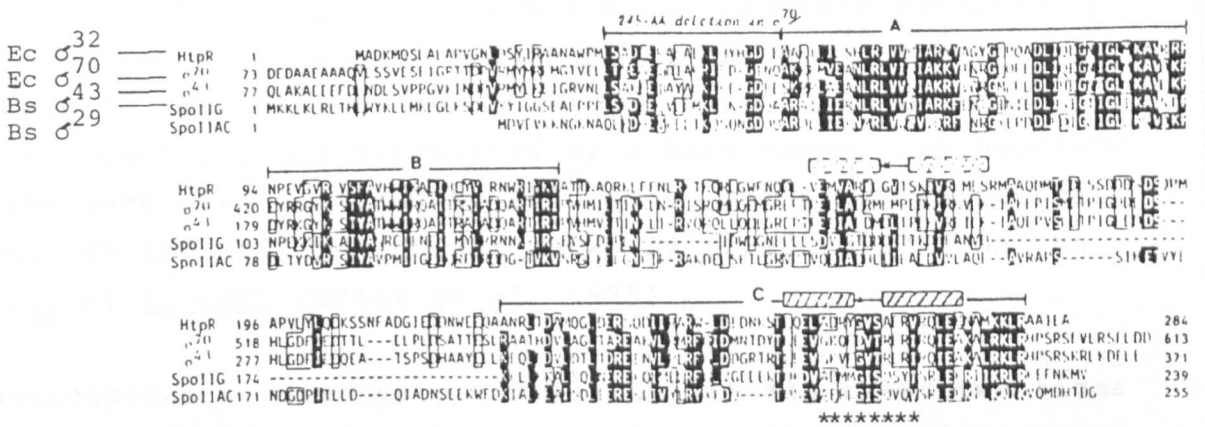
Bs B. subtilis

* Total homology, i.e. includes identical as well as conservative replacements

******* "DNA BINDING SITE"

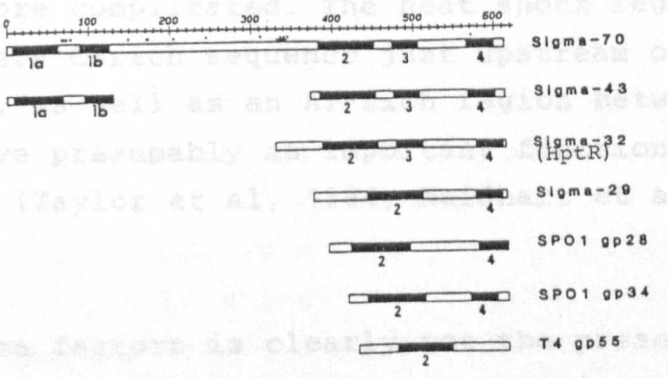
		Ref.
Ala/GlyNNNGlyNNNNNVal/Ile		
Ec σ^{70}	577-587	as above
Bs σ^{43}	336-346	
Ec σ^{32}	158-168	
	257-267	
Bs σ^{29}	210-217	Stragier (1986)

(B) FROM STRAGIER (1986)



Alignment of the amino acid sequences of the *E. coli* HtpR and σ^{70} proteins with the *B. subtilis* σ^{43}, SpoIIG and SpoIIAC proteins. A 245-amino-acid deletion corresponding to residues 130–374 has been introduced in the σ^{70} sequence at the position shown by the arrow. The one-letter amino acid notation is used. Chemically similar amino acids are defined as the following groups: D and E; K and R; S and T; F and Y; I, L and V. Gaps have been introduced to maximize the alignments and are shown by hyphens. White letters on a black background indicate positions occupied by identical or chemically similar amino acids in 4 or 5 proteins. Positions at which 3 residues are identical or similar are boxed. The homologous regions discussed in the text are indicated by brackets (A, B and C). The locations of potential α-helix–β-turn–α-helix structures typical of DNA-binding proteins are schematized.

(C) FROM GRIBSKOV and BURGESS (1986)



Location of Highly Conserved Regions.
The shaded section of the boxes indicate the locations of the highly conserved regions discussed in the text. The locations of amino acid differences between E. coli and S. typhimurium sigma-70 are indicated by dots above the sigma-70 bar. The location of the rpoD800 mutation is shown as a short line above the sigma-70 bar.

the product of the htpR/rpoH gene, and controls heat shock promoters in E. coli (Grossman et al, 1984). Homology exists between the deduced AA sequences of σ^{70} and σ^{32} (Landick et al, 1984; see Table 3).

In B. subtilis, the entire rpoD operon is transcribed during growth from the σ^{43}-type P1 and P2. During stationary phase transcription is from P3, which lies 130 bp further upstream. The rpoD gene here too is preceded by a primase gene (dnaE) (Wang & Doi, 1986).

Examination of respective recognition sequences shows that B. subtilis σ^{28} and E. coli σ^{32} have similar promoter specificity. Transcription in E. coli of cloned B. subtilis σ^{28}-specific promoters is dependent on the $HtpR^+$ phenotype, but is neither dependent on, nor stimulated by a heat shock, and proceeds from the same site utilized by $E\sigma^{28}$ in B. subtilis. Reciprocal recognition and utilization was shown by purified $E\sigma^{28}$ on the rpoD P_{HS} of E. coli (Briat et al, 1985).

One anticipates a reasonable similarity between these two sigma factors, but this must await the cloning of the gene for sigma-28.

It is clear however from above experimental data that, in the same way as other factors apart from a canonical promoter and the appropriate sigma factor are sometimes required for expression of B. subtilis genes in vivo, so the E. coli heat shock promoters are more complicated. The heat shock regulated genes all have a very C-rich sequence just upstream of the -35 and/or -10 region, as well as an AT-rich region between -40 and -60. These have presumably an important function in heat shock regulation (Taylor et al, 1984; Neidhart et al, 1984).

A multiplicity of sigma factors is clearly not the preserve of Bacillus: other Gram-positive differentiating organisms have

also proved to have several RNA polymerase forms eg. Streptomyces coelicolor (Westpheling et al, 1985), and it is anticipated that, as for HtpR, sigmas will be found for special sets of genes having in common conserved promoters eg nif genes of Klebsiella (Buck et al, 1986).

In Table 3A&B the regions of significant homology between B. subtilis and E. coli sigma factors are shown. There are two main regions that seem to be conserved. One is presumably the RNA polymerase core binding area, the other shows the parallel α-helices which are a general feature of well-characterized DNA-binding proteins (Takeda et al, 1983).

I.3.12 Homologies Between 'Phage-Encoded Sigma Factors

Among the 'phage-encoded sigma factors the SPO1 σ^{gp28}, and $\sigma^{gp33-34}$ genes (whose gene products seem to act synergistically) have been cloned (Constanzo & Pero, 1983; Constanzo et al, 1984 resp.). The latter 2 genes overlap by 4 bp.

Surprisingly, while gp^{28} competes with host σ^{43} efficiently for core polymerase also in vitro (Chelm et al, 1982), the two proteins share little or no homology (Gitt et al, 1985). Similarly, the M_r 28,000 peptide (analogue of σ^{gp28}) of SPO1-related phage SP82 can associate with E. coli RNA core polymerase to allow specific binding to and transcription from middle gene promoters (Achberger & Whiteley, 1980). σ^{gp28} has however very limited homology with the major E. coli σ factor, σ^{70} (Burton et al, 1981; Constanzo & Pero, 1983). gp^{33} and gp^{34} showed neither significant homology to each other nor to gp^{28}, which is rather larger.

However, Gribskov & Burgess (1986) have made a more sophisticated statistical analysis of the 7 known sigma factor sequences, and find that all belong to a family of homologous proteins. Each protein could be shown to be significantly related to at least one of the others, for each element of structure tested. Though not included in this study, sigma-37 clear-

ly also belongs to this family (Binnie et al, 1986).

Conserved structural features are shown in Table 3C. Region 1, highly conserved and present only in sigmas -43 and -70, is proposed to account for the weaker binding of the major sigma factors to core, relative to the high affinity shown by the minor or phage-encoded forms. Region 2 is proposed to involve core binding. Sigmas -70, -43 and HtpR each have 2 copies of the helix-turn-helix characteristic of DNA-binding proteins (Regions 3 & 4). In the other sigmas, only one such was detected by the analytic method employed. These are proposed to interact directly with -10 and -35 regions, possibly in co-operation with other polypeptides.

Isolation of mutants with predictable phenotypes would strengthen these hypotheses.

I.3.13 Sigma Factors as Regulatory Targets

Mutations in the E. coli rpoD (sigma-70) gene have been identified as increasing the expression of araBAD in the absence of CAP-cAMP (Hu & Gross, 1985). Although the basis for catabolite repression in B. subtilis is different, the crsA47/ rpoD47 mutation similarly abolishes the glucose repression of sporulation in B. subtilis (Takahashi, 1979; Price & Doi, 1985) and suppresses spoOB, spoOE, spoOF, spoOK mutations (Leung et al, 1985).

I.4: General Regulation of Gene Expression in Bacillus

I.4.1 Control of Exoenzyme Synthesis

Many genes, including those coding for the sporulation process, and exoenzymes, are under general regulation eg catabolite repression, and possibly stringent control.

Together with sigma factor modification of RNA polymerase, these general mechanisms either directly, or indirectly by their effect on sporulation, together with other regulatory loci combine to control exoenzyme synthesis in Bacillus.

The map shown in Figure 3 indicates the chromosomal location of some of these loci.

Below is a summary of currently-held perceptions of how these different factors contribute to exoenzyme regulation.

I.4.2 Catabolite Repression

In B. subtilis, provided that a utilizable nitrogen source is available, glucose (or another rapidly metabolizable carbon source) will inhibit sporulation (Schaeffer et al, 1965). The addition of glucose after $T_{0.5}$ however (i.e.later than 30 mins. after the cessation of exponential growth) has no effect i.e. the cells are then "committed" to sporulation, and resistant to glucose repression (Schaeffer, 1969).

Glucose also exerts an inhibitory effect on exoenzyme synthesis, both directly and indirectly (reviewed by Priest (1977)) and an expression of a number of genes unrelated to sporulation eg. that for acetoin dehydrogenase (Takahashi, 1979).

The mechanism of glucose repression in bacilli is obscure relative to that in E. coli, but even in this system much is still conjectural.

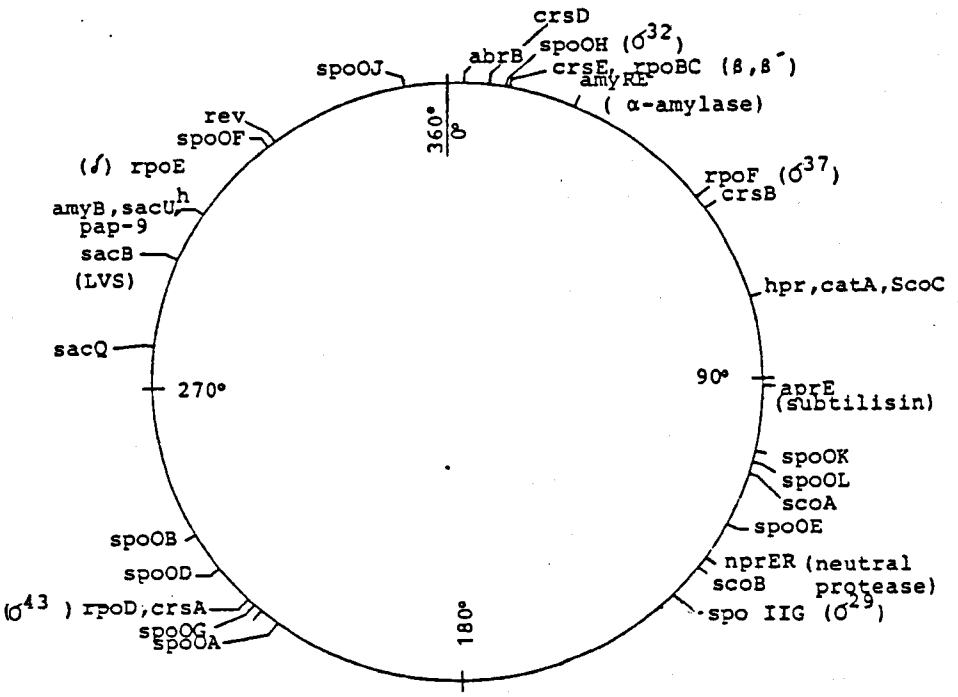
I.4.2.1 Cyclic Nucleotides

In E. coli, the cAMP intracellular concentration rises when glucose is limiting. cAMP together with its receptor protein CRP or CAP, acts as a positive effector for genes sensitive to catabolic repression, binds to the gene close tc the transcriptional start site, and facilitates RNA polymerase binding (Rickenberg, 1974). However, cAMP seems not to be a universal regulator and has not been found in B. megaterium (Setlow, 1973), B. licheniformis (Bernlohr et al, 1974) or B. subtilis (Hanson, 1975).. Furthermore, enzymes for its synthesis and degradation were not detectable in B. pumilis, B. cereus, B. subtilis (Ide, 1971) or B. megaterium (Setlow, 1973).

FIGURE 3

PARTIAL GENETIC MAP OF B. SUBTILIS

adapted from Piggot & Hoch (1985)



Some of the structural genes mentioned in the text, and regulatory loci influencing their expression, are indicated.

Another cyclic nucleotide besides cAMP viz. cGMP has also been implicated in catabolite repression. It seems to act as an antagonist of cAMP: Levels of intracellular cGMP in E. coli were inversely proportional to cAMP, and showed a direct relationship to growth rate (Bernlohr et at, 1974). A direct connection between cGMP levels and expression of catabolite-repressed enzyme beta-galactosidase was demonstrated in vivo: addition of cGMP inhibited transcription in intact cells and this required functional cAMP receptor protein (Artman & Werthamer , 1974). A mutant CRP has also been found, which was altered so as to be activated in vitro by cGMP instead of its normal effector cAMP (Sanders & McGeoch, 1973).

In contrast to cAMP, cGMP has been found in B. licheniformis. As with E. coli, levels are proportional to the growth rate, but there is also a sharp increase in sporulating cells (Bernlohr et al, 1974).

I.4.2.2 Catabolite Repression of Inducible Enzymes

glnA and gltA mutants have been found where several genes involved in carbohydrate metabolism, and normally under catabolite control, are pleiotropically affected (Fisher & Sonenshein, 1984; Fisher & Magasanik, 1984 resp.). It is known that glutamine synthetase, the glnA product, is autoregulated, and that both ATP and glutamine are required in addition to the enzyme in order to repress synthesis (Schreier & Sonenshein, 1986).

It has been demonstrated that in mutants displaying a "catabolite resistant" sporulative ability, catabolite repression of the inducible enzyme acetoin dehydrogenase is not necessarily similarly affected (Takahashi, 1979), suggesting that several mechanisms for catabolite repression may operate.

A clear demonstration that in E. coli and B. subtilis the catabolite repression systems are mechanistically different comes from examination of the expression of the B. subtilis α-amy-

lase gene in the 2 organisms. On a high copy plasmid, in B. subtilis the cloned gene is expressed in the same way as the chromosomal copy viz. post-exponentially and subject to repression by added glucose. The amyR1 allele gra-10 abolishes glucose repression, but permits temporal activation. In E. coli, however, the synthesis of α-amylase from the wt clone is constitutive and not glucose-repressible (Nicolson & Chambliss, 1985).

I.4.3 The Stringent Response, Highly Phosphorylated Nucleotides and Sporulation Initiation

In E. coli under conditions of amino acid starvation, the "stringent response" is invoked to adjust the cell's metabolism to suit environmental conditions. ppGpp and pppGpp are synthesized with the help of "Stringent Control Factor" from GTP (Fiil et al, 1977). The beta subunit of RNA polymerase has been shown to interact directly with ppGpp: some beta mutants are "relaxed" in the presence of high levels of ppGpp (Glass et al,1986).

RelA mutants, defective in stringent control, do not synthesize (p)ppGpp (Cashel & Gallant, 1969). B. subtilis cells starved of either a carbon or nitrogen source, eg. amino acids, will sporulate. Highly phosphorylated adenosine nucleotides were found in sporulating cells, but not under conditions where sporulation is inhibited, and it was suggested that they somehow trigger sporulation. Ribosomes from vegetative cells synthesize (p)ppGpp. However, "relaxed" mutants of B. subtilis were found to sporulate normally even although the highly phosphorylated guanosine nucleotides were not synthesized, so it appeared that the stringent response was not needed for sporulation initiation (Rhaese et al, 1976).

When B. subtilis is starved for amino acid(s), an increase in (p)ppGpp is seen, as well as an inhibition of IMP dehydrogenase, the first enzyme in the synthesis of guanosine nucleotides.

It was demonstrated by the use of specific inhibitors decoyinine and mycophenolic acid (which inhibit GMP synthetase and IMP dehydrogenase resp.), where a drop in GDP or GTP could be induced without a concomitant rise in "magic spots" (p)ppGpp occurring, that sporulation could be triggered in both Rel^+ and relA strains, even in the presence of excess nutrients. It is believed that it is the drop in GTP levels that is the actual initiation trigger (Vasantha & Freese, 1980; Ochi et al, 1981).

It has been suggested that the morphologically altered protease-nonproducing colonies that commonly arise in large-scale fermentations of B. licheniformis are rel^+ revertants (Frankena et al, 1985).

In addition to changes in levels of guanosine nucleotides, factors which can be supplied extracellularly are required for sporulation initiation. Studies with preconditioned media show that factors dependent on spoOC,-OE,-OH, and -OJ, but not spoOA, or -OB are needed, and will allow cells at low density to sporulate (Grossman, personal communication). (Note that spoOC denotes lesions situated C-terminally in the spoOA gene product, which are less pleiotropic in effect than spoOA).

I.5: Specific Regulatory Loci for Exoenzyme Synthesis

Inducible enzymes such as levansucrase (LVS) are not considered separately here. Recent accounts of the control of sacB are available (Steinmetz et al, 1985; Aymerich et al, 1986;Shimotsu & Henner 1986).

I.5.1.1 α-Amylase Production: amyR1:-R2

The 3 strains B. subtilis 168, B. subtilis natto and B. amyloliquefaciens (SAC) secrete α-amylase post-exponentially approximately in the ratio of 1:5:15 respectively. Transformation experiments indicated that a regulatory element that was tightly linked to the structural gene amyE controlled the levels of enzyme produced. Relatively low synthesis was attributed to the amyR1 allele of B. subtilis 168, and the higher production of B. subtilis natto to amyR2 (Yamaguchi et al, 1974b).

The amyR3 regulator of SAC is identical to amyR2 (Yoneda et al, 1974); the exceptionally high production by this strain is due to a mutation at another locus (pap) that also affects other functions (see I.6.1).

The amyR1 locus has been characterized, and comprises about 280 bp upstream of amyE. There are 3 possible promoters, based on examination of the sequence, with 3 typical SD regions within 150 bp of the translational start. Apart from P2 (σ^{43}-type), they are difficult to categorize. Mapping of the transcriptional start has not been reported, but the removal of P3 and all material downstream of it in a transcriptional fusion did not affect either temporal activation or glucose repression of transcription (Nicholson et al, 1985).

The gra-10 allele of amyR1 abolishes catabolite repression selectively, while allowing temporal activation of α-amylase in stationary phase. gra-10 exerts this effect only in cis (Nicholson & Chambliss, 1985; 1986).

The amyR2 locus has also been studied (Yamazaki et al, 1983). In addition to conferring a catabolite repression-resistant expression in the same way as gra-10 (Nicholson & Chambliss, 1986), it has much in common with the regulatory region of the α-amylase gene from B. amyloliquefaciens (Lehtovaara et al, 1984). An AT-rich region including an inverted repeat resembling a transcription terminator was proposed to account for the amylase hyperproduction of amyR2 strains, by elimination of putative interference by transcription read-through from a site far upstream. Removal of the repeat reduced expression of B. amyloliquefaciens α-amylase by 70%. Removal of the promoter of an upstream operon restored high yields, confirming Yamazaki's hypothesis (Kallio, 1986). The amyR2 locus is resistant to catabolite repression but only allows expression of α-amylase in stationary phase. (Nicholson & Chambliss, 1985).

It should be noted that α-amylase synthesis in most bacilli

seems to be partially inducible - yields are higher when the growth medium contains starch (Priest, 1977; Sibakov, 1986).

I.5.1.2 Tunicamycin Resistant Mutants & amyE Amplification

After mutagenesis, tunicamycin resistant mutants with specifically raised α-amylase production were isolated, and the mutation mapped near amyR (Nomura et al, 1978). It has subsequently been shown that amplification of the region of the chromosome carrying both the tmr^R gene and amyE had occurred (Hashiguchi et al,1985, 1986). The sequence that is required for amplification has been identified (Mori et al,1986).

I.5.2 NPR Production: nprR1:-R2

The strain B. subtilis natto secretes about 15-20 times more neutral protease than does B. subtilis 168.

The ratio of NPR/APR secreted by strain natto is 13, and in the case of B. subtilis 168 it is 1. Increased NPR production can be conferred on strain 168 by transformation with DNA from strain natto. Isolation of mutants with altered enzyme etablished that nprR and the structural gene nprE are tightly linked loci (Uehara et al, 1979). By analogy with α-amylase, two regulatory loci nprR1 and -R2 were proposed, where the latter is responsible for hyperproduction (Uehara et al, 1974).

Experimental confirmation of this hypothesis has recently been obtained. The nprR2 locus from an overproducer lacks an area of dyad symmetry (centred about 100 bp upstream from the initiation codon), which is present in nprR1. This is due to the lack of a 66 bp sequence present in the wild-type, and which may be the target site for a regulatory protein. Start points for transcription of npr are apparently the same in both wt and overproducer, but transcription in the latter rises more dramatically after exponential growth is over (Toma et al, 1986).

I.6 Pleiotropic Regulation of Exoenzymes

A number of pleiotropic mutations affecting not just exoprotein production, but also other seemingly unrelated features have been described by numerous authors. A bewildering array of phenotypes has been described, but in recent years it has become clear that many of the independently investigated mutants are alleles of one and the same locus. Difficulties in correlating disparate reports arise as various groups select their mutants and investigate them from a slightly different viewpoint, emphasizing certain aspects more than others, so the same features are not always studied. The cloning of some of these loci is just beginning to contribute to the understanding of the mechanism by which they exert their effect.

I.6.1 pap, sacU, amyB

pap mutants, displaying simultaneously increased protease and α-amylase synthesis, were first described by Yoneda & coworkers. Other features apparently conferred by a single lesion are loss of competence, filamentous growth (Yoneda et al, 1973; Yoneda & Maruo, 1975), decreased autolytic enzyme activity, and loss of flagella (Ayusawa et al, 1975).

While amylase levels were raised 2-3 times, exoprotease activity was increased 5-16 times, and this was shown to involve both major enzymes; although the effect on APR levels was more marked (Yoneda & Maruo, 1975).

A positive synergistic effect on amylase or NPR production is achieved by combining the pap mutation with either amyR2 or nprR2 resp. (Yoneda et al, 1973; Yoneda & Maruo, 1975).

Two amyB sequential mutants both had increased α-amylase secretion, but one was motile and produced lower amounts of exoproteases, while the other was non-motile and had greatly raised exoprotease production (Sekiguchi et al, 1975).

The $sacU^h$ mutation leads to increased levansucrase (50-100

times), exoprotease (5-10 times) and α-amylase (2-3 times) production, to the ability to sporulate under conditions inhibitory to the wild-type, and sometimes to the loss of motility (Kunst et al, 1974).

It has been established that pap-9, amyB and $sacU^h$, with their similar phenotypes, have a common locus (Steinmetz et al, 1976).

The sacU wild-type locus has been cloned, and in E. coli minicells, a protein of M_r46,000 is the gene product. A component of the same size is missing from membrane extracts of $sacU^-$ mutants ($SacU^-$ phenotype: greatly reduced levansucrase levels), and is found in higher than wild-type amounts in $SacU^h$ membranes. This protein is not flagellin (M_r 36,000). Comparison of levansucrase mRNA levels in $sacU^-$ and $sacU^h$ cells suggested that the sacU gene product acts post-transcriptionally (Aubert et al, 1985). However, it has been shown more recently that sacB mRNA levels are in line with LVS secretion in the various sacU backgrounds (Aymerich et al,1986). This data has been supplemented by the demonstration that a presubtilisin translational fusion with lacZ involving only 8 codons of the signal sequence, when present in a single copy in a sacU32 background, gives 7-fold higher levels of β-galactosidase than does the parental strain (Ferrari et al, 1986). If the sacU product indeed affects a post-transcriptional step, the target of action lies 5' of the 8th codon in the signal sequence; initiation of translation is one possibility. Secretion cannot be the regulatory target in the case of this cytoplasmic fusion protein.

A mutant div-341 has been described which grows filamentously at the non-permissive temperature 45°C. At 37°C the mutant has lowered exoenzyme synthesis, sporulation and competence relative to wt. Div-341 maps very close to sacU and may be an allele of the same locus. In a double div-341$sacU^h$ mutant a mixed phenotype is obtained (Sadaie & Kada, 1985).

I.6.2 sacQ

The sacQ36 mutation is less pleiotropic than is $sacU^h$, and leads "only" to increased production of levansucrase, APR, NPR and α-amylase (Kunst et al, 1974).

It was thus believed to have an effect on secretion, as apart from all being exported, the affected proteins are under rather different control: while α-amylase and proteases are produced post-exponentially, LVS is synthesized only in vegetative growth on sucrose.

The gene has been cloned from B. amyloliquefaciens (Tomioka et al, 1985; Yang et al, 1986) and from B. subtilis wild-type and sacQ36 (Yang et al, 1986). The 2 genes are almost identical, and encode a 46 AA polypeptide.

The sacQ36 mutation was found to be an "up"-promoter mutation: it is a single base change at -10 from the transcriptional start, and leads to higher levels of sacQ message in stationary phase. This was demonstrated with a sacQ-lacZ integrated translational fusion (Yang et al, 1986). The alteration does not lie in a sequence resembling a characterized RNA polymerase recognition site.

Hyperproduction of exoprotease can also be achieved with the wild-type sacQ locus on a high copy plasmid, so elevated levels of the sacQ gene product evidently account for the sacQ36 phenotype.

A translational fusion of B. subtilis pre-subtilisin (6 codons of the signal sequence) and β-galactosidase in a sacQ36 background leads to higher levels of the fusion products, suggesting that sacQ36 effects the hyperproduction phenotype at a stage prior to secretion, since here the fusion product is cytoplasmic (Yang et al,1986).

The role of sacQ in wild-type cells is obscure: partial deletion of the coding sequences did not discernably affect wild-type levels of (neutral)protease or α-amylase production, and sporulation was unaffected. The sacQ36 phenotype was however abolished from the mutant, and exoenzyme production returned to wild-type levels.

I.6.3 hpr, scoC and catA

These loci map in the same area and are probably alleles. The phenotypes include hyperproduction of exoprotease and defects in sporulation control eg. delayed, or catabolite- repression-resistant sporulation.

A high protease (Hpr) phenotype (16-30 fold higher) was described where α-amylase production was unaffected. NPR levels were elevated 10-fold, APR about 2-fold (Prestidge et al, 1971; Higerd et al, 1972).

In various hpr mutants, an integrated translation fusion of presubtilisin (first 8 codons) and lacZ, under the control of subtilisin upstream regulatory sequences, gave 16-fold higher β-galactosidase levels than did wild-type (Ferrari et al, 1986).

catA mutants, selected for the ability to sporulate in the presence of glucose, secrete elevated levels of protease and this secretion is likewise not repressible by glucose, whereas histidase remains under glucose control. catA is hypostatic to spoOA,-OB mutations as double mutants are Spo^- and do not secrete protease (Ito & Spizizen, 1973).

Some (not all) spore control mutants (Sco), selected for the ability to secrete protease under catabolite inhibitory conditions, mapped at a locus, scoC that could not be distinguished from catA in genetic crosses (Milhaud et al, 1978). APR production was especially affected. Careful experiments showed that the "catabolite-resistant" phenotype is in fact a matter of degree. While scoC and catA mutants are also sensitive to addi-

tion of glucose to the medium (it delays sporulation and inhibits protease production), they develop resistance to catabolite repression earlier in growth than does wt, and recover more quickly from glucose addition (Dod & Balassa, 1978). The timing of glucose addition was crucial. In general, the glucose effect was less marked in mutants if added prior to To. Addition of glucose between T_o and T_1 actually stimulated APR production in the mutants dramatically. It appeared that exoprotease control was under catabolite control both directly, and indirectly by control of sporulation.

The ability to sporulate under conditions where the wild-type cannot (eg. catA, scoC) can be associated with the hyperproduction of protease also in B. cereus, where mutants able to sporulate in exponential growth have been isolated. These "derepressed" sporulation mutants included many purine auxotrophs, and the other phenotypes were correlated with the auxotrophy. Selection for purine auxotrophs instead of protease overproduction also yielded such mutants (Levisohn & Aronson, 1967). Presumably the GTP levels needed for stringent control are not available (Ochi et al, 1982).

I.6.4 Other Cloned Pleiotropic Loci

There are reports of cloned fragments from other Bacillus spp. that enhance production of exoenzymes. The prtR locus, from B. subtilis natto leads to increased NPR, APR and LVS secretion, as well as causing filamentous growth in the strain carrying the clone in multiple copies. An open reading frame of 60 codons, encoding a highly charged polypeptide (42% charged AAS) is believed to constitute the gene (Nagami & Tanaka, 1986). The resemblance of the PrtR phenotype with $sacU^u$ is striking, but the size of the gene product cannot be correlated with the sacU data of Aubert et al (1985).

A gene from B. licheniformis enhances APR, NPR, α-amylase and RNAase production in B. subtilis cells carrying the locus on a plasmid (Okada et al, 1984).

No other data such as mapping, is as yet available, so the relationship, if any, to the loci already discussed is still unclear.

I.6.5 Stage 0 Mutants

I.6.5.1 Hierarchy of spo0 Loci

There are 10 spo0 genes: spo0A,-0B,-0D,-0E,-0F,-0G,-0H,-0J,-0K, -0L.

The stage 0 sporulation (spo0) mutations have pleiotropic effects on exoprotease and antibiotic production, the development of competence, susceptibility to 'phage, and on transcription by minor forms of polymerase (Brehm et al, 1973 and refs. below). Cloning of these genes, identification of products and lesions, as well as characterization of intergenic suppressors is beginning to elucidate their mode of action.

Spo0A and -0B mutations are the most pleiotropic and were believed to be concerned with sporulation initiation (Hoch, 1976; Hoch et al, 1973). The spo0 loci were placed in a hierarchy by Piggott & Coote (1976).

Table 6: Hierarchy of spo0 Loci

	PROTEASE	ANTIBIOTIC	COMPETENCE
Spo0A	-	-	-
-0C	+	+	+
Spo0B	+/-	-	+
Spo0G	-	-	+
Spo0D	+	-	+
Spo0E	+/-	variable	-
Spo0F	+/-	-	+
Spo0H	+	+	-
	+/-	variable	-
Spo0J	+	+	-
Spo0K	+	-	-

While spo0A and -0C could be differentiated (Michel & Cami,

1969; Hoch et al, 1978) they are alleles of one locus. spoOC mutations are more weakly pleiotropic, presumably since they are now known to be alterations in the extreme C-terminus of the spoOA coding sequence (Ferrari F.et al, 1985).

I.6.5.2. spoO Gene Products: Low Abundance Transcription Activators

It has already been mentioned that spoVG transcription in vivo requires spoO functions in addition to $E\sigma^{43}$, $E\sigma^{37}$, and $E\sigma^{32}$ (Ollington et al, 1981; Johnson et al, 1983; Zuber & Losick, 1983). It is clear that the spoO gene products are present in very low amounts: propagation of the spoVG promoter region on a multicopy plasmid inhibits sporulation (Banner et al, 1983). The level of expression of a spoVG-lacZ transcriptional fusion on a high copy plasmid is only 2-3 fold higher than in cells with one copy of the fusion (Zuber & Losick, 1983). The multicopy phenomena are most easily explained by assuming titration of a trans-acting factor. At the same time, propagation on a plasmid abolishes the need for certain spoO functions (spoOA, -OB),but not for spoOH (which encodes σ^{32}, Leighton, personal communication) and also affects the timing of expression: significant β-galactosidase synthesis is seen prior to T_o (Zuber & Losick, 1983 & personal communication regarding the use of a spoOH deletant). It was suggested that supercoiling of promoter DNA in the plasmid was supplanting the possible actions of spoOA or -OB, but similar disproportionately low levels of β-galactosidase activity were obtained when integrated copies of the spoVG-lacZ fusions were caused to amplify.

It should be noted, however, that having many copies of a spo-regulated promoter does not always eliminate the need for one or more sporulation gene products in order to get expression. One such example is the cat-86 gene of B. pumilus (Mongkolsuk & Lovett, 1984), another is a cryptic gene from an alkalophilic strain that is expressed at T_8 (Kudo et al, 1985). ctc however, which is, like spoVG, transcribed by $E\sigma^{32}$ and $E\sigma^{37}$, is independent of spoO functions (Ollington et al, 1981).

The promoters P28-1 and P28-2 "read" by $E\sigma^{28}$ are dependent on certain spoO (-OA,-OB,-OE,-OF) functions, but this dependence is abolished by intergenic suppressor rvtA11 (Gilman & Chamberlin, 1983). rvtA11 (Sharrock et al, 1984) has exactly the same phenotype as does sof-1 (Hoch et al, 1985) viz. it confers the ability to sporulate in the presence of alcohol, glucose, and suppression of spoOB, -OE and -OF even in the presence of glucose, but not of spoOJ, -OH, -OK. Sof-1 has been shown to be a single base change in the 12th codon of the spoOA gene (Hoch et al, 1985).

The spoOA and spoOF genes have been cloned and code for proteins of M_r 29,000 and 14,000 resp. (Ferrari et al, 1985; Ikeuchi et al,1986; Trach et al, 1985). The spoOF gene product is 56% homologous (AA) with the N-terminal region of spoOA, and 2^o structure predictions are identical for the 2 proteins. Both are homologous to the N-terminal domains of E. coli proteins OmpR and SfrA that are known to regulate transcription of membrane components. In addition, the C-terminal region of SpoOA shows a significant homology to SpoOB. A direct interaction with RNA polymerase is proposed for spoOF and spoOA products.

The synthesis of SpoOF itself is under glucose control (Saito, personal communication): An integrated fusion of spoOF - α-amylase gives α-amylase synthesis only in the absence of glucose or in a crs background. Moreover, low-abundance positive regulatory factors must be involved in activating spoOF: having the gene in high copy inhibits sporulation, for which the gene product is absolutely required. A deletion of spoOF coding sequences abolishes sporulation entirely (Kawamura & Saito, 1983).

I.6.5.3 Effect of spoO Mutations on Subtilisin Expression

The effect of spoO mutations on levels of a translation fusion between the first 8 codons of presubtilisin and β-galactosidase has been examined. The fusion gene was under the control of the subtilisin upstream regulatory sequences, and was present in

only 1 copy per chromosome.

The decrease in expression of β-galactosidase, which expression was strictly confined to stationary phase in all experiments, was qualitatively in accordance with expectations in the various spo0 backgrounds (see "Hierarchy" of Pleiotropy in I.6.5.1). However, the effects were undramatic by comparison with earlier data on characterization of the spo0 mutants themselves: spo0A exerted a 10-fold inhibition, spo0H,-0E,-0F, -0B were similar in giving approx. a 5-fold inhibition, and spo0J none at all with reference to the wild-type strain (Ferrari et al,1986). When the subtilisin promoter (or gene fusion) is present on a high copy plasmid in a spo0A background, β-galactosidase from an integrated (or plasmid-borne) fusion gene is expressed at "wild-type" levels, suggesting the titrating of a negative regulator by the promoter region . Even where spo0A regulation is thus "overcome", another level of control operates to confine expression to stationary phase. Either catabolite repression, or an inducer that first appears in late log phase may be responsible. These results contrast with those from multicopy spoVG-lacZ fusions, where expression is seen in exponential phase (Zuber & Losick, 1983).

I.6.5.4 The Involvement of σ^{43} in Sporulation Initiation

As discussed above, the spo0A allele sof-1 (rvtA11) confers a catabolite-resistant sporulation phenotype (Crs), the ability to sporulate in the presence of alcohol (Ssa phenotype, which otherwise induces a spo0-like phenocopy), as well as being able to suppress spo0B,-0E and 0F mutations fully (Hoch et al, 1985; Sharrock et al, 1984; Kawamura & Saito, 1983).

It has been shown that crsA47,-4,-1 mutations, that likewise confer catabolite-resistant sporulation, relief from catabolite repression of some enzymes, and the ability to sporulate in the presence of alcohol, all involve a two base change resulting in a single AA alteration (Pro-->Phe) in the rpoD gene product viz.

sigma-43. crsA47 will suppress spoOE,-OF,-OK fully, and spoOB, -OJ and -OD partially (Price & Doi, 1985; Leung et al, 1985; Kawamura et al, 1985). In turn, crsA47 can itself be suppressed by some spoOA,-OD,-OF,-OK and non-spoO mutations (Kawamura et al, 1985; Sun & Takahashi, 1985).

I.6.5.5 Other Extragenic Suppressors of Sporulation Mutations: Network Concept

The Crs phenotype (crsC,-D,E,F loci) can be suppressed by the rif-11 allele of the RNA polymerase β-subunit (Sun & Takahashi, 1985). crsE has been shown to constitute the rpoBC operon, encoding subunits β and β' (Leighton, personal communication).

In E. coli, one ts β' mutant displays a "stringent response" as regards rRNA synthesis in the absence of ppGpp accumulation, and can be considered a similar phenomenon (Oostra et al, 1981).

Everything thus points to the action of spo mutations and many of their intergenic suppressors acting via the transcriptional machinery.

Another partial suppressor, abrB, restores some features to spoOA and -OB mutants (but not sporulation ability), and is believed to act via ribosomal protein alterations (Trowsdale et al, 1979). rev-4 partially compensates for sporulation defects caused by alterations in the transcriptional and translational machinery, and is believed to involve membrane alterations (Sharrock & Leighton, 1982).

It seems though, that sporulation is rather flexible, and should be considered to be a regulon, or "network"; it can take place by "bypassing" a step rather than, as was first assumed, being a process involving a rigid consecutive developmental program (Ballassa, 1969). The extragenic suppressors demonstrate that some spoO products are neither unique nor essential. As yet,however, no suppressors of spoOH (σ^{32}) have been isola-

ted.

Thus certain mutations in spoOA and rpoD can render environmental signals (eg. nutrient availability) and other spoO functions redundant and allow sporulation to proceed (Price & Doi, 1985).

Another example of flexibility is the gene product of spoIIG, viz. sigma-29. It can accumulate in one early-blocked strain (spoOJ, but not in most other spoO strains), and yet fails to be expressed in several stage II (and possibly stage III mutants), which have progressed to a much more advanced stage (Trempy et al, 1985b).

I.7: Translational Control of Expression

Physiological control of expression can take place at the level of translation. Two of the many possible regulatory variables are (a) the rate of translation initiation and (b) regulation of elongation rate by codon choice (ie tRNA availability). Both need to be considered when evaluating B. subtilis as a host for heterologous gene expression, for such mechanisms could prove an effective barrier to expression.

I.7.1 Translational Initiation in E. coli

Initiation of protein synthesis by procaryotic ribosomes involves selection of a site on mRNA by the 30S ribosomal subunit.

This site includes the following features: a polypurine Shine-Dalgarno (SD) region (Shine & Dalgarno, 1974) which is complementary to a variable extent to the 16S rRNA 3'-end, an initiation codon, and the appropriate spacing between these features, commonly 3-11 nucleotides (Gold et al, 1981).

There is no correlation between spacer length and the "strength" of the ribosome binding site (RBS); moreover,application of

the structural criteria listed above cannot distinguish functional from non-functional E. coli ATG initiation codons (Stormo et al, 1982).

It is evident that other, less easily defined features influence initiation. The sequence of the "spacer region" can contribute to initiation efficiency (De Boer et al, 1983, Hui et al, 1984). An examination of 29 mRNAs showed that many are characterized by the potential for stem-loop formation , where the SD is found in a non-helical section, but the initiation codon often in a stem (Selker & Yanofsky, 1979). Alterations upstream from the RBS can have dramatic effects on translation efficiency (Iserentant & Fiers, 1980; Queen & Rosenberg, 1981; Hall et al, 1982; Kastelein et al, 1983), and secondary structure considerations have been invoked to explain these observations.

I.7.2 Codon Usage in E. coli and B. subtilis

It is known that in E. coli, iso-accepting tRNA species are present in different molar amounts (Ikemura, 1981). Highly expressed genes are made up of codons corresponding to abundant tRNA spp. This is not the case for weakly expressed genes whose products are needed in relatively low amounts (Grosjean & Fiers, 1982). It has been suggested that codon choice and relative tRNA abundance have more to do with the energetics of proof-reading than with elongation rates per se as far as abundant proteins re concerned (Holm, 1986). In B. subtilis, different profiles of iso-accepting tRNA species have been found for vegetatively growing cells, and spores (Vold, 1973).

Comparison of a data base of 16,351 and 6,121 codons for E. coli and B. subtilis respectively, indicates that whereas E. coli often has a pronounced bias towards use of one codon, B. subtilis tends to use alternative codons more evenly (Piggot & Hoch, 1985). A more detailed analysis, where homologous genes from the two organisms were compared, as well as vari-

ous categories of Bacillus genes (such as spoO genes, those for secreted enzymes, and genes expressed late in sporulation), came to a similar conclusion (Ogasawara, 1985). In B. subtilis ribosomal protein genes however, and gyrA,B, there were several codons that were not used at all, or were underused, in parallel with the counterpart genes in E. coli. From this study, it would appear that "rare" codons are unlikely to limit the translation of foreign messenger RNAs in Bacillus.

It should be mentioned that a non-AUG translation-initiating codon (viz. GUG or UUG) is significantly more frequently found in B. subtilis (and other Gram-positives) than in E. coli (29 vs 3%; Hager & Rabinowitz, 1985a).

I.7.3 Translation of Gram-Negative mRNAs in Bacillus

It has already been mentioned that B. subtilis, does not share E. coli's characteristic ability to express genes from a wide variety of sources. Transcriptional differences between the organisms were discussed. It is however, at the translational level that considerable restriction of heterologous gene expression takes place in Bacillus.

Failure to obtain phenotypic expression of genes from Gram-negative sources in Bacillus subtilis is reflected in in vitro experiments with cell-free transcription-translation systems. Thus while salt-washed ribosomes of both Gram-positives and Gram-negatives will synthesize up to 30% as much protein in the absence of salt-wash as they do when salt-wash is present on mRNAs of Gram-positive origin, Gram-positive ribosomes are extremely inefficient in translation of mRNAs from Gram-negative sources, even in the presence of salt-wash (Stallcup et al, 1976; Sharrock et al, 1979; McLaughlin et al, 1981b).

Ribosomal salt-wash from Bacillus subtilis can substitute fully for the E. coli wash in effecting translation of T7 mRNA by

E. coli ribosomes (McLaughlin et al, 1981b) and can therefore be presumed to contain similar translation-stimulating factors. Conversely, the results from heterologous systems point to the salt-wash-independent translation of Gram-positive mRNAs being due to some inherent property of these mRNAs.

E. coli S1 ribosomal protein (30S subunit) was demonstrated to improve utilization of f2 RNA by B. stearothermophilus ribosomes (Isono & Isono, 1975), but has no effect in the B. subtilis system with the same template or T7 mRNA. Addition of purified IF factors to the transcription-translation system indicates that IF3 is largely responsible for the stimulatory effect of salt-wash (McLaughlin et al, 1981b). This factor has been proposed to enhance translation by dissociating the 70S subunit (Godefroy-Colburn et al, 1975), and as this can be promoted by Bacillus mRNA, it may account in part for the organism's independence from this factor (Sharrock & Rabinowitz, 1979).

I.7.4 The Gram-Positive Shine-Dalgarno Homology

It had been observed that the degree of dependence on IFs and S1 for effective initiation at different R17 RNA translation initiation regions was inversely proportional to the complementarity of templates to the 3'-end of 16SrRNA: the greater the extent of complementarity, the more stable the 30S - mRNA complex, and the less essential were the IF factors (Steitz et al, 1977). The hypothesis was thereafter put forward by Rabinowitz & coworkers that for translation to occur in the B. subtilis system, a stronger Shine-Dalgarno complementarity was required. Examination of several Gram-positive initiation regions supported this view, and explains why poor translation of Gram-negative mRNAs occurs in Gram-positives (McLaughlin et al, 1981a; Moran et al, 1982a). Indeed, since then a large number of genes from Gram-positives have been characterized, and these invariably show impressive Shine-Dalgarno homology, while the "spacing" of RBS relative to the initiation codon is

similar in E. coli and B. subtilis. The free energies of formation of hybrids between mRNA and B. subtilis 16S rRNA average - 16.7 (+/- 2.3) kcal/mol, whereas in Gram-negatives this figure is significantly lower viz. -10.9 (+/- 3.4) kcal/mol (Hager & Rabinowitz, 1985a).

Experimental in vivo support for the strong Gram-positive SD hypothesis came from the use of a "shuttle-vector" carrying a promoterless CAT gene, which was transcribed at a low level by "read-through" from elsewhere on the plasmid. The gene retains its Tn9-derived SD sequence. This was sufficient to confer chloramphenicol resistance on E. coli, but B. subtilis transformants became Cm^R only when homologous DNA was cloned in front of the gene, thereby presumably providing an SD that could be utilized (Goldfarb et al, 1981). When a B. subtilis promoter only is positioned 5' to the gene, there is no phenotypic expression of CAT (Lin et al, 1985). Interestingly, spontaneous Cm^R mutants resulting from tandem dupli- and triplication of a portion of the gene encompassing the SD sequence have been obtained. Since the SD homology remains unchanged, it is proposed that, as the hypothetical alternative stem-loop structures that can be drawn for the mutant CAT mRNAs place the SD region in the loop, they make it more available for binding, and thereby compensate for the weak homology (Lin et al, 1985).

But it seems that more subtle forces are involved: a strong SD, even where positioned "freely" in a secondary structure prediction, is not in itself a guarantee of efficient translation in a Gram-positive system. An example of such mRNAs are the T7 late transcripts, some of which have powerful SD sequence homology. Nonetheless, only a low level of translation of the late proteins can be achieved in B. subtilis. These messenger RNAs cannot compete with Ø29-derived mRNA in an in vitro translation system (Hager & Rabinowitz, 1985b). These authors point out the extreme A-richness about the translation initiation site in Gram-positives, which is even more pronounced than it is in

Gram-negatives (Stormo et al, 1982). It is not clear though whether this constitutes a factor for ribosomal utilization.

While it is clear then that sequence alterations 5' to the SD may be influential, and that putative 2^{o} structure where SD and/or initiation codon are found in single-stranded areas can be important, but not decisive, in determining whether efficient translation can take place, it should also be borne in mind for work with heterologous genes that the coding sequence itself is also important. A given 5' untranslated region, giving good expression of one gene product, will not necessarily be equally suitable in juxtaposition with another gene, presumbly due to unfavourable secondary mRNA conformations (Stanssens et al, 1985; Warburton et al, 1983).

Finally, by supplying the appropriate promoter(s) and ribosome-binding site(s), where necessary, expression of foreign genes is made possible in B. subtilis. This is most elegantly illustrated by the transfer of the E. coli lac operator, and repressor gene to B. subtilis. The latter was brought under the control of the host regulatory signals, and allowed IPTG-inducible expression of the gene downstream from the lac operator (Yansura & Henner, 1984).

I.8: Zymogens

I.8.1 Inactive Protease Precursors

Microbial proteases and the eukaryotic digestive proteases are believed on the basis of structural similarity to have a common ancestry.

By their very nature, proteolytic enzymes, if uncontrolled, can pose a threat to the cell. In fact, most secreted proteases are synthesized in the form of an inactive zymogen precursor, which is activated by limited proteolysis. This process is as yet better understood for eukaryotes.

A classic example of autocatalytic activation is pepsinogen, which is converted to pepsin by pepsin already in the gastric lumen. Activation by specific cleavage of a peptide lying N-terminal to the active site from the inactive precursor, by a different endopeptidase, forms the basis of several physiological systems. Typically, a cascade of proteases are activated in turn. The product of one cleavage reaction, having adopted an active conformation, is the catalyst for the next. Regulatory proteases usually are multichain, with the chains covalently bound by disulphide bridges.

An amino-terminal extension is one mechanism for inhibiting protease activity. Trypsinogen, for example, is activated by the removal by enterokinase of residues 1 to 6. Trypsinogen, like some other zymogen forms, is weakly catalytic.It will react with DFP at the active site serine with an efficiency about 10^3-fold lower than trypsin. The reduced activity seems to be due to poor substrate binding ability rather than to impaired catalysis. The "activation peptide" or "pro-peptide" may be short, as in above example, or may even exceed the size of the enzyme itself.

Secreted proteases in common with other exported proteins, also have an N-terminal leader peptide or signal sequence, which is cleaved off during transport from the cytoplasm by a membrane-bound signal peptidase. This "pre-sequence" also ensures that the precursor is maintained in an inactive conformation form during compartmentalization (Neurath, 1975, 1984).

I.8.2 Bacillus Exoproteases

Cloning of the genes for the Bacillus APR and NPR revealed the presence of a sizeable coding region between sequences coding for a recognizable signal sequence and the mature secreted enzyme. Since, unlike the mammalian digestive proteases, only a single active extracellular form had been observed, this finding was greeted with surprise. It was established by changing

the 5th codon by site-specific mutagenesis to a stop codon, and comparing expression in the wt and a suppressor mutant, that this pro-region indeed is translated (Vasantha et al,1984).

When the "pro" regions of subtilisin are (re-)defined as starting C-terminal to the signal consensus cleavage site Ala-X-Ala (von Heijne, 1984a), they are 76 AA (B. licheniformis, Jacobs et al, 1985) or 77 AA (B. subtilis and amyloliquefaciens, Stahl & Ferrari, 1984; Wells et al, 1983 resp.). Remarkably, codons for charged AAs comprise 35-39% in the pro-region vs. only 10-12% for the mature enzyme. Most of the charged AA's are found in blocks of several together. There is a net positive change (+4 -->+7).

Similarly, in the "pro" region of the metalloprotease, charged amino-acids are frequent, comprising 28% of the large activation peptide of approx. 190 AAs. There is a net positive charge (Vasantha et al, 1984; Yang et al, 1984).

It has been suggested that this excess charge helps to bind the proenzyme to the oppositely charged membrane. Attention has been drawn to the fact that the peptide bond between "pro" and mature region is a preferred site for the mature APR, NPR, resp. suggesting autocatalytic activation of a membrane-bound form (Vasantha et al, 1984).

Experimental evidence supporting this hypothesis is available from studies with site-specific mutagenized subtilisin genes (Powers et al,1986). Alteration of Asp32, (in the catalytic triad) to Asn gives a non-active product. In Western blots, a mutant precursor of M_r 42,000 associated with the membrane can be detected with anti-subtilisin. In the course of incubation, this precursor disappears provided that a background level of subtilisin activity is present, and enzyme is released into the medium. In a mutant host where apr had been deleted, release of the mutant inactive "mature" enzyme (of M_r

27,000) from the membrane could be effected by the addition of exogenous subtilisin. Subtilisin thus probably undergoes autocatalytic activation.

Studies where increasing lengths of the NPR prepro-region are fused to α-amylase have suggested the presence of several processing sites internally situated in the "pro" regions, which need to be included in the fusion in order to achieve export of the α-amylase (Shimada, Mitsui, personal communication; Honjo et al, 1985). However, no forms of subtilisin other than prepro- and mature forms have been detected to date (Powers et al, 1986), although processing after the pre- region has been demonstrated in signal sequence fusions to the spa gene (Vasantha & Thompson, 1986). Such processing resulted in heterogenous N-termini.

I.9: Secretion of Proteins

The process of protein transport in both prokaryotes and eukaryotes has been the subject of intensive study in recent years, and new developments are regularly reviewed (Silhavy et al, 1983; Oliver, 1985a; Benson et al, 1985; Pugsley & Schwartz, 1985).

Only the basic features of secreted proteins and the secretory machinery will be discussed here.

I.9.1 Signal Sequence

With few exceptions, proteins that are translocated from the cytoplasm (be it to endoplasmic reticulum in eukaryotes, to periplasm or other compartment in Gram-negative bacteria, or into the medium in the case of Gram-positive organisms) are synthesized on membrane-bound polysomes (Redman et al, 1966; Randall & Hardy, 1977) in the form of a precursor which has a 15-25 AA N-terminal extension (signal peptide). The signal peptide is normally cleaved off during translocation (Perlman & Halvorson, 1983).

Signal sequences from diverse sources have certain features in

common: A) the extreme N-terminus (1-5 residues) is very basic due to the presence of several Lys and/or Arg residues, B) a block of hydrophobic residues (7-17 AAS) follows and (C) a more polar region terminating in a small amino-acid is found just before the cleavage site (Inouye & Halegoua, 1980).

Signal sequences do not otherwise display a high degree of sequence similarity, even where 2 proteins are otherwise homologous. eg.the α-amylases (Stephens et al, 1984) and subtilisins (Jacobs et al, 1985) of B. licheniformis and B. amyloliquefaciens. The signal sequence of proteins of Gram-positive origin tend to be longer than the average, and have a higher average positive charge (c.3) than do their Gram-negative counterparts (viz.2) (Watson, 1984; Abrahmsen et al, 1985).

This may reflect the fact that though Gram-positives have but one membrane,a thick cell wall must be traversed by secretory proteins.

From studies of characterized and presumed cleavage sites, general rules (the (-3, -1) rule) for the signal peptidase recognition site can be drawn: if AXB denotes the last 3 AAs of the signal sequence, where cleavage takes place after B, B can be Ala, Gly or Ser, Cys, Thr or Gln and A Ala, Gly, Ser, Leu, Val, Thr or Ile, ie the -1, -3 positions are usually taken by small, neutral AAs (except for Gln). Ala-X-Ala is most commonly found. The -2 position (X) is usually occupied by an aromatic, charged or large polar AA. Secondary structure disruptants, Gly and Pro, are commonly found at -4 and -5 resp. In eukaryotic signal sequences, Arg is found 4 times more frequently than Lys in the polar region.

The non-random AA occurrence about the cleavage site argues for evolutionary constraints that reduce the possibility of ambiguity in the form of nearby alternative sites. Certain residues are "forbidden" at certain positions: while proline

is common at -5, it is absent from -3 --> +1. Ala is rare at -2,+2 --> +5, and glycine very seldom seen at -3, though is often at -4. From work with mutant pre-proteins that are exported but not cleaved, it appears that the C-terminus of the hydrophobic core defines a "window" of approx. 7 residues that are available to peptidase (Perlman & Halvorson, 1983; von Heijne, 1983, 1984a, 1986).A weight-matrix approach has recently been described that can predict processing sites with an accuracy of c.75%, which represents a considerable improvement on the (-3,-1) rule (von Heijne, 1986).

Charge differences between eu- and prokaryotic signal sequences are possibly accounted for by assuming that an extra basic residue is required in prokaryotes to compensate for initiation with uncharged f-Met (von Heijne, 1984b).

The conserved features mentioned above are important for export: alterations to either the charged N-terminal (Inouye et al, 1982; Hall et al, 1983; Vlasuk et al, 1983) or disruption of the hydrophobic core (Emr et al, 1980; Michaelis et al,1986) interferes with transport and the coupling of synthesis and secretion. The amino acid composition of the hydrophobic core is not critical: the only requirement apparently is the length of the stretch, its hydrophobic character and the potential for alpha-helix formation (Kendall et al, 1986).

Comparison of codon usage in large number of Gram-negative genes for secreted proteins revealed a high incidence of rare codons in the regions coding for the hydrophobic core, and may possibly facilitate "coupling" by slowing translation (Burns & Beacham, 1985).

I.9.2 The Secretory Machinery

Signal sequences are to some extent interchangeable between secretory proteins and kingdoms. E. coli will for example correctly process both bacterial and eukaryotic signal sequences

in hybrid protein fusions (Talmadge et al,1980). Moreover, Signal Recognition Particle (SRP) is needed for translocation of a bacterial protein pre-beta-lactamase, across the mammalian endoplasmic reticulum, as is the case with homologous proteins (Muller et al, 1982). Thus some conservation of the translocational machinery during evolution is implied.

I.9.2.1 Secretory Machinery in Eukaryotes

Protein synthesis appears to be coupled to export by interaction of the signal sequence in the nascent chain with a M_r 54,000 component of a soluble complex of proteins and RNA (SRP) which causes a translational arrest (Walter & Blobel, 1981, 1983; Kurzchalia et al, 1986). This is only relieved by subsequent interaction with "docking protein", a component of the endoplasmic reticular membrane having a large cytoplasmic domain (Meyer et al, 1982;Lauffer et al, 1985). Translational arrest has subsequently been shown to be system-dependent, and is believed by some to be an *in vitro* artefact. The concept of obligatory cotranslation transport, with post-translational translocation rare, or invariably inefficient is no longer tenable either for pro- or eukaryotes (Schatz, 1986; Strauch et al,1986).

I.9.2.2. Secretory Machinery in E. coli

Indirect evidence for a limited number of translocational complexes in the plasma membrane comes from studies with hybrid proteins. Where these cannot be transported properly, they have a pleiotropic inhibitory effect on transport and processing of other secretory proteins, presumably by blocking all available export sites (Ito et al, 1981).

Mutations in the secretory machinery are often lethal in *E.coli*. Leader peptidase I, a component of both outer and inner membranes, is an essential enzyme (Date, 1983). (Lipoproteins are processed by a separate signal peptidase, the product of the *lspA* gene (Yamada et al, 1984)).

Genetic analysis of loci involved in secretion has primarily been achieved by exploitation of gene fusion technology. Several conditionally lethal alleles have been identified that 1) are required for excprotein synthesis and secretion, 2) can suppress signal sequence mutations or 3) suppress a conditionally lethal mutation. These categories are not mutually exclusive (Oliver 1985 a&b; Pugsley & Schwartz, 1985).
Gene products have thus far been identified for 2 wild-type components viz. PrlA/SecY and SecA.

PrlA/SecY protein is an integrated inner membrane 443 AA protein with several hydrophobic segments that span the membrane. ts variants lead to accumulation of precursors of exported proteins in the membrane at the non-permissive temperature. Other mutations are able to suppress signal sequence mutations, suggesting a direct interaction between PrlA and signal peptide (Oliver, 1985a; Akiyama & Ito, 1985). But the interpretation of the experiments on which these conclusions are based has recently become questionable, since a great multiplicity of suppressor loci is identifiable, some of which have been shown to directly affect protein synthesis rather than secretion (Strauch et al,1986; Lee & Beckwith, 1986; Shiba et al, 1986).

SecA mutations lead to pleiotropic defects in secretion of cell envelope proteins. It was believed that SecA helps to couple synthesis to secretion by recognition of especially the hydrophobic segment of a signal sequence but recently work with an amber secA mutant has cast doubt on such a role (Strauch et al, 1986). SecA has an M_r of 92,000, and is peripherally associated with the inner membrane. Many of the studies involving extragenic suppressors of secA mutations have been called into question (see above).

There may well be a system parallel to SRP and docking protein in prokaryotes to couple synthesis and translocation initiation,

as one LamB signal sequence mutation leads to decreased synthesis of the protein (Hall et al, 1983). SecA, PrlA and the as yet unidentified products of some of the other prl and sec loci implicated in secretion are likely to constitute part of such machinery (Oliver, 1986b).

Some different components are required however, for different types of proteins: secB is needed only for transport of a subset of E. coli envelope proteins (Kumamoto & Beckwith, 1985).

I.9.2.3 Secretory Machinery in B. subtilis

A M_r 64,000 protein present in cytosol, was found in a membrane fraction complexed to ribosomes, but was not present in the ribosome-free fraction. It appeared to be located between membrane and ribosome (Horiuchi et al, 1983a). This protein is also associated with free monosomes. It can be immunoprecipitated in a complex with 3 other polypeptides from ribosome-containing fractions.

A M_r 60,000 protein has been identified in S. aureus. It is similar to the B. subtilis 64kD protein immunologically, in its distribution in the cell, and can likewise be coprecipitated in a complex with other polypeptides (Adler & Arvidson, 1984).

The physiological role of the so-called secretory S complex is as yet unclear as the connection to protein transport has been based solely on location of the polypeptides in question (Horiuchi et al, 1983b; Caulfield et al, 1984; Caulfield et al, 1985).

I.10 Expression - Secretion Systems in Bacillus

I.10.1 Why Secretion of Foreign Products is Desirable

In E. coli, where experience with the expression of foreign proteins is most extensive, overproduction of heterologous eukaryotic polypeptides commonly leads to their aggregation in "inclusion bodies" in the cell (Williams et al, 1982). Recovery of the product in an active form after "renaturation" procedu-

res is not efficient. At high levels of expression of heterologous non-exported products, such inclusion bodies are seen also in B. subtilis. The accumulated product, moreover, often has a toxic effect on the cell (Brosius, 1984).

In addition, purification of product free of contaminating, toxic cell products is difficult, but is an obligation where the final product application is in the food or pharmaceutical industry.

These problems can be circumvented by achieving secretion of the heterologous product.

The Gram-positives have thus received much attention as potentially valuable expression hosts due to the ability of most species to secrete proteins in large amounts, It may therefore even be feasible to operate continuous cultures, with the exported desired products being recovered from the medium without the drawbacks mentioned above.

The secretion of foreign proteins by micro-organisms has recently been reviewed (Pugsley & Schwarz, 1985; Nicaud et al, 1986).

Although there are reports of glycosylation of some bacterial exoproteins (Bacillopeptidase F, Roitsch & Hageman, 1983; cellulases of C. fimi, T. Warren, personal communication), bacteria are not suited to the production of proteins where such correct post-translational modification is essential for activity.

I.10.2 Requirements of Expression - Secretion Systems

The successful application of any expression secretion system depends on the following factors having been adequately dealt with.

1) A stable vector must be provided

2) A functional promoter, RBS (see I.3.9 & I.7.3, I.7.4) and signal sequence (see I.9.1) must be provided.
3) the product-specific mRNA may have to be stabilized
4) The product should be recoverable either free of extraneous moieties such as a signal sequence, or "linker" peptides, or in a form from which it can be dissociated.
5) Proteolysis of sensitive products, a considerable problem in Bacillus, must be controlled (see Chapter I.2 & below). The systems that have been most widely studied in Bacillus, and with which expression and secretion of foreign proteins have been achieved are:

B. amyloliquefaciens α-amylase (Palva et al, 1981, 1982, 1983; Ohmura et al, 1984a & b; Nakamura et al, 1985; Fahnestock & Fisher, 1986); B. licheniformis penicillinase (Neugebauer et al, 1981; Mosbach et al, 1983), B. amyloliquefaciens neutral protease: pre-sequence (Yoshimura et al, 1986;Vasantha & Thompson, 1986); prepro- sequence (Honjo et al, 1985, 1986), alkaline protease: pre-sequence (Vasantha & Thompson, 1986), and the S. aureus protein A (Spa) (Lofdahl et al, 1983; Fahnestock & Fisher, 1986; Fahnestock et al, 1986).

Note that only part of the B. licheniformis penicillinase sequence was used in the studies cited . Modification of,and linkage to the membrane via Cys-27 in the signal has been problematical (Gray & Chang,1981; Hayashi & Wu,1983; Hayashi et al, 1986).

Only results that have a direct bearing on the discussion below will be mentioned.

I.10.3 Different Ways of Meeting Some of the Requirements

I.10.3.1 Vector Stability and Chromosomal Integration

Without a stable template, no reliable expression can reasonably be expected. Stability is discussed in depth below because of its relevance to the work presented in this thesis.

Plasmid functions such as par (Tucker et al, 1984; Gerdes et al, 1986), which ensure stable inheritance of the plasmid by daughter cells at cell division, are not considered here, as information about such functions on Gram-positive vectors is not yet available

At present, for most commercial applications involving plasmids, segregational stabilization of multicopy vectors in bacteria is usually achieved by antibiotic resistance selection pressure, where the resistance marker is carried on the plasmid.

The use of antibiotics for this purpose is less than desirable where a product will ultimately have a clinical application or be used in the food industry for various reasons including cost, stability of antibiotic during protracted large-scale fermentations, and difficulties in elimination of traces of it during product purification.

Novel systems employing complementation by plasmid of lesions in essential genes have been developed. For example, the gene for D,L-alanine racemase (dal) on a plasmid ensures absolute segregational stability in Dal^- strains in (complex) media lacking D-ala (B. Diderichsen, personal communication; Ferrari E. et al, 1985).

Structural stability is however more difficult to guard against in Bacillus using multicopy vectors. Especially where the plasmid supports a high level of expression of a product not required for cell survival, recE-independent deletions often occur with the elimination of expression of the gene (Kreft & Hughes, 1981; Uhlèn et al, 1981; Lopez et al,1984; Espinosa et al,1984).

High expression can lead to plasmid instability either because the overproduced gene product has a toxic effect (Brosius, 1984) or because the high level of transcription per se interferes with plasmid replication (Stassi & Lacks, 1982; Lopez et al, 1984). Read-through transcription has been shown to compromise expression of distal genes (see I.5.11, amvR2). In B. subtilis, plasmids seem to be more vunerable to overtranscription than is the corresponding clone in E. coli, where the foreign gene derives from another Gram-positive (Fahnestock et al, 1986; Sibakov & Palva, 1984), presumably due to more efficient trans-

cription in the former (the spa gene in the first reference was stably transcribed in E. coli though it was cloned without its own terminator), and possibly also to different plasmid replicative mechanisms in the two organisms (see below), or both.

Such difficulties can be overcome in several ways: a) by using a weaker promoter (B. amyloliquefaciens instead of B. lichenformis α-amylase promoter, Sibakov & Palva, 1984; B. amyloliquefaciens amylase instead of S. aureus spa promoter, Fahnestock & Fisher, 1986); b) by reducing copy number of the foreign gene B. licheniformis penicillinase on a low copy number plasmid, Imanaka et al, 1981; chromosomal integration of S. aureus blaZ (Saunders et al, 1984 a & b; or spa (Fahnestock et al, 1986); or c) by placing efficient terminators after the gene (Gentz et al, 1981).

It seems that some plasmids are inherently unstable in B. subtilis: Ehrlich & coworkers have demonstrated using duplicated vector sequences carried either extrachromosomally or integrated into the chromosome, that recombination between the homologous seguences takes place at a significantly lower frequency in the chromosome (viz. 10^{-4}) compared with a plasmid (10^{-1}; Niaudet et al, 1984). Using a "portable" ts ori it was demonstrated that the proximity of a functional ori stimulates recombination (S. Ehrlich, personal communication). It is suggested that the propensity of most Gram-positive plasmids, in contrast to their Gram-negative counterparts, to replicate via single stranded molecules provides the trigger for recombination (te Riele et al, 1986).

Direct repeats as short as 3 bp, and imperfect 11 bp homologies have been implicated in illegitimate intramolecular recombination in Bacillus (Lopez et al, 1984). Such recombination is entirely rec-independent, and is a ubiquitous phenomenon, having been described inter alia in E. coli (Albertini et al, 1982), and in higher organisms (Hogan & Faust, 1984). Evidently though, features other than short directly repeated sequences are req-

uired (Hahn & Dubnau, 1985).

These investigations have prompted the development of chromosomally integrated expression systems, where integration is directed to a locus by the incorporaton of homologous sequences carried on the plasmid (see IV.6.1, -2, -7).

Once a non-replicable plasmid has been integrated, no selective pressure is required to retain it (Janniere et al, 1985; Yang et al, 1986).

There is no evidence for DNA amplification being a spontaneous phenomenon among Bacillus species, as opposed to Streptomyces for example (Fishman & Hershberger, 1983), unless a selective pressure is applied. Under antibiotic selective pressure amplification of chromosomal (tmr, see I.5.1.2; Wilson & Morgan, 1985) and plasmid (European Patent Application 84200990.4, Gist-Brocades; Young, 1984; Janniere et al, 1985) sequences can take place, resulting in gene-dosage related expression.

Production strains of Bacillus spp. secrete enormous amounts of certain enzymes presumably from a single structural gene. Thus there is reasonable expectation of achieving high yields of heterologous gene products using chromosomal integrates in strains otherwise optimized for synthesis and export of the normal gene product of the system in use. Such optimization no doubt involves changes at several intergenic loci concerned with regulation.

The above-mentioned patent-application deals with exploitation of the B. licheniformis α-amylase producer strains and α-amylase regulatory machinery for the expression of α-amylase and eukaryotic genes.

Chromosomal integration confers structural stability on

strongly-expressed heterologous prokaryotic genes of Gram-pos-positive origin in B. subtilis where plasmid vectors could not be come established (Saunders et al, 1984a & b, Fahnestock & Fisher, 1986).

Expression levels from a single integrated copy of the spa gene (GX3357) compare very favourably with those from similarly integrated copies where the amylase promoter drives a fusion amylase-protein A gene. Increasing the strength of the comparatively weak P_{amy} by a tandem promoter duplication (GX3360) raises expression of the fusion product 10x to levels slightly above those obtained from GX3357 and accordingly renders the construction unstable on the replicative vector, pUB110.

When the intact spa gene is present in the chromosome in several copies (4 copies, GX33305), it directs a correspondingly higher level of protein A synthesis and secretion; equivalent to that achieved with the fusion gene under P_{amy} control, on pUB110. These levels are however, still 5x lower than those obtained when the amylase promoter and RBS are fused precisely to the spa gene, and the construction propagated on pUB110. In a spo0A strain, which has markedly lower APR expression than the wt, gram per liter levels of native protein A are accumulated. This high accumulation is due to the fact that under the amylase promoter control, expression continues through stationary phase whereas the protein A promoter itself is only active during exponential growth (Fahnestock & Fisher, 1986; Fahnestock et al, 1986).

I.10.3.2 3' mRNA Stabilization

It was first demonstrated in E. coli that placing a correspondingly strong (fd) terminator after a powerful promoter could stabilize plasmids (Gentz et al, 1981). These authors anticipated the potential of this approach in achieving high levels of heterologous gene expression.

Where an mRNA has a short half-life, yield of its product can be boosted by enhancing its resistance to exonucleolytic degradation by judicious use of a heterologous terminator.

This expedient has been useful in achieving higher B. licheniformis penicillinase and interleukin-2 synthesis with the aid of the crystal protein (cry) terminator both in B. subtilis and E. coli (Wang & Chang, personal communication). In similar fashion, the E. coli rrnB terminators improved chymosin yield in B. subtilis (Sven Hastrup, personal communication). It seems that the Bacillus transcriptional machinery will recognize a diversity of E. coli-derived terminators (Peschke et al, 1985).

In all these cases, the heterologous terminator was the only one present. It remains to be seen whether these improvements can be detected in cases where the foreign gene's own terminator is also present.

Unpublished data cited in Fahnestock et al (1986) mention that a strain having a single integrated copy of the S. aureus spa gene with its own terminator sequences sometimes produced as much protein A as another having 4 copies of the gene, which however lacked these sequences.

I.10.3.3 Processing of Secreted Fusion Proteins

Where the foreign protein to be produced is itself a secretory protein, it may be best, where possible, to allow it to use its own signal sequence in order to obtain correct processing, with liberation outside the cell of a product identical to the native protein. By utilizing only the regulatory elements (P, RBS), or these in addition to part of a signal sequence, combined with the entire coding region for the product and its own pre(pro)- region, synthesis and export of authentic proinsulin and protein A have been achieved (E. coli: Talmadge et al, 1980; B. subtilis; Neugebauer et al, 1981; Mosbach et al, 1983; Fahnestock & Fisher, 1986), after correct processing in the hete-

rologous expression host.

Results with gene fusions where regulatory elements and signal sequence are fused to the product coding region at the end of the signal are more equivocal: such fusions may give rise to secretion of heterogeneous products non-identical with the natural product (eg Yoshimura et al, 1986 and refs. in I.10.2).

Depending on the details of fusion gene construction, several AAs encoded by linkers which facilitate construction, or which belong to the gene donating the signal sequence, may remain as "extra" residues after leader peptidase cleavage. Provided that the product has activity this may or may not be important, but certainly in the case of potential therapeutic agents, unacceptable immunological consequences may ensue in recipients if "foreign" residues are present.

While signal sequences and cleavage sites have some conserved features (see I.9.1), the 5' ends of genes coding for the mature moiety of secretory proteins have little in common thus making strategical choices in fusion construction difficult. A systematic study of product yield in a series of constructions where the B. amyloliquefaciens α-amylase gene was fused at its codon +3 (ie after the signal cleavage site) to B. licheniformis α-amylase mature sequences at different sites from bp +1 --> +15 revealed that the small differences led to considerable variations in product yield and activity (Sibakov, 1986b).

Another strategy for liberating the authentic protein from a fusion product involves engineering a specific protease cleavage site at the junction and employing a highly selective protease. Factor Xa for example, cleaves after XYGlyArg, where the possibilities for residues X,Y are limited. Such a tetrapeptide occurs seldom in protein sequences, and has been applied to the production of Hu-β -globin in E. coli (Nagai &Thøgersen, 1984; Celltech European Patent 161937 (1985)). The experience of others indicates that factor Xa does not display absolute speci-

ficity (Nordisk Insulin, personal communication).

In much the same way, chemical cleavages can be tailored to result in cleavage of only the desired bond (see I.10.5).

I.10.3.4 The Protease Problem

When expressed in B. subtilis from pUB110, the B. amyloliquefaciens α-amylase gene directs the production of α-amylase to a level of 50% of protein synthesized by the cell, or corresponding to 1.5 gL^{-1} secreted into the medium (Palva et al, 1981, 1982;Palva,1982). When the α-amylase regulatory signals and signal sequence are used to direct synthesis of α2-interferon, only 0.5-1 ug L^{-1} were obtained (equivalent to approx. 10^9 UL^{-1}) (Palva et al, 1983). Similarly poor results have been observed with inter alia, fusions of B. licheniformis penicillinase signal and preprosinsulin (Mosbach et al, 1983), and B. amyloliquefaciens NPR signal and IFN-β (Honjo et al, 1985).

A careful study of α-amylase fusions by the Finnish group has shown that the half-life of the "eukaryotic" mRNA (for Semiliki Forest Virus glycoprotein E1) was the same as that of a prokaryotic mRNA (for E. coli β-lactamase) and that translation rates were equal, so the 10-fold and 100-fold lower yield of the β-lactamase and E1 protein resp. compared with α-amylase from the secretion vector was attributed to proteolytic degradation in the medium (Ulmanen et al, 1985). Even where a strain deficient in NPR or APR or deleted for both major exoprotease genes viz. apr and npr is used, appreciable degradation of the β-lactamase begins at the end of exponential growth (Palva, personal communication; M. Sibakov, 1986a). However, in similar experiments Nakamura & coworkers achieved accumulation of β-lactamase in the medium both in wt and apr - npr mutants (which have about 3% original exoprotease activity) by manipulation of growth conditions. Higher levels were achieved in the protease mutant (Nakamura et al, 1985). Honjo et al (1985) showed by the addition of 10 nm PMSF to the medium that their hIFN-β production was

sensitive to the low level of APR secreted by the host strain.

The yield of protein A, which would otherwise be totally degraded after To, was greatly improved by utilizing a spo0A strain, (ïS53), which was deficient in APR, but produced normal amounts of NPR (Fahnestock et al, 1986).The product nonetheless showed signs of limited proteolysis. Still better results have been obtained in a strain where the spo0A mutation was combined with deletions of apr and npr genes (Fahnestock, personal communication).

Where a product (Human Growth Hormone) is produced intracellularly by induction at high cell density (A_{450} 80), 1.5 gL^{-1} undegraded HGH was made and accumulated in refractile bodies in a spo0A strain (Ruppen, Genentech, personal communication).

In general, proteolysis is more noticable where expression levels are low. Besides the 2 major exoproteases, APR and NPR, other activities clearly compromise yields of foreign proteins attainable. Perhaps the esterase, which accounts for only a small proportion of exoprotease activity (see I.2.2.8), or other enzymes, yet to be discovered, are responsible for product degradation. Esterase levels are impaired concomitantly with APR in spo0A mutants . Should it not be possible to overcome the protease problem satisfactorily either by identification and deletion of the protease genes, or by manipulation of growth conditions, it might be necessary to turn to Bacillus spp., which while not as well characterized genetically as B. subtilis, are transformable, and able to secrete proteins, but lack appreciable exoprotease activity eg. B. brevis. (Kurotsu et al, 1982; Takahashi et al, 1983).

I.10.4 Optimization

I.10.4.1 Culture Medium and Growth Conditions

Optimization of protein production from apr and npr signal fusions revealed that variation from 10 mg to 3gL^{-1} accumulated, dependent on culture medium and conditions. Unexpectedly, a complex medium containing glucose, which is not optimal for NPR

production, gave highest yields of processed protein A exported with the aid of the NPR signal (Vasantha & Thompson, 1986; (Vasantha, personal communication). Growth rate limitation, achieved by addition of succinate to the medium and providing only poor aeration, results in high levels of E. coli β-lactamase which accumulate over 48h in the medium even with a wild-type strain, whereas total degradation of the product otherwise occurred between 12 to 24 h under more favourable growth conditions (Nakamura et al, 1985). The improvement due to manipulation of culture conditions was estimated to be 50-60-fold when using an NPR^- APR^- mutant. Transferring the cells to fresh medium has also proved useful (Honjo et al, 1986).

I.10.4 Use of Intergenic Loci that Enhance Secretion

Where expression is already high eg. B. amyloliquefaciens α-amylase on multicopy pUB110 in B. subtilis, there was no advantage in having an amyS and pap-9 background, (ie. having the mutant alleles present in one copy per chromosome) rather against expectations (Sibakov et al, 1983). The effect of multiple copies of some cloned regulatory genes on expression has been discussed (see I.6.2 and I.6.4).

I.10.5 E. coli as Host for Expression-Secretion Vectors

E. coli is the best characterized micro-organism available. Due also to its ability to express genes from a wide variety of organisms, it is often the system of choice for initial cloning and expression regardless of the source of the gene in question. A multiplicity of effective expression systems is available which can give high yields of product.

It should be noted that although the Gram-negative bacteria are not noted for their ability to secrete proteins into the medium, many, including E. coli K12 are able to do so (reveiwed in Pugsley & Schwartz, 1985). This ability extends in some cases also to foreign proteins. "Leaky" (Lky) mutants can be isolated after mutagenesis that leak both periplasmic and heterologous secretory proteins into the medium (Gilkes et al, 1984; reviewed in Ni-

caud et al, 1986).

E. coli HB101 will "secrete" cloned penicillinase from an alkalophilic Bacillus species into the medium . It accumulates to higher levels than are obtained when the gene is expressed in B. subtilis, attributable mainly to the absence of significant proteolytic activity in the E. coli medium. Appreciable loss of periplasmic alkaline phosphatase however occurs. It has been established that in these studies, the integrity of the outer membrane was affected by a low level of transcription through the pMB9 kil gene (Kato et al, 1983, 1984; Kudo et al, 1983; Kobayashi et al, 1986).

With such "passive secretion" from the periplasm, where expression is high enough, the periplasmic contaminants do not constitute a significant contaminating factor, especially where a strong specific purification is possible. Such results have been achieved with the protein A system of M. Uhlèn originally intended for staphylococcal hosts. The signal peptide is correctly processed during translocation. The remaining protein A moiety of the fusion product forms the basis for an affinity chromotographic purification and the desired product released by a specific chemical cleavage at the engineered splitting site. (Abrahmsen et al, 1985; Nilsson et al, 1985).

With a fusion of OmpF, an outer membrane protein, and human β -endorphin, bona fide selective transport was achieved without causing leakage of periplasmic markers. The OmpF signal peptide was correctly processed, and the hybrid peptide released into the culture medium, where it accumulated (Nagahari et al, 1985). Similar results have been achieved with the S. marcescens serine protease. In this case moreover, a large C-terminal polypeptide was correctly cleaved to yield the mature enzyme (Yanagida et al, 1986).

As a consequence of these recent results, there is a growing consensus of opinion that E. coli or other Gram-negative orga-

nisms possibly represent a more promising system for achieving high yields of exported heterologous gene products than does B. subtilis.

I.11 Reasons for Undertaking this Project

As already mentioned in I.1, the cloning of subtilisin genes was undertaken for several reasons. The most immediate of these was the planned use of transcriptional and translational signals, as well as the genes' signal sequence, in order to direct synthesis and secretion of foreign gene products in Bacillus host strains.

As is described in Chapters III and IV, this initial goal was not realized by the author. Due both to difficulties in isolating the subtilisin gene from one organism, and to concurrent developments in other laboratories, the emphasis of the work was changed from application/exploitation of the isolated sequences, to that of characterization, and investigation of regulation of the gene's expression. Associated with the original intention to use Bacillus as a host for heterologous gene expression, the cloned gene from B. subtilis was utilized to create a chromosomal deletion, resulting in a strain that produces no subtilisin. As many proteins are subtilisin-sensitive, such a deletion mutant presents an obvious improvement as host for the expression of secreted foreign gene products.

Chapter II: Materials & Methods

II.1.1 Enzymes, Chemicals & Biochemicals

Restriction enzymes, T4 DNA ligase, polynucleotide kinase, DNA polymerase I and Klenow fragment were obtained from New England Biolabs, unless otherwise stated. Asp718 was supplied by Boehringer Mannheim. All enzymes were used in accordance with the supplier's instructions. Chemicals and biochemicals were obtained from Merck and Sigma resp., unless otherwise stated. Radiochemicals were supplied by New England Nuclear. Oligomers were synthesized in the organic chemistry laboratories of either Kjeld Norris (Novo Research Institute, Denmark) or Les Bell (at the former Searle Research Laboratories, High Wycombe).

II.1.2 Computer Software

For VAX and Hewlett-Packard mini-computer resp., programs by R. Staden and M. Jonson were used, unless otherwise stated.

II.1.3 Strains & Plasmids

Strains and plasmids used in this study are listed in Tables 4 and 5 resp.

II.2 General Media and Buffers

Most components supplied by Difco.

L Broth (LB): Bactotryptone 10 gL^{-1}, Yeast extract 5 gL^{-1}, NaCl 10 gL^{-1}

L Agar: As above, with 15 gL^{-} Bacto-Agar.

Modified Schaeffers Medium: (Leighton & Doi, 1971)

Nutrient Broth (Difco) 16gL^{-1}, $MgSO_4.7H_2O$ 0.5gL^{-1}, KCl 2gL^{-1}

After autoclaving, the following additions were made from sterile stocks, just prior to use.

$Ca(NO_3)_2$	10 ml	of 0.1M stock	(final conc.	10^{-3} M)
$MnCl_2$	1 -	- -	(	10^{-4} M)
$FeSO_4$	10 µl	- -	(	10^{-6} M)
Glucose	5.0 ml	20% stock	(	0.1%)

Tryptone Agar (1L): Tryptone 10 g, NaCl 2.5g, 15 g agar

An industrial medium, BP-X, Bacto-Penassay Broth and Base Agar (Difco), and Tryptone Blood Agar Base (Difco) were also used.

TABLE 4: LIST OF BACTERIAL STRAINS USED IN THIS STUDY

B. subtilis Strains

STRAIN	RELEVANT PHENO-/GENOTYPE	SOURCE/REF.
MFJ93	=DN497=DN223XDN234 . Produces APR of BPN' type DN223=RUB200; DN234=trpC2,Spo$^-$,Prt$^{+/-}$	B. Diderichson
MFJ135	=MFJ93(pMJ6). Overproduces BPN' from cloned MFJ93 gene.	this thesis
MFJ157	=MFJ93Δapr Cm^{6R}, ex MFJ93::pMJ13 Secretes no detectable BPN'	this thesis
MFJ158	=MFJ93::pMJ14 (Cm^R,APR$^+$)	this thesis
MFJ163	=MFJ93::pMJ20 (Cm^R,APR$^+$). Plasmid integrated 5' to CAT	this thesis
MFJ164	=MFJ93::pMJ20 (Cm^R,APR$^+$). Plasmid integrated 3' to CAT	this thesis
B. subtilis 168 wt	=BGSC 1A1. trpC2	BGSC
MFJ195	=MFJ93(pMJ28) Cm^R. Secretes BPN' from endogenous subtilisin gene	this thesis
MFJ196	=MFJ93(p820) Cm^R . Secretes reduced amount BPN'	this thesis
MFJ198	=MFJ157(pMJ28). Cm^R. Secretes no detectable subtilisin	this thesis
RUB200	aroI906,amyE07,amyR2	Yoneda et al (1979)
MFJ205	=MFJ210(pMJ29). Cm^R. Secretes subtilisin Carlsberg in stationary phase in presence of xylose, from cloned gene ,XynB promoter	this thesis
MFJ210	=SHa28. metB$^-$Xyl$^-$ MFJ93 derivative. Protease production as MFJ93	S. Hastrup
MFJ213	=MFJ210(pMJ32). Cm^{R}Km^R. Secretes subtilisin Carlsberg in stationary phase from cloned gene. α-amylase promoter	this thesis
MFJ219	=MFJ93(pMJ33). Cm^R.	this thesis
MFJ221	=MFJ93(pMJ42). Cm^{R}Km^R.	this thesis
MFJ231	=168wt X MFJ157. Cm^{6R} Δapr	this thesis
MFJ239	=MFJ210(pMJ50). Cm^{6R} Xyl$^-$	this thesis
MFJ240	=MFJ210(pMJ50/P$_{amy}$). Cm^{6R} Xyl$^+$	this thesis
MFJ246	=MFJ210(pMJ51). Cm^R. Secretes subtilisin Carlsberg from cloned gene,own promoter	this thesis
MFJ248	=MFJ157(p820). Cm^R. Secretes no detectable subtilisin	this thesis
MFJ249	=MFJ198	this thesis
MFJ250	=MFJ157(pMJ51). Cm^R. Secretes subtilisin Carlsberg from cloned gene,own promoter	this thesis

TABLE 4 (continued)

E. coli Strains

STRAIN	RELEVANT GENO-/PHENOTYPE	SOURCE/REF.
GM33	dam33	N. Fiil
MFJ55	=WT264=MC1061recA56Tn10 (Tet^R)	W. Tacon
MFJ92	=NF3081=MC1000dam33Rec$^-$	N. Fiil
HB101	hsdS20(r_B^- m_B^-)rec13,ara-14,proA2,lacY1,galK2, rpsL20 (Str^R),xyl-5,mtl-1,supE44,λ^-	Maniatis et al (1982)
JM101	R^+M^+ Δ(lac,pro),supE,thi,F'traD36,proAB,$lacI^Q$ ZΔM15	Yanisch-Perron et al (1985)
M182	Δ(lacIPOZY)X74,galK,galU,strA	Casadaban & Cohen (1980)
MFJ233	MFJ55 (Tet^S)	this thesis
MC1000	araD139,Δ(ara,leu)7697,ΔlacX74,galU,galK,strA	Casadaban & Cohen (1980)
MC1061	araD139,Δ(ara,leu)7697,ΔlacX74,galU$^-$,galK$^-$,strA, hsm$^+$,hsr$^-$	Casadaban & Cohen (1980)

TABLE 4 (continued)

B.licheniformis strains

STRAIN	Alternative name, Relevant geno-/phenotype	SOURCE/REF.
MFJ 29	= NCIB6816. Produces subtilisin of the Carlsberg type.	NCIB, this thesis
BGSC5A4	M28-RT-M20-RT-T10=FD03. pep-1 pur-1 Tet^{5R} Emy^{5R} Cm^{5R} Ap^{R} Km^{R}	BGSC, this thesis, Thorne & Stull (1966)
BGSC5A5	=FD6.pep-1 pur-1 leu-2 Tet^{5R} Emy^{5R} Cm^{6R} Ap^{R} Km^{6R}	BGSC, this thesis
BGSC5A13	=M8R-1. pep-1 met-4 Tet^{5R} Emy^{5R} Cm^{6R} Ap^{R} Km^{20R}	BGSC, this thesis, Thorne & Stull (1966)
BGSC5A15	=M18. pep-1 arg Tet^{5R} Emy^{S} Cm^{S} Ap^{S} Km^{6R}	BGSC, this thesis, Leonard & Mattheis (1965)
BGSC5A16	=M28. pep-1 gly-3 Tet^{5R} Emy^{5R} Cm^{6R} Ap^{R} Km^{6R}	BGSC, this thesis, Thorne & Stull (1966)
BGSC5A20	=749.prototroph. Tet^{5R} Emy^{5R} Cm^{15R} Ap^{R} Km^{S}	BGSC, this thesis, Sherratt & Collins (1973)
BGSC5A21	=749/C. as 5A20, but constitutive beta-lactamase	as for 5A20
BGSC5A23	=749/C-72. pen-72 (beta-lac^{-}) CRM^{+} Tet^{5R} Emy^{5R} Cm^{6R} Ap^{S} Km^{6R}	as for 5A20
BGSC5A24	=11946. prototroph. Tet^{5R} Emy^{5R} Cm^{6R} Ap^{R} Km^{20R}	BGSC, this thesis
AL20222	Tet^{5R} Emy^{20R} Cm^{15R} Ap^{R} Km^{S} prototroph	R. Rogers, Apothekernes Laboratorium

TABLE 5: LIST OF PLASMIDS USED IN THIS STUDY

NAME	Phenotype conferred	SOURCE/REF.
GRAM-NEGATIVE PLASMIDS		
pBR322	Ap^R, Tet^R	Sutcliffe (1979), Peden (1983)
pBR322 DERIVATIVES		
pAT153	Ap^R, Tet^R	Twigg & Sherratt (1980)
pMJ1	Ap^R. =pAT153 & 3' end of Carlsberg gene from codon +55	this thesis, Fig. 8
pMJ2	Ap^R. =pUR222 & 5' end of Carlsberg gene	this thesis, Fig. 10
pMJ3	Ap^R. pBR322 deriv.& BclI cloning site followed by stop codons	this thesis, Fig. 9
pMJ5	Ap^R. = deletion derivative of pBR322	this thesis, Fig. 36
pMJ7	Ap^R. =pBR322 & C-terminal region of Carlsberg gene with BclI cloning site	this thesis, Fig. 9
pMJ8	=pMJ5 & B. subtilis subtilisin gene	this thesis, Fig. 36
pMJ13	pMJ8 deriv.Part of subtilisin BPN' gene substituted with CAT gene. $Ap^R Cm^{6R}$	this thesis, Fig. 36
pMJ14	as pMJ13,but different segment of BPN' gene exchanged for CAT gene	this thesis, Fig. 36
pMJ15	Ap^R. pN47.11 with HindIII deletion in BPN' gene	this thesis, Fig. 36
pMJ16	$Ap^R Cm^{6R}$. =pMJ15 derivative carrying CAT gene	this thesis, Fig. 36
pMJ19	Ap^R. =pMJ3 derivative & 5S RNA gene and terminators	this thesis, Fig. 9
pMJ20	$Ap^R Cm^{6R}$. =pMJ16 derivative , CAT gene beside partially deleted BPN' gene	this thesis, Fig. 36
pMJ22	$Ap^R Cm^{6R}$. =plasmid resulting from excision of pMJ14 from its site of integration at BPN' locus; carries additional 5'sequences	this thesis, Fig. 41
pMJ24	Ap^R. =pUC18 derivative & 5' part of Carlsberg gene	this thesis, Fig. 16
pMJ26	Ap^R. =pUC18 derivative & Carlsberg gene; includes transcriptional start & terminator (5'-StuI to HaeIII-3')	this thesis, Fig. 16
pMJ31A	Ap^R. =pUC18 derivative & Carlsberg gene & its promoter (5'-AvaI to HaeIII-3')	this thesis, Fig. 24
pMJ50	$Ap^R Cm^{6R}$. Single copy Bacillus promoter cloning vector	this thesis, Fig. 28
pN47.11	Ap^R. =pBR322 & BPN' gene	this thesis, Fig. 31
pUC18	Ap^R.Polylinker cloning sites in lacz	Yanisch-Perron et al (1985)
pUR222	Ap^R.Polylinker cloning sites in lacz	Ruther et al (1981)

TABLE 5 (continued)

NAME	Phenotype conferred	SOURCE/REF.
GRAM-POSITIVE REPLICATIVE PLASMIDS		
p820	Cm^R. =pBD64 derivative	B. Diderichson, Fig. 17
p900	Cm^R. =low cop analogue of p820	B. Diderichson, Fig. 17
p1050	Cm^R. =pBD64 derivative	B. Diderichson, Fig. 34
pBD64	$Cm^R Km^R$, pUB110 ori	Gryczan et al (1980a)
pBF827	Km^R. =pUB110 & Carlsberg fragment from λs6	J-I. Flock, Fig. 13
pC194	Cm^R	Horinouchi & Weisblum (1982) Dagert et al (1984)
pMJ6	Cm^R. =p1050 & subtilisin BPN' gene	this thesis, Fig. 34
pMJ27	Cm^R. =p900 & subtilisin Carlsberg gene, including transcriptional start & stop sites (5'-StuI to HaeIII-3')	this thesis, Fig. 17
pMJ28	as pMJ27, but is p820 derivative. No detectable expression of Carlsberg enzyme.	this thesis, Fig. 17
pMJ29	Cm^R. pSX50 derivative & Carlsberg gene (5'-StuI to HaeIII -3'), preceded by xylose-inducible promoter	this thesis, Fig. 18
pMJ32	$Cm^R Km^R$. pBD64 derivative carrying Carlsberg gene (as in pMJ27), preceded by P_{amy}	this thesis, Fig. 18
pMJ33	$Cm^R Km^R$. Carries promoterless xynB gene, preceded by sequences upstream from Carlsberg gene on pMJ2. No detectable xylosidase activity	this thesis, Fig. 23
pMJ42	$Cm^R Km^R$. Carries Carlsberg gene (5'-StuI to HaeIII-3'), preceded by same sequences as in pMJ33, with linker of 30 bp between. No detectable Carlsberg synthesis.	this thesis, Fig. 23
pMJ51	Cm^R. =p820 & Carlsberg gene & its own promoter (5'-AvaI to HaeIII-3'). Supports Carlsberg synthesis	this thesis, Fig. 24
pPX7	$Cm^R Km^R$. pBD64 derivative & promoterless B. pumilus xynB gene	S. Hastrup, Fig. 23
pSX50	Cm^R. pBD64 derivative & B.subtilis xynB OP region in front of B.pumilus xynB gene	S. Hastrup, Fig. 18
pUB110	Km^R	Gryczan et al (1978a), McKenzie et al (1986)
GRAM-POSITIVE INTEGRATIVE PLASMIDS		
pHV60	≃ pHV32 . $_R Cm^R$. pBR322 ori. In Gram-negatives, Ap^R & Tet^R also expressed.	Niaudet et al (1982)
pMJ13 pMJ14 pMJ16 pMJ20 pMJ50	see list of pBR322 derivatives	this thesis

TE: 10 mM TRIS 1mM EDTA pH 7.5

II.3 DNA Preparations

II.3.1 Preparation of Chromosomal DNA from B. licheniformis, B. subtilis & E. coli

II.3.1.1 Miniscale

Cells from 5 ml of an L Broth overnight culture were pelleted. The cells were resuspended in 250 µl 25% sucrose-100mM NaCl-50 mM TRIS-HCl (pH 7.5), spun to rinse, and resuspended in 100 µl of the sucrose buffer.

100 µl of a 25 $mgml^{-1}$ lysozyme suspension (in sucrose buffer), and 150 µl EDTA pH8, were added, and the suspension incubated for 30 mins. at 37^0C with agitation.

120 µl 1M TRIS-HCl pH 7.5, 10 µl 5M NaCl and 50 µl 10% SDS were added, and the suspension incubated at 60^oC for 5 mins, after which it was extracted several times with 530 µl Phenol/$CHCl_3$ (1:1). Once the inter-phase was clear, the nucleic acids were precipitated on ice with 2 volumes cold EtOH. After incubation on ice for 5 mins. the nucleic acids were pelleted by centrifugation in an Eppendorf centrifuge for 2 mins, and the supernatant removed carefully with a pipette.

The pellet was rinsed with cold 70% EtOH, and recentrifuged as above, after which it was resuspended in 100 µl 50 mM TRIS-HCl (pH 7.5) - EDTA 0.5 mM at 4^oC overnight.

The resuspended nucleic acids were incubated with 2 µl of 1 $mgml^{-1}$ RNAase A solution, that had been heat-treated (viz. incubated at 80^{oC} for 1h), for 60 mins. at 37^oC, after which they were again phenol extracted and pelleted as described above. Finally the DNA was resuspended in 50 µl TE. 5µl of the suspension was normally sufficient for 1 track in a Southern blotting experiment.

II.3.1.2 Large scale (only differences from above noted).

The method described above was scaled up by a factor of 10.

The incubation at 60°C was for 10 mins., and DNA was spooled out rather than spun down after ethanol precipitations. For the RNAase digestion, the DNA threads were resupended immediately in 2 mls TE, and 0.5 mg RNAase A added. Incubation was for 30 mins. at 37°C, whereupon 0.1 mg proteinase K (Boehringer) was added, and incubation continued for 1 h. NaCl was added to 0.5 M, and the suspension extracted twice with an equal volume of phenol/$CHCl_3$, and twice with $CHCl_3$. After spooling out from ethanol, and rinsing in 70% ethanol for the last time, the DNA was resuspended in 2 mls TE as described above.

II.3.2 Plasmid Preparation

Plasmids from either E. coli or B. subtilis were extracted by a modification of the method described by Kieser (1984).

Solution 1: 0.3 M sucrose, 25 mM TRIS pH8, 25 mM EDTA pH8. Lysozyme was added to 2 $mgml^{-1}$ prior to use.

Solution 2: 0.3 M NaOH, 2% SDS
Made fresh once a month, stored at 4°C, and warmed prior to use to redissolve SDS.

Solution 3: phenol 100 g, chloroform 100 ml, H_2O 20 ml.

Unbuffered sodium acetate (3 M)

Neutral phenol/chloroform: A 1:1 v/v solution made with phenol, saturated and equilibrated first with 1M TRIS pH8, then with 0.1 M TRIS pH8.

10 x TE: 100mM TRIS pH8-10mM EDTA pH8

RNAase A: 1 $mgml^{-1}$ stock heated at 80°C for 60 mins. Aliquots were stored at -20°C.

II.3.2.1 Miniscale

Cells from 1 to 5 mls of an overnight culture were pelleted by spinning briefly in an Eppendorf centrifuge, or in a Sorvall SS34 rotor for 5 mins. at 5000 rpm. 500 µl of solution 1 was added to the cell pellet, followed by vortexing to resuspend the cells, which were then incubated at 37°C for 15 mins. 250 µl solution 2 was added, the suspension mixed by inversion, and incubated at 70°C for 30 mins. The sample was then cooled to room temperature. 100 µl of solution 3 was added, mixed in gently, and the sample centrifuged for 5 mins. as above. The supernatant was aspirated to a fresh tube. 70 µl 3M sodium acetate and 700 µl isopropanol were added, and the sample incubated briefly at -80°C. Nucleic acids were pelleted by spinning for 2 mins. in an Eppendorf centrifuge, and the supernatant discarded.

The pellet was resuspended in 100 µl 10XTE/0.3M sodium acetate, and extracted once with an equal volume of neutral phenol/$CHCl_3$. The suspension was spun for 2 mins., and the aqueous phase transferred to a fresh tube, where it was extracted once with an equal volume of chloroform. The plasmid was precipitated finally with 2 volumes of ethanol. After chilling for 5 mins. at -80°C , or 30 mins. at -20°C, the plasmid was spun down in an Eppendorf centrifuge, and the supernatant aspirated. The pellet was rinsed without resuspension with 70% EtOH, and chilled at -20°C. The pellet was respun briefly (2 mins.) as above, and the supernatant removed carefully. The pellet was then lyophilized, and resuspended in 50 µl TE. 1 µl of plasmid was usually sufficient for a slot on a minigel.

II.3.2.2 For Low Copy Plasmids

Some modifications were necessary for plasmids present in low copy number. Only differences from the above method are noted.

10 mls of an overnight culture were used, and respectively 1, 0.5 and 0.4 ml of solutions 1,2 and 3 added. For the first precipitation 140 µl 3M Na-acetate and 1.4 ml isopropanol were added. After the lyophilization, an extra step was required, viz. the DNA was resuspended in 400 µl TE, and 35 µl 0.5M EDTA pH8 and 20 µl RNAase A were added. After incubation at 37°C for 30 mins, the DNA was phenol extracted, precipitated and dried. After resuspension in 50 µl TE, 10 µl were normally required for one application on a minigel

II.3.2.3 Large Scale

Solutions as for mini-prep, but in addition:

n-butanol saturated with H_2O and CsCl

CsCl/TE Stock: 100 g CsCl, 100 ml TE, 3 ml EtBr 10 $mgml^{-1}$.

Starting from 100 mls of an overnight cell culture, method II.3.2.1 was followed, using 10-fold greater volumes of reagents, as far as the isopropanol precipitation. The resulting pellet was resuspended in 400 µl 10XTE/o.3M sodium acetate, and extracted once with the neutral phenol/$CHCl_3$ solution, and once with $CHCl_3$. The nucleic acids were then precipitated with 2 volumes of EtOH, chilled at -80°C for 5 mins., pelleted by spinning at 10,000 rpm for 20 mins., and dried.

The DNA was resuspended in 500 µl TE. 40 µl 0.5M EDTA pH8 was added, and the sample incubated at 37°C for 30-60 mins. with 25 µl RNAase A. 0.5 g CsCl and 15 µl EtBr (10 $mgml^{-1}$ stock) were added, and the mixture transferred to VTi65 tubes. These were filled up to the neck with CsCl solution, and the tubes were sealed. After centrifuging in a VTi65 rotor at 45,000 rpm and 15°C for 18-20 h, the rotor was decellerated without braking. The lower plasmid band was "pulled" with a syringe, under UV illumination at 366nm. EtBr was extracted with the n-butanol solution. The colourless plasmid suspension was diluted with 3 volumes of water, and precipitated with 2.5 vols. EtOH in Corex tubes. After chilling at -80°C for 5 mins., the DNA was pelleted by centrifugation at 10,000 rpm for 30 mins. The supernatant was discarded, and the pellet dried. Plasmid DNA was resuspended in 400 µl TE, transferred to an Eppendorf tube, and reprecipitated at an NaCl concentration of 0.2 M with 2.5 vols. EtOH. After chilling at -80°C, and spinning down in an Eppendorf centrifuge for 5 mins., the pellet was rinsed with cold 70% EtOH, respun for 2 mins. and dried. The plasmid was resuspended in 200 µl TE. 1 µl was usually sufficient for a slot on a minigel.

Plasmids for cloning of chromosomal fragments were always banded on CsCl gradients.

II.3.3 Lambda DNA Preparation

Lambda DIL: 0.01% glycerol, 0.01M Tris-HCl pH 7.4, 0.01M $MgSO_4$.

NZY (1 L): 10 g NZamine, 5g yeast extract, 5g NaCl, 2g $MgCl_2.6H_2O$.

Lambda Top (1 L)

10 g tryptone, 2.5 g NaCl, 6.5 g agar.

This was dispensed into 50 ml portions. Just prior to use, a portion was melted and cooled to 45°C. A portion was never remelted.

TMB (1 L) (or L Broth containing maltose 0.4% and $MgCl_2$ 10mM):

10 g tryptone, 2.5 g NaCl, 2 g maltose.

Indicator cells

Recipient strain grown overnight in TMB .

II.3.3.1 Titrating Stock: Rapid One-Plate Method

A fresh tryptone plate was overlaid with 100 µl indicator cells in 2.5 ml lambda top (c. 45°C). The plate was dried at 37°C for 30 mins. with the lid off. A fresh plaque was picked into 200 µl lambda DIL, and left at room temperature for 30 mins. to elute the 'phage; when available, a 'phage stock was used instead. 10-fold dilutions of 'phage in lambda DIL were made in the range 10^{-1} to 10^{-10}, and 10 µl of each dilution was spotted on to the plate overlaid with indicator cells. Twenty mins. was allowed for adsorption, before the plates were transferred to 37°C for incubation overnight. Next day, plaques resulting from a dilution giving rise to a convenient number were counted, and the titer estimated.

II.3.3.2 Preparation of Plate Lysates of Lambda

100 µl indicator cells were preadsorbed with 10^5 pfu (or 5% of an eluted plaque) for 20 mins. at 37°C. Then 3 mls molten lambda top at c. 42°C was added, and the mixture spread on a fresh tryptone plate. The plate was inverted and incubated at 37°C for 8 h, or until lysis was confluent. 5 mls of lambda DIL were added to the plate, and it was left at 4°C for several hours with occasional shaking to elute the 'phage. The suspension was sucked up, and the plate rinsed with about 1 ml of lambda DIL, which was pooled with the first suspension. 100 µl $CHCl_3$ was added, and the harvested 'phage suspension vortexed briefly, after which it was centrifuged at 4,000 rpm and 4°C for 10 mins. A drop of $CHCl_3$ was added to the supernatant and it was stored at 4°C.

II.3.3.3 Large Scale Preparation of Lambda DNA (B. Old)

Day 0 The 'phage stock was titrated, and a fresh indicator culture was set up.

Day 1 1 ml fresh indicator, 1 ml 10 mM $MgCl_2$, and 100 µl lambda DIL containing 10^7, or 10^8 pfu were preadsorbed for 10 mins at 37°C (in triplicate for each of the two infection ratios). After preadsorption, each sample of cells was added to a 1L baffle-flask containing 300 mls fresh L Broth or NZY, which had been prewarmed to 37°C, and the flasks were incubated at 37°C for 16 h with slow shaking (120 rev min^{-1}).

Day 2 500 µl $CHCl_3$ was added to each flask, and the flasks were swirled gently. Their contents were centrifuged at 10,000 rpm for 10 mins. Afterwards the appropriate supernatants were pooled, and 1 ml of both pools set aside for titration. 40g NaCl per liter supernatant was added, and the suspension swirled to dissolve the salt. 111g PEG 6000 per liter was added, and the suspension stirred slowly at room temperature using a magnetic stirrer for 1 h. The suspensions were kept at 4°C overnight, while their titers were estimated as described in II.3.3.1.

Day 3 Suspensions having a titer below 10^{10}pfu/ml were discarded. Those of $\geq 10^{10}$pfu/ml were stirred slowly as above for 1.5 h to resuspend the particles, after which they were centrifuged at 10,000 rpm for 10 mins. The supernatant was discarded, and the position of the pellet marked on each tube. 5 mls of lambda DIL were added to each pellet, and the tubes rocked on a shaker for several hours, with regular rotation of the tubes by a few degrees until the entire pellet was resuspended. Suspensions were pooled, and incubated at 37°C for 1 h with 10 $\mu g ml^{-1}$ DNAase I to degrade chromosomal DNA, and with 1 mg RNAase A per 25 ml. They were then kept at 4°C overnight.

Day 4 The suspensions were given a clearing spin at 10 K rpm for 10 mins. The supernatants were then centrifuged in a swingout rotor at 40K rpm for 1 h, resulting in a yellow gelatinous pellet. This was resuspended in 500 µl lambda DIL by standing the tubes at 4°C for several hours. Finally, the suspensions were heated at 65°C for 10 mins. to inactivate DNAase I, and then cooled on ice. 30 µl of a solution of Pronase (12.5 $mg ml^{-1}$) was added to each ml of lambda, and the 'phage incubated at 37°C for 1 h. SDS was added to a final concentration of 1%, Na-acetate to 0.4M, and the suspension heated at 65°C for 10 mins. After this treatment, the suspension was phenol extracted,without vortexing, very gently several times until a clear interphase was achieved. After a single $CHCl_3$ extraction, 2.5 vols. ice-cold EtOH were added, and the lambda DNA spooled out. It was rinsed with 70% EtOH, and resuspended in 250 µl TE at 4°C overnight.

1L 'phage suspension usually gave about 2 mg DNA.

II.4 Transformation Procedures

II.4.1 Competent Cell Preparation for B. subtilis was performed by a modification of the method of Boylan R. J. et al (1972)

10 x MM (1L): $(NH_4)_2SO_4$ 20 g, KH_2PO_4 60 g, $K_2HPO_4.3H_2O$ 140 g.

Salt mix (20 mls)

0.40 ml 0.5M stock $CaCl_2.2H_2O$

0.20 ml 0.1M stock★ $FeCl_3.6H_2O$

o.20 ml 0.1M stock $MnCl_2.4H_2O$

★$FeCl_3$ stock was filtered, not autoclaved, and was stored at 4°C.

87% glycerol(sterile)

KM stock

Sterile solutions were mixed as follows:

900 ml sterile H_2O

100 ml 10xMM

10 ml 10% tri-sodium citrate (dihydrate)

2 ml 1M $MgSO_4.7H_2O$

KM1 & KM2

Sterile solutions were mixed as follows:

	KM1	KM2
KM stock	480 ml	480 ml
20% glucose	10 ml	10 ml
20% casaminoacids	500 µl	500 µl
5% yeast extract	10 mls	10 mls
0.1 M $MnCl_2$	15 µl	-
Salt mix	-	0.5 ml
0.5 M $CaCl_2$	-	0.5 ml
1 M $MgCl_2$	-	0.4 ml

Procedure

A culture was grown in L Broth for about 5 h. The following serial dilutions were made in KM1 and grown overnight in 15 ml tubes with shaking: 4 serial 25-fold dilutions, where 0.2 ml was added to 5 ml KM1, followed by 6 serial 50-fold dilutions, where 0.1 ml was added to 5 mls KM1.

An SS34 rotor, sterile oakridge tubes, KM2 and 87% glycerol were warmed to 37° C, and a centrifuge to 20° C. 2.5 ml of the highest cell dilution showing growth after the overnight incubation was diluted in 22.5 ml prewarmed KM2, and the incubation was continued for 1 h in a flask, with agitation. The cells were harvested by centrifugation at 7,000 rpm for 3 mins. in the prewarmed tubes. The supernatant was reserved under sterile conditions. 2.2 mls of it was used to resuspend the pellet, and 0.6 ml 87% glycerol was added and gently mixed in. 400 µl aliquots

were frozen in liquid nitrogen in cryotubes, and stored at -80° C until needed. The transformability of the cells was checked by transformation with a standard plasmid (here pHV33). The transformation frequency was typically 10^5 μg^{-1}, but occasionally was as high as 10^6 μg^{-1}.

II.4.2 Transformation of B. subtilis Competent Cells

BC medium

900 ml H_2O
100 ml 10 x MM (see II.4.1)
10 ml 10% Sodium citrate
2 ml 1M $MgSO_4$
1 ml salt mix (see II.4.1)

BTF (freshly made just prior to use)

350 µl BC medium, 50 µl 0.4% glucose, 50 µl 40 mM $MgCl_2$, 50 µl 2 mM EGTA.

Procedure

A portion of frozen competent cells was thawed quickly at 37° C, and one volume BTF was added. Up to one µg DNA was incubated with 100 µl of cell/BTF mix, with vigorous agitation, at 37° C for 20 mins. in tubes that had been prewarmed to 37° C. The transformed cells were spread on suitable selective plates.

II.4.3 B. subtilis Protoplast Transformation (Lars Christiansen) was performed by a modification of the method of Chang & Cohen (1979).

2 x SMM

1.5 M sucrose	100 ml	(final conc. 1.0 M)
0.2 M Maleic acid pH 6.5	30 ml	(0.04 M)
1.0 M $MgCl_2$	6 ml	(0.04 M)
Water	14 ml	

adjusted to pH 6.5 and sterile-filtered.

MC buffer

1 vol. 2xLB
1 vol. 2xSMM

PC plates (1 L): Mixed from stocks prewarmed or cooled to 60°C

(excepting BSA).

8 g Noble agar in	300 ml
1 M Na-succinate pH 7.3	500 ml
10% yeast extract	50 ml
20% casamino acids	25 ml
10 x MM	25 ml
20% glucose	25 ml
5% BSA	2 ml
1 M $MgCl_2$	20 ml

Antibiotic was added when required eg. 3 $\mu g ml^{-1}$ chloramphenicol

PEG-C: 40 g PEG-6000 and 15 mls water. Autoclaved and 50 mls of 2xSMM added.

Procedure

L Broth was inoculated from a fresh plate and the cells grown to A_{450} c.1 (reckoning 1.5 ml per transformation). The cells were harvested and washed in an equal volume of warmed MC Buffer, then resuspended in 400 µl MC containing 1 $mg ml^{-1}$ lysozyme for each 1.5 ml of original culture.
They were incubated at 37°C for 10-20 mins. while protoplast formation was monitored regularly under the microscope.
The protoplasts were sometimes frozen at this stage, by freezing slowly in cryotubes packed in ice and then placed at-80 °C. Each tube had a hole in the cap. Thawing prior to use was rapid (performed at 65° C).
10 µl DNA containing 1 µg or more and 600 µl PEG-C were added to the protoplasts, mixed gently and incubated at 37 °C for 2 mins. 400 µl MC Buffer was added, and the sample centrifuged for 30 secs. The supernatant was aspirated, and the pellet carefully resuspended in 200 µl MC, and incubated for 5 mins. as above, before plating out on PC plates.

II.4.4 Transformation of B. licheniformis Competent Cells

II.4.4.1 By a modification of the method of Sherratt & Collins (1973).

NBSG-X, BLSG, TM as described by Thorne & Stull (1966).
Penassay (Difco) plates were used. They were identical to NBY

plates (Thorne & Stull, 1966) except for having a lower Beef extract concentration (1.5 gL^{-1} vs. 3 gL^{-1})

Schaeffer's medium (1L): peptone 5 g, beef extract 3 g, $MgSO_4.7H_2O$ 0.25 g, KCl 1 g.

Added after autoclaving from sterile stocks:

$MnSO_4$	1 ml	50 mM stock
$FeSO_4.7H_2O$	0.37 ml	0.1 M stock
$Ca(NO_3)_2.4H_2O$	10 ml	of 2.4 g % w/v stock.

Procedure

1) Spore inocula were prepared by growing cells in 25 mls Schaeffer's medium for 3 days, harveting spores, and rinsing twice with sterile H_2O. These were resuspended in 1 ml H_2O, incubated 30 mins. at 65°C, and stored at 4°C.

2) Spores were titrated: dilutions in Tryptone Broth were plated out on Penassay and colonies scored.

3) 10^6 spores were inoculated into 25 ml BLSG or NBSG-X in 250 ml flasks and were shaken at 200 rev/min at 37°C for 18 h.

4) Cells were diluted 1/30 in TM. DNA approx. 1 $mgml^{-1}$ was added in 100 µl 2 M NaCl and the cells were shaken vigorously for 45 mins.

5) Aliquots were plated out on Penassay (omitting the DNAase I step described for chromosomal DNA transformation) with antibiotic. As control, a 10^{-4} dilution was plated on Penassay not containing antibiotic, for viable cell estimation.

II.4.4.2 By a method suggested by R. Runi (Apothekernes laboratorium A/S)

BL Broth (and Plates)(1L): Tryptone 10 g, Yeast extract 10 g, NaCl 5 g, (15g agar).

Adjusted to pH 7.0 - 7.5 prior to autoclaving.

An overnight culture was grown up in BLSG, diluted 1:10 in TM, and incubated 3 h. The cells were concentrated 1:10 in the used medium. 10 µl DNA was added to 90 µl concentrated cell suspension, and shaken slowly at 37 °C for 30 mins. 900 µl BL Broth was added, and the cells incubated with vigorous shaking for not less than 3 h.

The cells were then plated out on BL containing 15 µg ml^{-1} kanamycin.

II.4.5 Transformation of B. licheniformis Protoplasts was performed by modifications of the methods of Chang & Cohen (1979), Imanaka et al (PVP addition, high Km concentration; 1981) and Akamatsu & Sekiguchi (HMP medium; 1982, 1984).
I am grateful to Poul-Erik Pedersen for the tip concerning the addition of 1% starch to HMP-1.5 medium.

Solutions as for B. subtilis (II.4.3), and in addition:
6% w/v PVP: polyvinylpyrrolidone K90 (Serva). Autoclaved only 10 mins.
20% glucose (autoclaved)
2M Na-succinate pH7.3 (autoclaved)
K_2HPO_4-KH_2PO_4 3.5-1.5% w/v (autoclaved)
HMP-1.5 Plates (1L):
The following were mixed together, warming the solutions slightly to reduce viscosity.
25 ml glucose 20%
250 ml Na-succinate 2M
100 ml K_2HPO_4-KH_2PO_4 solution
250 ml 6% PVP solution
10 ml L-tryptophan 1%
20 ml $MgCl_2$ 1M (sterile)
The following solutions were cooled slightly before mixing with the salts:
50 ml Na-citrate-$(NH_4)_2SO_4$
5 g starch (autoclaved separately in a flask in 100 ml H_2O).
8 g agar autoclaved separately in 100 ml H_2O.
Antibiotics were added as required. The plates were not dried before use.
HCP-1.5 Medium
As HMP, but instead of Na-citrate-$(NH_4)_2SO_4$, 0.5 g% casamino acids were included (starch and agar were omitted).

Procedure

20 mls L Broth was inoculated from a fresh overnight culture and grown to A_{600} c. 0.5.

The cells were diluted 20X by addition of 1 ml to 20 ml fresh L Broth, and were agitated vigorously again at 37 °C until the A_{600} was again c. 0.5. The cells were harvested and resuspended in 1 ml MC Buffer containing 5 mg lysozyme, and incubated at 42 °C with gentle agitation. Protoplast formation was monitored under the microscope. After about 45 mins., when all the cells had protoplasted, harvesting was performed gently at room temperature by spinning at 4,000 xg for 5 mins. Protoplasts were gently resuspended in 1 ml MC Buffer, and pelleted as above. The wash was repeated once, and the protoplasts resuspended in 200 µl MC Buffer. 50-100µl were transformed by addition of DNA (in 10 µl), to which an equal volume of 2 x SMM had been added. Immediately, 500 µl PEG-C was added and mixed in carefully. After 2 mins., 4.5 ml MC Buffer was added, and the protoplasts harvested as before. They were then resuspended in 1 ml HCP-1.5 which had been prewarmed to 30 °C, and 10^{-1} and 10^{-2} dilutions made in the same medium. The protoplasts were incubated at 30 °C for 1.5 h without agitation, to allow phenotypic expression of the selective marker, before plating out 100 µl aliquots directly on to HMP-15 plates containing respectively eg. 0;250;500 and 750 $\mu g ml^{-1}$ Km and on L Agar control plates containing $10 \mu g ml^{-1}$ Km. The plates were incubated 2-3 days at 37 °C.

II.4.6 Preparation of E. coli Competent Cells

Cells were prepared as described in Maniatis et al (1982), with modifications, using sterile glassware and solutions.

An L Broth overnight culture was grown up , and 100 mls L Broth inoculated next day 1:100. The cells were shaken vigorously in a 1L flask at 37 °C to A_{550} 0.5 (corresponding to approximately 5×10^7 cells/ml Rec^- cells), then chilled on ice for 10 mins. before harvesting at 4 °C. The cells were resuspended in 50 mls ice-cold 50mM $CaCl_2$ -10mM TRIS pH8, then stored at 4 °C for 24 h. The cells were either used immediately, or else frozen until required. They

were frozen as follows: 1.25 ml glycerol from a cold 80% solution was added, 200 µl aliquots were dispensed into cryotubes, then frozen in liquid nitrogen and stored at -80 °C.

II.4.7 Transformation of E. coli Competent Cells

Either freshly aged cells were used, or an aliquot of frozen cells was thawed slowly on ice (takes 15 - 30 mins). DNA was added in a volume not exceeding 10 µl. Approx. 1 ng intact plasmid or up to 40 ng plasmid DNA in a ligation mix was added. The cells were incubated 30-60 mins. on ice with DNA before being heat-shocked by incubation for 2 mins.at 37 °C. An expression period was found not to be required for Ampicillin resistance, but for selection by other markers, L Broth was added to 1 ml, and the cells incubated at 37 °C for 30 mins. without shaking.
10 µl, 50 µl and the remaining volume of the transformation mixture were spread on suitable plates (or 5x200 µl if L Broth had been added).

II.5 Preparation of RNA from B. licheniformis was performed by a modification of the method of Berk & Sharp (1977) described in Gilman & Chamberlin (1983), with a few modifications.

Diethylpyrocarbonate (DEP) - treated H_2O

1L H_2O + 1ml DEP (Sigma). Was shaken occasionally, and autoclaved after 24 h.

DEP-treated water was used to make all solutions used for the mRNA preparation.

Disruption Buffer

30 mM TRIS-HCl (pH 7.5), 100 mM NaCl, 5 mM EDTA, 1% SDS
100 µg/ml proteinase K (Boehringer) was added prior to use (see below).

Preincubation Buffer

20 mM TRIS pH7.5-10 mM CaCl

DNAase I Digestion Buffer for RNA

20 mM TRIS-HCl (pH8), 10 mM $MgCl_2$, 2 mM $CaCl_2$

100 $\mu g ml^{-1}$ proteinase K-treated DNAase I added (see below).

DNAase I (Worthington, RNAase-free)

Procedure

1) B. licheniformis was grown in modified Schaeffer's medium. Samples were taken at A_{450}=1 (exponential growth; 40 ml); and T_6 or T_{14} (stationary phase; 20 ml).

2) The cells were spun down in autoclaved 250 ml tubes 1/3 full of freshly made ice, pre-chilled to -20°C. Cells were spun down in a chilled rotor at 0°C, and resuspended in 5 ml Disruption Buffer containing proteinase K.

3) Cells were sonicated 3x 10 secs. at 30 W and incubated 60mins. at 37° C.

4) The lysate was extracted twice with phenol: chloroform, once with chloroform: isoamylalcohol (24:1), once with chloroform , and precipitated with 2.5 vols. ethanol, spun down and dried.

5) DNAase I was pretreated with proteinase K to remove any residual RNAase activity as follows: (see Tullis & Rubin, 1980): DNAase I was made 1 $mg ml^{-1}$ in 20mM TRIS pH 7.5-10 mM $CaCl_2$ (Preincubation Buffer) and preincubated 15 mins. at 37°C. This was added to proteinase K to give a concentration of proteinase K of 1 $mg ml^{-1}$ and incubated 1 h at 37°C, after which it was immediately used on the nucleic acid preparation.(As an alternative, a DNAase I preparation such as RQ1^{TM} (Promega Biotech) can be used).

6) The nucleic acids were redissolved in 300 µl DNAase I Digestion Buffer, containing 1/10 vol. proteinase-K treated DNAase I (ie. 100 $\mu g ml^{-1}$) and digested at 37°C for 60 min. SDS was then added to 1% (1/10 vol 10%), EDTA to 50 mM (1/5 vol. 0.25 M) and NaCl to 0.2 M (1/10 vol. 2 M). The mixture was extracted once with phenol/chloroform, twice with chloroform, and precipitated with EtOH. RNA was resuspended in DEP-H_2O and A was measured.

260

II.6 S1 Nuclease Mapping

was carried out as described in Gilman & Chamberlain (1983)

Hybridization Buffer: 40 mM Pipes (pH 6.4),0.4 M NaCl,1 mM EDTA

Loading Buffer: 80% formamide, 10 mM NaOH, 1 mM EDTA, BPB, XCFF (Sigma, Biorad resp.) 0.04%.

S1 Digestion Buffer: 30 mM sodium acetate (pH 4.6), 0.25 M NaCl, 1 mM $ZnSO_4$, 5% glycerol, 20 µg/ml sonicated denatured salmon sperm DNA.
S1 Nuclease (Boehringer RNAase-free)

S1 Nuclease Mapping Procedure

Labelled hybridization probe was mixed with approximately 20 µg mRNA or 20 µg yeast tRNA (Calbiochem), and precipiated with EtOH.

The precipitate was rinsed with 70% EtOH, lyophilized and redissolved in 10 µl hybridization buffer, which was overlaid with paraffin. The sample was heated at 90 °C for 5 mins, and transferred to temperatures from 45 °C to 66° C, where it was incubated for 3 h. 200 µl ice-cold S1 Digestion Buffer containing 50U S1-nuclease (Boehringer) was added. 1 unit was defined as producing 1 µg of free Pi per minute. Incubation with S1-nuclease was for 30 mins. at 37° C, after which the nucleic acids were precipitated with EtOH, and rinsed with 70% EtOH. They were redissolved in Loading Buffer, Heated for 2 mins. at 90° C to denature, and plunged into ice. The samples were applied to 8% acryl-amide-urea sequencing gels (II.12.2).

II.7 Labelling Techniques

II.7.1. Nick-Translation

was carried out as described below.

Nick translation buffer (see II.7.2)

DNAase I (Worthington) was kept at -20°C as a stock solution of 1 $mgml^{-1}$ (in 50% glycerol-TE-50 mM KCl-1 mM DTT-100 µg ml^{-1} BSA). An aliquot was diluted in H_2O just prior to use. 2µl of 10^{-4} dilution was added to the reaction).

DNA c.	0.2 µg
10 x Nick Translation Buffer	10 µl
α-^{32}P dNTP eg. dATP	30 µCi
2 mM cold dNTPS (as needed) eg.dC,dG,dTTP:	4 µl of each
H_2O up to	100 µl
DNAase I	0.2 ng

The mixture was incubated at room temperature for 5-15 mins. (period giving good incorporation was determined empirically for each DNAase I suspension). The sample was heated at 65°C for 10 mins., then cooled on ice. 1 µl DNA polymerase I was added, and incubation proceeded at 16°C for 40-60 mins. The labelled DNA was phenol/$CHCl_3$ extracted once.

Removal of Unincorporated Nucleotides

Desalting to remove unincorporated labelled nucleotides was performed by standard methods eg. on NENSORB 20 Nucleic Acid Purification Cartridges (NEN Research Products).

II.7.2 End-Labelling

II.7.2.1 Labelling of Recessed 3' Ends ("Filling-Out")

was performed as described by Maniatis et al (1982), p. 113.

DNA	⩽ 12.5 µg
10 x Nick-Translation Buffer★	2 µl
α-^{32}P dNTP (2 pmole)	5 µCi
2 mM dNTP cold (as needed) (2nmole each)	1 µl
DNA polymerase (large fragment i.e. Klenow enzyme):	1 unit
H_2O up to	25 µl

After incubation at room temperature for 30 mins., 1 µl 0.5M EDTA (pH 8) was added, and the sample extracted with phenol/$CHCl_3$. Desalting was carried out as in II.7.1, or else the labelled fragment was used as a probe directly after boiling.

★10 x Nick Translation Buffer

0.5 M TRIS pH 7.2, 0.1 M $MgSO_4$, 1 mM DTT, 500 µgml^{-1} gelatine

II.7.2.2 Labelling Ends With 3' Overhang

was performed as described by Maniatis et al, 1982 p. 117.

10 x T4 DNA Pol Buffer

0.33 M TRIS-acetate pH 7.9, 0.66 M K-acetate, 0.1 M Mg-acetate, 5 mM DTT, 1 mgml^{-1} BSA. Stored at -20° C.

DNA	0.5-2 µg
10xT4 Buffer	2 µl
eg. α -^{32}PdATP	2 µCi
2mM each eg. dGTP, dCTP, dTTP	1 µl
H_2O up to	19 µl
T4 DNA polymerase (PL Biochemicals)	2.5 units

The mixture was incubated at 37°C for 5 mins. 1 µl 1mM cold ATP was added, and incubation continued for 10 mins. The reaction was terminated by heating for 5 mins. at 70°C.

II.7.3 Kinasing (& Labelling) of Linkers & Oligomer Probes was performed as described below.

Up to 1 A_{260} unit of linker DNA was labelled in 50 µl TE (7-10 nmoles) or 40 pmol of oligomer probe (=0.15 µg 14-mer).

For probes:

DNA in	25 µl
1M TRIS-HCl (pH 7.5)	6 µl
0.1M $MgCl_2$	10 µl
H_2O	46 µl

The mixture was heated at 70°C for 2 mins., and cooled to 37°C.

1µl 10 mM ATP (or 70 µCi γ-^{32}P ATP).

10 µl β-mercaptoethanol 0.1 M (i.e. 0.78 g/100 ml)

1 µl kinase

were added, and the mixture incubated at 37 °C for 30 mins. The 70 °C heating, addition of reagents thereafter, and incubation were repeated once, but omitting a second addition of labelled nucleotide. The sample was then phenol/chloroform extracted, extracted with $CHCl_3$, and precipitated with EtOH.

For linkers:

Linkers were kinased as above, but extra ATP was added when the kinase step was repeated. At the end of the second incubation, the sample was heated again at 70 °C in a water bath (eg. a beaker). The sample and water bath were transferred to 4 °C and the strands allowed to anneal overnight. Such samples were stored at -20°C.

II.8 Hybridization Techniques

II.8.1 Southern Blotting

was carried out as described below.

Transfer of DNA to filters was by osmosis as described by Southern (1975) in 20xSSC. PALL Biodyne™ transfer membranes were used and the accompanying manufacturer's instructions followed regarding membrane treatment, hybridization, "stripping" and re-hybridization. Gels were denatured in 1.5 M NaCl-0.5 M NaOH for 30 mins., and neutralized in 3 M Na-acetate (pH 5.5) for 30 mins. Filters were not rinsed after transfer, and were baked for 1h at 80° C.

1xDenhardt's solution: 0.02% w/v BSA (Sigma Fraction V), PVP, Ficoll-400.

1xSSC = 0.15 M NaCl, 0.015 M Na-citrate

Hybridization Buffer (Large Probes)

5xSSC, 5xDenhardt's solution, 50 mM Na-phosphate (pH 6.5), 0.1% SDS, 250 $\mu g ml^{-1}$ salmon sperm DNA, (Sigma) (sonicated 5 mins.), 50% formamide (deionized), 10 $\mu g\ ml^{-1}$ polyA.

Filters were preincubated 1 h at 42°C prior to addition of the denatured probe. Incubation was at 42°C, in sealed plastic bags.

Washing Procedure (Large Probes)

4 washes in 2xSSC -0.1% SDS at room temperature (30 mins./wash).

2 washes in 0.1xSSC - 0.1% SDS at 50° C (30 mins./wash).

Hybridization Buffer (Oligomeric Probes)

6XSSC, 10x Denhardt's solution, 0.1% SDS, poly A 10 $\mu g ml^{-1}$

Preincubation was at appropriate temperature overnight prior to addition of probe.

Washing Procedure (Oligomeric Probes)

1 wash in 6xSSC / 0.1% SDS (1 h)

4 washes 2xSSC / 0.1% SDS (30 mins./wash) at appropriate temperature

II.8.2 Colony Hybridization (E. coli) (Whatman 541 Filters)

Filter Preparation

The colonies were replica-plated, reserving one plate for later isolations. A Whatman filter was placed on one plate, and it was incubated for a further 2 h at 37° C.

The filter was transferred, colony-side down, to Cm^{200} plates, and incubated for 24 h at 37°C. The filters were treated, with shaking, for 2x5 mins. in the following solutions, at room temperature:

(a) 0.5 M NaOH, (b) 0.5 M TRIS pH 7.4, (c) 2xSSC pH 7.0.

Each filter was briefly dipped in 96% EtOH, and air dried.

Hybridization

The filters were hybridized in 6x NET for 1 h at 55°C (6xNET = 0.90 M NaCl; 0.09 M TRIS HCl pH 7.5, 0.06 M EDTA + 0.5% Nonidet P40 (BRL)).

Prehybridization was in above mixture containing also 100 $\mu g ml^{-1}$ sonicated denatured E. coli DNA for 2-16 h at 55°C. Hybridization was in 6xNET containing 250 $\mu g ml^{-1}$ tRNA, and the denatured labelled probe at an appropriate temperature (probe-dependent).

Washing Filters

Filters were washed 5x (10 mins./wash) in 250 mls 6xSSC at 0°C, and 3x as above at the hybridization temperature.

In my hands, only oligomeric or end-labelled (not nick-translated) probes could be used effectively in this procedure.

II.8.3 Colony Hybridization (B. subtilis) (PAL Biodyne filters) was carried out by the following protocol.

Filter Preparation

Colonies were replica-plated, and transferred to PAL Biodyne filters by laying the filter on a plate for 1 minute. The filter was placed, colony-side up, on a pad of filter paper soaked with 0.5M NaOH-1.5M NaCl, and left 5 mins. It was then placed on a pad soaked in 3M Na-acetate pH5.5, and left for 5 mins., after which it was blotted well between two sheets of filter paper, and air dried. The filter was baked for 1 h at 80°C.

Hybridization to Kinased Oligomeric Probes (see II.8.1).
5% of a portion of oligomer labelled up as per II.7.3 was needed per filter.

Hybridization to End-Labelled Probes
Preincubation was performed as above. The filter was incubated with denatured labelled probe in: 50% deionized formamide, 5x Denhardt's solution, 5x SSC, 50 mM Na-phosphate (pH 6.5), 0.1% SDS, 20 $\mu g ml^{-1}$ polyA, 100 $\mu g ml^{-1}$ E. coli sonicated denatured DNA, for 24 h at 42°C, and washed as above.

II.8.4 Northern Blotting
was performed as described below.
The RNA denaturing gels were run according to the method of Lehrach et al (1977), exactly as described in Maniatis et al (1982), p.202, in a gel apparatus reserved for RNA work.
The 20x20 cm agarose-formaldehyde gel was run at 60 V overnight with recirculation of 1xMOPS (Sigma) buffer.
Hybridization to DNA Probe was performed in 50% formamide as in II.8.1, but with 1% SDS, 500 $\mu g ml^{-1}$ salmon sperm DNA instead of 250 $\mu g ml^{-1}$, and omitting polyA.
Prehybridization: a minimum of 4 h at 42° C.
Hybridization: overnight at 42° C.
Washing Procedure

2 washes in	2xSSC	(5 mins. per wash, room temp.)
2 " "	2xSSC, 0.5% SDS	(30 mins. per wash, 65° C).
2 " "	0.1xSSC	(30 mins. per wash, room temp.)

II.9 Enzyme Assays
II.9.1 Proteolytic Activity Determinations
Exoproteolytic activity was assayed in 2 ways.
II.9.1.1 Plate Assay
Culture supernatants were applied to wells in 1% skim-milk, 1% agarose in 25mM sodium phosphate (pH8). These plates were the most sensitive, and were made by boiling the agarose in buffer, followed by cooling and the addition of skim-milk from a 10%

sterile stock solution.
Alternatively, in a modification of the dual-substrate plate assay described by Montville (1983), they were applied to wells in a plate containing 1% skim-milk, 1% gelatine, 1.5% agar in 25 mM sodium phosphate. The gelatine and agar were autoclaved in buffer, and after cooling the skim-milk was added from a 10% stock. This method gave clearer, better defined halos that photograph more easily but were not sensitive to small changes in exoprotease activity.

II.9.1.2 Azocoll Assay

Azocoll (Calbiochem 50-100 mesh). The grains were ground further in a pestle: the more finely ground, the more intense the colour that developed. For consistent results, the same batch had to be used.

Buffer: 10 mM TRIS HCl - 100 mM NaCl, pH 6.8

EDTA: 0.5 M pH8

PMSF: 10 $mgml^{-1}$ i.e. 1% in isopropanol.

Samples: Aliquots of culture supernatant (spun again to remove residual cells), preferably fresh. Otherwise the supernatant was frozen at -20°C, and thawed just prior to use.

Note: About 30% activity was lost in one freeze - thaw cycle. Samples kept overnight at 4°C also lost considerable activity.

Procedure

1. 100 mg Azocoll was weighed out into conical flasks.
2. The sample (30-2000 µl) was preincubated at 37°C with or without inhibitor for 30 mins. The buffer was also equilibrated to 37° C for 30 mins.

Inhibitor concentrations in the sample were 1 mM PMSF, 10 mM EDTA, and correspond to the addition of 25 µl of each of the stock solutions per ml sample.
Blank samples were prepared with 1 ml medium instead of culture supernatant.

3. At the end of the preincubation, 60 mls of buffer were added

to the Azocoll grains, and the sample was added.
4. Incubation was for 15 mins. at 37°C with vigorous shaking.
5. After precisely 15 mins., samples were filtered free from the remaining Azocoll through Whatman filters.
6. The A_{520} nm was read.

Note: Preincubation with inhibition was essential for a specific effect.
Azocoll could not be added as a suspension.

A mini-assay with 5 mg Azocoll in a total reaction volume of 3 mls similarly gave a linear response up to an OD of c.0.6, which corresponded to proteolysis by 6 µg of a preparation of subtilisin Novo under standard conditions. The mini-assay however, gave less reproducible results than the method described above.

II.9.2 Xylosidase Assay (Sven Hastrup)

Intracellular activity was detected and quantified as follows.

Plate Assay

The plates were sprayed with 1 mM MUX i.e.4-Methyl-umbelliferyl-β-D-xylopyranoside in 10 mM K-PO_4 buffer, pH7.0-1mM EDTA, and visualized under UV light. Positives fluoresced with a blue glow.

Colorimetric Assay

Samples of culture were frozen at -20° C until assayed.

100 µl culture was added to 1 ml PE*

*PE = 10 mM K-phosphate - 1 mM EDTA (pH7.1).

100 µl PE containing a pinch of lysozyme was added and the mixture incubated at 37°C for 5 mins. When lysis was evident, 100 µl PNPX was added, and incubation continued for 1 h. 1 ml 1M Na_2CO_3 was added, and A_{430} read vs. a reference (100µl medium, 1.1ml PE, 100µl PNPX, 1ml Na_2CO_3).

*PNPX = 12 mM p-nitro-phenyl-β-D-xyloside

II.10 Selection of Tet^S Colonies (Bochner et al. 1980)

A freshly grown culture was diluted in phosphate buffered salt solution to about 10^7 cells/ml. (A_{550} = 0.5 corresponds to approx. $5x10^7$ Rec^- cells/ml; A_{550} = 0.2 corresponds to approx. $5x10^7$ Rec^+ cells/ml)

0.1 ml of such a suspension was plated on fusaric acid plates.

Fusaric Acid Plates (1L) : agar 15 g, bactotryptone 10 g, yeast extract 5 g, NaCl 10 g, glucose 2 g, chlortetracycline-HCl 50 mg (Sigma), $NaH_2PO_4H_2O$ 10 g.

Above ingredients were autoclaved and cooled. 6 ml fusaric acid (2 mg/ml), 5 ml 20mM $ZnCl_2$, and auxotrophic requirements were added.

II.11 P1 Transduction

Sources: N. Fiil, T. Christensen, and Miller, J. H. (1972).

R-Agar (1L): 10 g tryptone, 8 g NaCl, 1 g yeast extract

Top Agar contained 8 gL^{-1} agar

Plates contained 10 gL^{-1} agar

Above components were autoclaved and the following added:

2 ml 1M $CaCl_2$, 5 ml 2% glucose (per liter)

The plates were always freshly prepared ie.at most one day old.

LB containing 5 mM $CaCl_2$

Sodium citrate 1M

McConkey Agar (Difco) Plates contained 1% maltose (added from a 10% sterile-filtered solution after autoclaving).

AB Minimal Plates (Phosphate Buffer)

A: per 200 ml: $(NH_4)_2SO_4$ 2g, Na_2HPO_4 6g, KH_2PO_4 3g, NaCl 3g.

B: per 800 mls: 1 M $MgCl_2$ 1ml, 0.1 M $CaCl_2$ 1ml, 3 mM $FeCl_3$ 1ml.

The 3 solutions comprising B were autoclaved separately.

A and B were mixed in the ratio 1:4 (pH7-7.2).

Agar (final concentration 2%) was autoclaved in 800 ml H_2O.

After cooling to 60°C, the following were mixed together:

200ml AB, 800ml agar, 1ml each 1 M $MgCl_2$, 0.1 M $CaCl_2$, 3 mM $FeCl_3$.

The selection substrate, here maltose, was added to 1% from a 10% sterile-filtered solution and sodium citrate added to 20mM from a 1 M stock to prevent reinfection. In addition, auxotrophic

requirements and thiamine (1 ml of a 1 $mgml^{-1}$ stock) were added.

Methodology for Preparation of P1 Lysate was exactly as in Miller (1972) viz. One drop of an overnight culture of the donor strain was subcultured into 5 ml of L Broth containing 5 mM $CaCl_2$, and grown till the cells reached a density of $2x10^8$ cells/ml (OD_{450}=1 corresponds to approx.3 $x10^8$ cells/ml). Plvir was preadsorbed by adding 10^7 phage to 1 ml of this exponential culture and incubating 20 mins. at 37°C.

2.5 mls molten R-Top Agar was added, and the cells plated onto a fresh R plate, which was incubated at 37°C, face up, for 8 h. The Top Agar was scraped off, the plate rinsed with 1 ml L Broth, which was pooled with the Top Agar. 5 drops of $CHCl_3$ were added, and vortexed for 30 secs. After standing at room temperature for 10 mins., the debris was centrifuged down, and the supernatant stored at 4°C with a few drops of $CHCl_3$.

Generation of Recipient Strain - see text (III.4.4.2.3)

Transduction Procedure (N. Fiil)

An overnight culture of the recipient strain was grown in L Broth containing 5 mM CaCl . A portion of the P1 lysate was allowed to air for a while to remove traces of $CHCl_3$. The lysate and cells were incubated together for 20 mins. at 37°C in the following ratio v/v: Lysate: Cells:: 10:1 & 1:10.Usually, 10 µl, and 100 µl were used. The cells were spread directly onto the selective AB plates containing 20 mM sodium citrate to prevent reinfection and incubated 48 h at 37°C.

II.12 Gels

II.12.1 Gels: Non-Denaturing

PB Buffer= Running Buffer

10X PB (2L): Tris-HCl 216 g, boric acid 110 g, EDTA 18.6 g (pH approx. 8.3).

For analytical or preparative purposes the following were used:

0.8-1% agarose: intact plasmids, fragments down to 1 kb

1.5% agarose (Sigma): fragments 400 bp - 1 kb

1.7% agarose : fragments 100 - 400 bp

5% polyacrylamide : fragments smaller than 400 bp

Agarose gels and Running Buffer contained ethidium bromide at 1 μgml^{-1}.

Acryl-Bis: acrylamide-N,N methyl bisacrylamide (Serva) 29:1% w/v
5% Polyacrylamide Gel: 16.7 ml acryl-bis, 10ml 10xPB, 77.7ml H_2O, 40 µl TEMED (Biorad), 0.64 ml 1% ammonium persulphate
Prerun 1 h at 200 V, run at 200 V.
Stained 10 mins. with EtBr solution 1 $\mu g ml^{-1}$
Samples: Applied after addition of 1/5 vol. Ficoll mixture (35% Ficoll-35;0.1% BPB (Sigma) & XCFF (Biorad), 40 mM EDTA).

II.12.2 Gels: Denaturing
0.15 mm thick, 25 x 40 cm or 80 cm
Acryl-bis 40 - 1.3% w/v
Buffer = 3/4 PB

Acrylamide conc.	20%	8%
Urea	72 g	72 g
acryl-bis	80 ml	32 ml
H_2O	25 ml	70 ml

The above ingredients were dissolved over gentle heat with stirring. 1 tsp. Amberlite Monobed Resin (Serva) was added, and stirred for 10 mins. The beads were filtered off using a Whatman filter, and the volume adjusted to 154 ml with water. 12.5 mls 10xPB were added, and the mixture degassed briefly. 40 µl TEMED (Biorad), and 0.9 ml of a freshly made 10% APS solution were added. The gel was cast, and run at 1W/cm. The position of the XCFF dye corresponds to that of a 67b fragment on 8% gels, to 29b on 20% gels.
Samples: Applied after boiling 2 mins. in 8M urea (deionized) containing 0.1% XCFF.

II.12.3 SDS Protein Gels
20x20 cm x 1 mm
Lower Gel Buffer: 1.5 M TRIS-HCl pH 8.8 contained 0.4% SDS
Upper Gel Buffer: 0.5 M TRIS-HCl pH 6.8 contained 0.4% SDS.
Acryl-Bis : 29.2-0.8% w/w i.e.100g-2.74g made up to 343ml.

Lower Gel		Upper Gel
14 ml	Acryl/bis	1.5 ml
10.5 ml	appropriate buffer	2.5 ml
17.5 ml	water	6.0 ml
252 µl	10% APS	30 µl
33 µl	TEMED	10 µl

The mixture was degassed prior to the addition of TEMED.

2x Sample Buffer

25% Upper Gel Buffer, 20% glycerol, 4.6% SDS, 10% beta-mercaptoethanol (added just prior to use), and a sprinkling of BPB.

Sample Dilution Buffer with Inhibitors

(for protease-containing samples)

1 vol 2x Sample Buffer, 1/5 vol PMSF (sat. solution in EtOH), 1/5 vol 1% EDTA (in water). Made up to 2 vols with water.

SDS Buffer (Running Buffer) (1L)

0.025 M TRIS 3 g, 0.192 M glycine (pH8.3) 15 g, 0.1% SDS 1g

Molecular Weight Standards (Pharmacia): phosphorylase b, 94kdal; BSA, 83kdal; ovalbumin, 43kdal; carbonic anhydrase, 30kdal; soybean trypsin inhibitor, 20.1kdal; alpha-lactalbumin, 14.4kdal.

Subtilisin Markers: 1 $mgml^{-1}$ in Sample Dilution Buffer. Applied 1-20 µl.

Gel Preparation

A 3 ml plug of lower gel mixture was poured, after addition of extra TEMED and APS. The Lower Gel was poured to within 4 cm of the top, and overlaid with water. Once it had polymerized, the water was removed and the Upper Gel, in which the comb was inserted, was cast. The gel was run at 150 V (c. 20 mA) while the BPB dye moved through the Upper (stacking) Gel, thereafter at 250-300 V until the BPB dye reached the bottom.

Staining of SDS Gels

Brilliant Coomassie Blue R at 200 mg/l in 7.5% acetic acid, 50% EtOH, was used. Gels were stained for 1 h, then destained overnight in 7.5% acetic acid. After photography, the gels were transferred to filter paper (Whatman/Chr) and dried down.

II.12.4 Preparation of Samples for SDS Gels

From Bacillus spp.

Culture Supernatants

20 ml L Broth overnight cultures were used. The A_{450} was read, and a volume normalized for cell density differences was taken from each sample (c. 10 mls). The cells were harvested at 7k rpm for 10 mins, and 1/10 vol. ice-cold 50% TCA added to the supernatant, which was set on ice for 1 h. The proteins were collected by centrifugation at 10k rpm for 20 mins., and the pellets drained well. 10 µl 10N NaOH and 50-100 µl 2x Sample Buffer were added. If the sample turned yellow, 5 µl 1M TRIS pH8 was added. Inhibitors were added if proteases were present, and the volume was made up (see II.12.3, Sample & Dilution Buffers, as appropriate).

From E. coli Cells

Cells equivalent to 0.25 A_{450} units were resuspended in 20 µl 1x sample buffer and applied to one slot.

All samples were boiled for 3 minutes prior to application to gel. I samples were used a second time, they were reboiled.

II.13 Western Blot

Transfer of Proteins from SDS Gel to Filter

0.1, 0.2 or 0.45 micron nitrocellulose filters (Schleicher & Schull), and a Biorad Transblot apparatus were used, and run at 25-30 V (100-150 mA) overnight. A control strip of filter was stained with Amido Black to check transfer, and visualize marker bands.

All steps at room temperature unless otherwise stated.

ELISA Procedure

Filters were incubated with 200 mls Blocking Buffer, for 2 mins. with gentle agitation, then with 50 mls Washing Buffer containing the specific antiserum in different concentrations: 1000-10000x dilutions were incubated overnight; 300-3000x dilutions for 1-4 h. 100 µl PMSF from a saturated EtOH solution, and 100 µl 1% EDTA were added. The filters

were washed 3x in 100 ml Washing Buffer (10 mins./wash) with gentle agitation. They were incubated for 2 h with 50mls Washing Buffer and the 2^{o} antiserum-peroxidase (Dako), then washed 3x as before. One 10 minute wash in 100 ml acetate Staining Buffer, was followed by a 5 minute incubation with freshly made Peroxidase Substrate Solution with gentle agitation. The filters were washed several times in H_2O, for 10 mins. at a time, then washed once in 50mM $Na_2S_2O_5$ for 5 mins. The filters were finally dried at room temperature, and photographed.

Solutions for Western Blotting

Transfer Buffer (pH 8.4) (3L):25 mM Tris 9.08 g, 192 mM glycine 43.2 g, methanol 600 ml, SDS 0.1% 30 ml 10% stock.

Washing buffer (pH 10.2) (5L) : 50 mM Tris 30.3 g, 150 mM sodium chloride 43.8 g, 5 mM sodium azide 1.6 g, 0.05% Tween 20.

Blocking Buffer (pH 10.2)(1L): 2% (w/v) Tween 20 20 g (Merck art. 822184), Washing Buffer ad 1000 ml.

Staining Buffer (pH5.5) (1L) : 50 mM sodium acetate 6.8 g, glacial acetic acid 1.75 ml, water ad 1000 ml.

Peroxidase Substrate Solution

1% (w/v) 3-amino-9-ethyl carbazole in acetone (Sigma no.A-5754) (freshly made),50 mM sodium acetate buffer, pH 5.0 50 ml, 30% (w/v) H_2O_2 2.5 µl

In spite of some precipitation no filtering was necessary.

$Na_2S_2O_5$ 0.95g/100ml

Amido Black Staining Solution : 0.5 % (w/v) Amido Black 5 g, methanol 450 ml, water 450 ml, glacial acetic acid 100 ml.

Destaining Solution : as above without Amido Black.

II.14.1 DNA Sequencing (Maxam & Gilbert (1977))

A modification of the original procedure (Michael Berman NCI-Frederick Cancer Resarch Facility, Maryland), was used.

Solutions

5 M NaCl

Carrier DNA: 10 mg/ml Calf thymus DNA (Sigma)

DMS PPt: 3 M NH_4-acetate, 0.1 M Mg-acetate, 100 µg/ml carrier DNA, 25 mM beta-mercaptoethanol

Pyridinium-formate (Pyr-For): (10 ml)

1 M formic acid (0.45 ml 88% formic acid adjusted to pH2 with a few drops of piperidine)

NH_4 PPt: 2 M NH_4-acetate

DMS Buffer: 50 mM Na-cacodylate-NaOH pH8, 10 mM $MgCl_2$, 0.1 mM EDTA pH8.

Hydrazine PPt: 2 M NH_4-acetate, 20 $\mu g/ml^{-1}$ carrier DNA.

Chemicals (all to be used with caution)

Piperidine (10 M stock), diluted 10x just prior to use; piperidine was added to water.

Hydrazine 95-98% (WAS INACTIVATED WITH 2M $FeCl_3$)

Dimethyl sulfate (DMS)(WAS INACTIVATED WITH 5N NaOH)

SEQUENCING PROTOCOL shown on the next page was followed.

drawn-out Pasteur pipettes,
dry ice - ethanol bath (-70° C),
and 90° C, 30° C water baths were required,

and tubes were marked with alcohol-resistant ink.

In the work scheme shown on the next page, " ' " signifies "minutes".

REACTION	A&G	G	C&T	C
DNA	10	10	10	5
carrier DNA	1	1	1	1
water	9	-	10	-
5M NaCl	-	-	-	15
DMS Buffer	-	200	-	-
Pyrfor	2			
	30°C/30'			
Hydrazine			30	30
			20°C/8'	20°C/8'
HYD.PPt			200	200
EtOH			800	800
-70°C/5'				
Spin				
DMS		1		
		20°C/4'		
NH_4PPT			200	200
EtOH			700	700
-70°C/5'				
Spin				
DMS PPT		50		
EtOH		800		
-70°C/5'				
Spin				
NH_4 PPT	200			
EtOH	800			
-70°C/EtOH				
Spin				
	A&G	G	C&T	C

After the last centrifugation, all samples were treated in the same way. They were rinsed without resuspension with 1 ml 70% EtOH, spun briefly, and the supernatant carefully removed with a drawn-out pipette. Pellets were lyophilized, then 60 µl of 1M piperidine (i.e. 10x diluted) was added, and samples incubated at 90 °C for 30 mins., with a plate of glass clamped over the tubes' lids to

prevent them opening. Thereafter the samples were frozen in a dryice/EtOH bath and lyophilized, which process took approx. 1 h. The samples were dissolved in 100 µl H_2O before freezing and drying as above. The pellets were resuspended as described in II.12.2 so that at least 15 cps in at most 2 µl could be loaded for all reactions, excepting 7.5 cps for the C reaction.

NOTE: When the first 10 bases were not required to be read, EtOH precipitation was carried out immediately after the piperidine incubation as follows: 0.5 µl carrier DNA, 9 µl 5M NaCl, and 450 µl EtOH were added, the samples frozen at -80°C for 5 mins., and DNA spun down and dried. The samples were run on appropriate denaturing gels.

II.14.2 Sanger Method in M13

A procedure adapted from Messing et al (1981) & Sanger et al (1977), was used.

M13 Host Strain: JM101 (see Table 4).

This host was maintained on minimal plates to ensure that the episome was not lost. Cells were always grown from a single colony.

Media: (1L)

YT : Bactotryptone 8 g; Bacto Yeast Extr 5 g; NaCl 5 g

2xYT: Bactotryptone 16 g; Bacto Yeast Extr 10 g; NaCl 5 g

II.14.2.1 Mini-Prep. for Double-Stranded Replicative Form

The following were grown overnight at 37°C with shaking :

a blue plaque of JM101 infected with the desired M13 vector which has been grown on Xgal-IPTG plates (see below), and picked into 2 mls 2x YT, and 10 mls 2xYT inoculated with a single colony of JM101 Add 100 µl of the former, and 1 ml of the latter to 100 mls YT. Harvest the cells after 3 h incubation at 37°C with vigorous aeration, and proceed as for plasmid preparation (II.3.2.3).

Cloning into M13 Vectors : As standard procedures.
Transformation into JM101

20 mls YT was inoculated with a single colony of JM101, and grown 5-6 h at 37°C, until the A_{450} was about 0.3. The cells were harvested, resuspended in 10 mls cold 50 mM $CaCl_2$, and left on ice for 20 mins. before resuspension in 2 mls $CaCl_2$. To the dregs of the culture that was used for competent cells, 2 mls YT was added, and the cells incubated at 37°C during the transformation to provide a lawn upon plating out. The ligation mixture , preferably in several aliquots, was added to 200 µl aliquots of competent cells, and incubated on ice for 40 mins. The cells were then heat shocked for 2 mins. at 42°C. During the shock, the following reagents were mixed in the given order: 3 mls YT soft agar at 42°C (per plating), 30 µl Xgal (stock 20 mg/ml in dimethylformamide), 20 µl IPTG (stock 24 mg/ml in water), 200 µl of exponentially growing cells. This combination was added to the shocked cells and plated out on YT. The plates were incubated overnight at 37°C. Colourless plaques indicated the incorporation of DNA into the cloning site.

II.14.2.2 Preparation of Single-Stranded DNA Template

20 µl of a saturated overnight culture of JM101 was added to 25 mls of 2xYT, and dispensed into 1 ml aliquots. Individual plaques were picked into 1 ml broth, and shaken at 37°C for 5-6 h. At the end of this time, the cells had grown to high density. They were poured into Eppendorf tubes (tt), and spun 5 mins. Approx. 0.8 ml was decanted into a fresh tt. Sterile glycerol was added to the pellets, and the suspensions kept at -20°C to provide 'phage stocks. 200 µl of 2.5M NaCl/20% PEG 6000 was added to the supernatants, which were incubated at room temperature for 15 mins. before being spun (5 mins.). The supernatants were poured off, and the tts spun briefly to collect droplets of PEG. All remaining PEG was aspirated with drawn-out pipettes. The pellets were usually clearly visible at this stage.

100 µl TE and 10 µl 3M sodium acetate were added, as well as an equal volume of saturated phenol. The samples were vortexed for 10 secs., left 5 mins. at room temperature, then vortexed again for a full minute. After spinning to separate the phases, the aqueous phase was re-extracted with phenol, then extracted 3 times with ether. EtOH was added to precipitate DNA. The samples were quick-frozen, spun, and the pellets rinsed with 70% EtOH before drying. The DNA was resuspended in 50 µl TE. 5 µl was run on a minigel, omitting XCFF dye from the loading buffer. The ssDNA bands were very intense to the eye, at about the position of XCFF dye that was run in a marker track, when the gel was visualized under UV light after EtBr staining. It was essential to check that ssDNA had survived the isolation procedure in this way. The ssDNA was stored at -20 °C. Whenever a particular fragment was cloned, ssDNA was always prepared from several colourless plaques as well as a few blue ones, and put through a "T-screen."

II.14.2.3 The Sequencing Reactions

ssDNA concn: ? as made up in standard way. Primer at 2 ng/µl.

Preparation of Annealed Clones: (viz. hybridization of primer to ssDNA) was performed as described below by mixing in a 400 µl tt: 5 µl ssDNA, 1 µl Hin* + DTT, 2 µl primer, 2 µl water.

*Hin was as follows: 0.1 M Tris pH 7.4, 0.1 M $MgCl_2$, 0.5 M NaCl. 1µl of 0.1 M DTT was added to 9 µl Hin immediately prior to use.

The samples were boiled 5 mins. and the cloned allowed to come to room temperature gradually. 2 µl of such hybridized clones was used for each reaction.

Note: New clones were usually "T-screened" by performing just the T reaction.

The Dideoxy Reactions:

For each base reaction the following was used

1-5 µCi of α-^{32}P dATP (Specific activity 400 Ci/mmol).

1 µl dNTP' solution (deoxy NTPs)

1 µl ddNTP solution (dideoxy NTP)

0.2 µl Klenow enzyme (0.5U/µl)

The dNTP' mixes were prepared as follows:

	dTTP'	dCTP'	dGTP'	dATP'	
0.5mM dTTP	1	20	20	20	microlitres
0.5mM dCTP	20	1	20	20	
0.5mM dGTP	20	20	1	20	
50 mM Tris, 1mM EDTA pH8	5	5	5	5	

0.5mM dATP in 5 mM TRIS, 0.1 mM EDTA was made for the "chase".

Preparation of ddNTP solutions: The optimal concentration was determined for each new batch of ddNTPs and enzyme. The optimal concentrations, were near the top of the following range: ddTTP 0.5-2.5mM; ddCTP, ddGTP 0.3-0.7mM; ddATP 0.1-1mM. Having optimised the concentrations of the ddNTPs, stocks were made of "A mix", "T mix", etc., by mixing equal volumes of the dNTP' and ddNTP mixes. This was dispensed into 10 µl aliquots and kept at 20°C. Unnecessary freezing and thawing of ddNTP stocks was avoided.

The Dideoxy Reactions:

1-5 µl of α-^{32}P dATP x the no. of reactions, 2 µl of "A mix" x the no. of reactions, and 0.2 µl of Klenow x the no. of reactions, were added to a tube marked "A". The reagents were mixed together well, and kept on ice. Using Gilson tips, the same procedure was carried out for G,T, and C. The hybridized clones (2µl) were dispensed into tubes marked with the name of the clone and A,G,T or C.

At t=0', 2 µl ^{32}P/dNTP'/ddNTP/Klenow mix was added to 2 µl of hybridized clone. It was mixed, and incubated for 15 mins. at room temperature.

At t=15 mins. 1 µl of 0.5 mM dATP was added to each tube, mixed in, and incubation continued for 15 mins. (the "chase").

At t=30 mins. 4 µl of formamide mix (0.03% XCFF,BPB; 20 mM EDTA in deionized formamide) was added, and the samples boiled for 2 mins. 2 µl was applied to sequencing gels (II.12.2).

II.15 Ouchterlony Double Diffusion Plates

The following method was applied:

Buffer: 0.1M potassium phosphate (pH8), 41.3ml KH_2PO_4 13.6g/L, 58.7 ml K_2HPO_4 17.4 g/L.

Gel: 1 g agarose, 5 ml glycerol, 95 ml Buffer.

The mixture was boiled, cooled, and poured onto clean, ethanol-rinsed glass plates. Wells 4 mm in diameter were punched.

Diffusion: After the samples had been applied to the wells, diffusion in a damp chamber for 24-48 h at room temperature was allowed.

Pressing, Washing and Drying: The plates were pressed as follows to remove non-precipitated proteins and achieve a low background.

The plate was placed on filter paper, and distilled water poured on so that all the wells were filled. Approx. 5 layers of filter paper were placed on the gel, followed by a glass plate which was weighted down. After 5 mins., the upper filters were replaced and the washing and pressing repeated. The gel was washed in 0.1M NaCl for 15 mins., then pressed for 10 mins as above. This washing and pressing was repeated 2-3 times. Finally, the plate was placed in distilled water for 15 mins., and dried in a stream of hot air.

Staining:

Stain Solution: Coomassie Brilliant Blue R 250 5 g, ethanol 96%, glacial acetic acid 100 ml, distilled water 450 ml. The reagents were mixed overnight, then filtered.

The gels were stained for 5 mins., then destained 3x 10 mins. with Destain Solution, which is as Stain Solution, but with the omission of Coomassie Blue. The destained plates were dried in a stream of hot air.

Samples:

15 µl of antibody preparation, or protease-containing supernatants (which had been spun twice), and diluted in buffer, when necessary.

Supernatants were treated with PMSF (2.5 µl of 1% PMSF in iso-propanol per 100 µl supernatant) prior to dilution in buffer.

Positive controls: 0.1% ($1 mg ml^{-1}$) " Alcalase" or subtilisin Novo.

II.16 Removal of 5'-Phosphate with Calf Intestinal Alkaline Phosphatase (CIP) (Boehringer).

was performed as follows: DNA, 0.5 µl CIP (10 units), 50 µl 50 mM TRIS pH8, 1mM EDTA, were incubated for 1 h at 37°C in a final volume of 100 µl. The DNA was then extracted with phenol, and recovered by precipitation.

II.17 DNA Extraction from Gels

1) DNA fragments were electrophoresed into DEAE-paper (Schleicher & Schull NA 45).
2) Paper carrying the fragment was rinsed briefly with cold H_2O.
3) The paper was incubated in 400 µl 1.5 M NaCl-TE at 65°C for 30 mins.
4) The buffer was then phenol-extracted, $CHCl_3$ extracted and DNA precipitated with EtOH.

II.18 cDNA Synthesis (One Strand)

Single-stranded cDNA was synthesized on RNA, using a synthetic oligomeric primer, by the following method (Esper Boel).
4x First Strand Buffer: 20 µl 1M Tris-HCl, pH7.5; 28 µl 2M KCl; 4 µl 1M $MgCl_2$; 2 µl 0.2M DTT; 10 µl each dCTP, dGTP 20mM ; DEP-water to 0.1ml.

Approx. 30 µg RNA was added to 50 pmol kinased oligomer. 5 µl 50 mM EDTA was added, and volume made up to 18 µl with DEP-water. The mixture was incubated at 75°C for 3 mins., then put on ice. 2.5 µl RNasin, 12.5 µl 4x First Strand Buffer, 2.5 µl each 10mM dATP and dTTP, and 3 µl Reverse Transcriptase were added, with water to 48 µl.The mixture was incubated for 1h at 42°C. The reaction was stopped by the addition of 2 µl 250 mM EDTA, and 50 µl TE. The mixture was phenol extracted once, then chloroform extracted once. 100 µl 4M NH_4-acetate and 500 µl EtOH were added. After 15 mins. in a dry ice/EtOH bath, followed by 10 mins. at room temperature, the nucleic acids were pelleted, and the pellet lyophilysed.

Chapter III: Cloning and Characterization of the Gene for Subtilisin Carlsberg

III.1 Cloning Strategy

For the reasons set out below it was decided to use an approach that did not rely on expression for detection of the cloned gene:

(A) *E. coli* was chosen as the cloning host, and

(B) the aim was to clone only a part of the gene at a time, in the first instance, screening clones with oligomeric hybridization probes. Overlapping clone(s) would then be isolated in a second stage to give all the desired sequences.

(A) The following considerations were taken into account regarding the choice of *E. coli* as the cloning host:

B. subtilis is not an ideal host for "shotgun cloning". It is clear from the work of Trautner and Dubnau that in the absence of appreciable homology either with the chromosome (Canosi et al, 1981), or with an already resident plasmid ("Plasmid Rescue System", Contente and Dubnau 1979; Gryczan et al, 1980), only multimers of incoming plasmids are active in transformation of competent *B. subtilis* cells (Canosi et al, 1978). Monomers are completely inactive (Mottes et al, 1979).

Although this is not the case in the protoplast transformation system (Mottes et al, 1979; de Vos and Venema, 1981), where monomeric species *are* able to transform, the procedure is cumbersome and in our hands had given equivocal results.

Furthermore, as was clear from the Introduction (Chapter I), well-characterized protease-deficient mutant *B. subtilis* cloning hosts were lacking.

Most available mutants of *Bacillus* with altered protease production were pleiotropic, leading to deficiencies in the production of other exoenzymes as well. The mutations involved pro-

bably lie at loci other than the protease-coding sequences. The likelihood of detecting a clone based on increased exoprotease activity in such a host was thought to be uncertain, as a factor necessary, either for expression, or export, might have been affected. Conversely, if increased activity were obtained in a cloning attempt, it could well be due to the cloning of one of the many regulatory loci. While such clones would be very interesting to examine, they would complicate analysis of results our experiments, where the primary goal was to clone the extracellular serine protease gene.

We were not confident of being able to detect increased exoprotease activity resulting from the presence of the cloned subtilisin gene in a wild-type B. subtilis host as i) the endogenous protease background is high, ii) the cloned gene, being present in several copies might titrate out positive trans-acting regulatory factors required for activity and iii) may well require species-specific factors for expression, as it would be derived from B. licheniformis.

B) Regarding the use of oligomeric probes:
Bruce Wallace had shown that in hybridization of a mixture of short oligomers to DNA the temperature conditions can be chosen so that hybridization of an oligomer with perfect homology to the DNA can be distinguished from less specific annealing. Only a single base mismatch leads to a significant decrease in the thermal stability of the hybrid (Wallace et al, 1979).

The use of such a mixture of oligomeric probes in order to identify subtilisin-specific fragments was open to us, as the amino acid sequence of subtilisin Carlsberg was available. Thus possible coding sequences could be deduced, and a suitable region chosen on which to base the sequence of the oligomeric probes to be synthesized. This region would ideally offer a relatively low number of coding combinations, and cover a stretch unique in, and specific to, the Carlsberg subtilisin enzyme.

The rationale for deciding to clone only part of the gene at a time is as follows:

Many Bacillus genes are phenotypically expressed in E. coli (Chi et al, 1978), though the reverse is seldom true (Goldfarb et al, 1981; I.3 & I.7).

It was therefore decided to clone a specific fragment that was shorter than that one could expect to carry the entire coding sequence (as estimated from the exoenzyme sequence and what was known of Bacillus signal sequences). In this way one could obviate possibly lethal over-expression of protease in the E. coli cloning host .

The primary translation product of most secreted proteins has an N-terminal extension (signal sequence) that is usually cleaved off during the secretory process (Silhavy et al 1983; I.9.1). Many degradative enzymes such as proteases, are synthesized in the form of an inactive "zymogen" precursor that must be processed to yield the active molecule (Neurath 1984; I.8). Although we did not know at the time whether this also held true for the Bacillus exoproteases, if phenotypic expression of a cloned exoprotease gene in an E. coli host were achieved, it was not certain whether the resulting protein would, or could, be transported from the cytoplasm by the host secretory machinery. We reasoned that a heterologous protease trapped in the cytoplasm or periplasm, and possibly in an active form, was unlikely to be beneficial to the host, and might well be lethal.

As will be described and discussed below, some of this circumspection was unjustified, but in the particular case of the subtilisin Carlsberg gene, the cloning strategy has indeed been vindicated.

Several groups have, since this work was begun, successfully shotgun-cloned, and directly detected cloned exoprotease genes

in Bacillus wild-type hosts, by virtue of exoprotease overproduction. Using "plasmid rescue" and protoplast transformation systems respectively, the subtilisin and neutral protease genes of B. amyloliquefaciens (Vasantha et al, 1984) and the neutral protease gene of an unidentified Bacillus species were cloned (Corfield et al, 1984).

Low protease-producing nitrosoguanidine-generated Bacillus mutants were however needed for the cloning of the B. subtilis subtilisin (Stahl and Ferrari, 1984) and thermostable neutral protease of B. stearothermophilus (Fujii et al, 1983).

An indirect approach based on cloning of specific fragments in E. coli prior to transferring the clone to B. subtilis for expression was used to clone the B. subtilis subtilisin (this thesis) and neutral protease (Yang et al, 1984).

The strategy outlined above for the B. licheniformis gene was successful in so far as overlapping E. coli clones yielded the entire coding sequence (this thesis). As is described below, only weak expression was achieved in B. subtilis, unless another promoter was provided. To my knowledge, no group seeking this gene by screening for activity in B. subtilis has been successful, despite intensive efforts in many laboratories. My own attempts to clone this gene in B. subtilis have so far yielded only defective cloned genes. To date there are no published reports of the cloning of this gene other than that resulting from this project (Appendix 1). Personally communicated reports of the cloning of the gene have been made by two other groups (Genentech, Gist-Brocades). In both instances, this was achieved only after many fruitless attempts. In the one case, a small segment of the gene was synthesized instead of cloned. Possible reasons for the difficulties experienced, as well as other approaches that could give the desired results are discussed.

III.2 Selection of a Strain that Secretes Subtilisin Carlsberg

As discussed in Chapter I, there has been some confusion regarding the classification of subtilisin secretors. Accordingly, before cloning work was commenced, the secreted proteases of several strains of B. subtilis and B. licheniformis were tested immunologically.

Anti-Carlsberg serum was kindly provided by Novo Industri A/S. It is an antiserum raised against the major component of the commercial preparation "Alcalase".

Figure 4 represents an immune double diffusion (performed as described in II.15) illustrating the strong cross-reaction of the stationary phase supernatant of one B. licheniformis strain, MFJ29 (well 5) with the antiserum, and the immunological identity of the cross-reactant with Alcalase (well 4). A B. subtilis subtilisin overproducer, MFJ 135 (well 2), on the other hand, shows no trace of cross-reaction.

This chapter describes the cloning of the subtilisin gene of MFJ29.

III.3 Cloning a Portion of the Coding Sequence

III.3.1 Synthesis of Oligomeric Probes

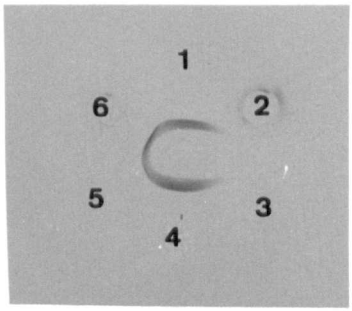
An oligomer of 14 bases had been shown to be the shortest that would bind reliably to the complementary sequence, even allowing for one mismatch (Wallace et al, 1979, 1981).

Using the published subtilisin Carlsberg amino acid sequence (Smith et al, 1968), all possible coding sequences were generated by computer. Accordingly, the coding sequence was searched for a 14-base stretch for which fewest possible permutations exist. The region arrived at, viz. residues 133-137, not unexpectedly, encoded a Methionine residue. Sixteen 14-mers were synthesized pairwise, as shown in Figure 5a, and the T_D of the expected hybrids was estimated according to the rule ($4^{o}C$ x no. of GC pairs) + ($2^{o}C$ x no. of AT pairs).

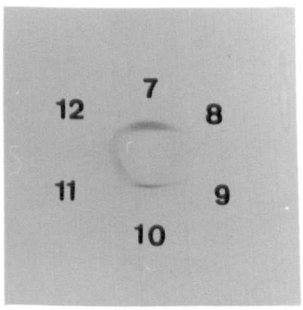
FIGURE 4: IMMUNODIFFUSION OF CULTURE SUPERNATANTS AND ANTI-SUBTILISIN

Samples were 4 day BP-X culture supernatants. The Ouchterlony plates were prepared as described in Materials and Methods.

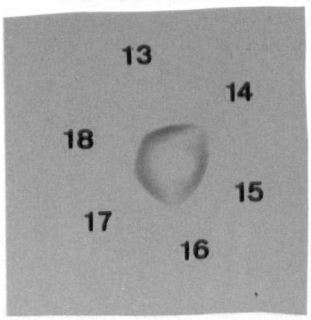
(A) central well: anti-Carlsberg serum diluted 1:4
well 1: Alcalase, 2: MFJ135, 3: MFJ198, 4: Alcalase, 5: MFJ29 diluted 1:2, 6: MFJ213 diluted 1:4.



(B) central well: anti-Carlsberg diluted 1:4
well 7: Alcalase, 8: MFJ210, 9: MFJ135, 10: MFJ 246, 11: MFJ213, 12: MFJ246



(C) central well: anti-NOVO serum diluted 2:3
well 13: subtilisin NOVO, 14: MFJ157, 15: subtilisin NOVO, 16: MFJ93, 17: subtilisin NOVO, 18: MFJ135 diluted 1:2, 19: subtilisin NOVO, 20: MFJ157
21 to 24: doubling dilutions of MFJ157 sample



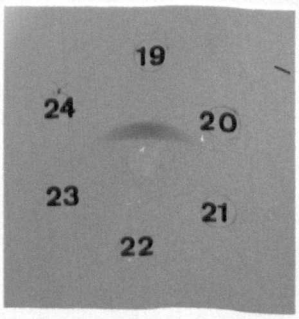


FIG.5: CHROMOSOMAL RESTRICTION MAPPING WITH OLIGOMERIC PROBES

(a) **Oligomeric Probes**

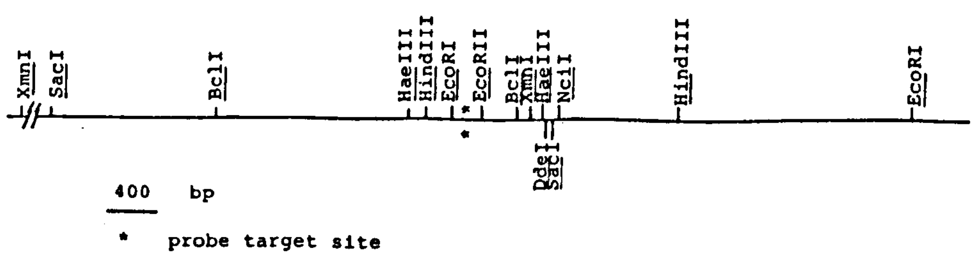
codon no.		Ala Met Lys Gln Ala 133 134 135 136 137	T_D (°C)
oligomer pair	1	GCC ATG AA$^{A}_{G}$ CAA GC	42 44
	2	GCA ATG AA$^{A}_{G}$ CAA GC	40 42
	3	GCT ATG AA$^{A}_{G}$ CAA GC	40 42
	4	GCG ATG AA$^{A}_{G}$ CAA GC	42 44
	5	GCC ATG AA$^{A}_{G}$ CAG GC	44 46
	6	GCA ATG AA$^{A}_{G}$ CAG GC	42 44
	7	GCT ATG AA$^{A}_{G}$ CAG GC	42 44
	8	GCG ATG AA$^{A}_{G}$ CAG GC	44 46

(b) **Results of Chromosomal Restriction Mapping**

Restriction enzyme		Fragment size (kb)
	XmnI BamHI SalI SphI ClaI	very large
	SacI EcoRI BclI HindIII HaeIII	5.0 3.7 2.4 1.9 0.9
HindIII &	EcoRI NciI SacI DdeI HaeIII XmnI BclI EcoRII	1.7 1.05 1.0 0.95 0.8 0.75 0.65 0.45

The HindIII fragment was not cut by the remaining enzymes giving very large fragments on their own or by EcoRV, SmaI, AvaI, NaeI, NruI.

(c) **Preliminary Restriction Map of Subtilisin Region**



III.3.2 Chromosomal Mapping

III.3.2.1 Finding the Correct Hybridization Conditions

The oligomers were kinased (as described in II.7.3), and used to probe restriction digests of MFJ29 chromosomal DNA (prepared as described in II.3.1.2) by Southern blotting (as described in II.8.1).

Duplicate strips were hybridized to probe and washed at temperature intervals of 5°C, over the range 30°C to 50° C. At 40°C only 1 band was seen in each track after autoradiography. This hybridization was abolished at 45°C, and higher temperatures. 42°C was found to give the most satisfactory results. By hybridization with individual pairs of probes, pair no.8 was identified as being responsible for the specific hybridization, and was used thereafter.

III.3.2.2 Preliminary Chromosomal Map

Using the probe, the area around the target hybridization site was mapped as shown in Figure 5 b&c. Once the proximity of the HindIII and EcoRI sites had been established, other enzymes which shortened the specific HindIII fragment were sought. Computer predictions from all the possible coding sequences revealed only one possible XmnI site, and its distance from possible EcoRI and HindIII sites was consistent with the map obtained. The same applied to the distance from EcoRI to BclI (Figure 5c).

The shortest subtilisin fragment that could be expected to be expressed and encode an active product was of the order of 1 kb long (structural gene = 274 codons + anticipated signal sequence of c.30 codons plus regulatory sequences of c.100 bp). In accordance with the rationale set out in III.1, it was decided to clone the HindIII-HaeIII 0.9 kb fragment (Figure 5).

III.3.3 Cloning the HindIII-HaeIII 0.9 kb Fragment

About 100 µg MFJ29 chromosomal DNA was size-fractionated on agarose gels after HindIII digestion. The recovered DNA of c.2 kb

(isolated as described in II.17) was digested with HaeIII, and again fractionated in the same way. The estimated yield of fragments in the "0.9 kb" fraction was less than 1 µg.

These were ligated with vector pAT153, a high copy derivative of pBR322 (it lacks a HaeII fragment between 1647 and 2352 bp), which had been cut with BalI and HindIII. The BalI recognition site includes that for HaeIII. Competent cells of MFJ28 (II.4.6) were transformed, and the resulting 92 Ap^R colonies were screened by colony hybridization (II.8.2), using probe pair no. 8. Figure 6 shows an autoradiograph of a Southern blot (II.8.1) of plasmid extracted from the 2 positive clones identified in this way. The closed circular plasmid DNA was efficiently transferred to the filter by irradiating the gel prior to denaturation under UV lamps (254 nm) for 20 mins.

III.3.3.1 Restriction Analysis of Probe-Positive Clones

From the known sequence of vector pAT153, the size of its HaeIII digestion fragments was deduced. There are 5 ranging from 267 to 587 bp, in addition to others too small to visualize on agarose gels. The cloned HindIII-HaeIII subtilisin fragment should appear intact at about 900 bp in a HaeIII digest, where the proximal HaeIII end is provided by that at 4344 bp in the pBR322 sequence.

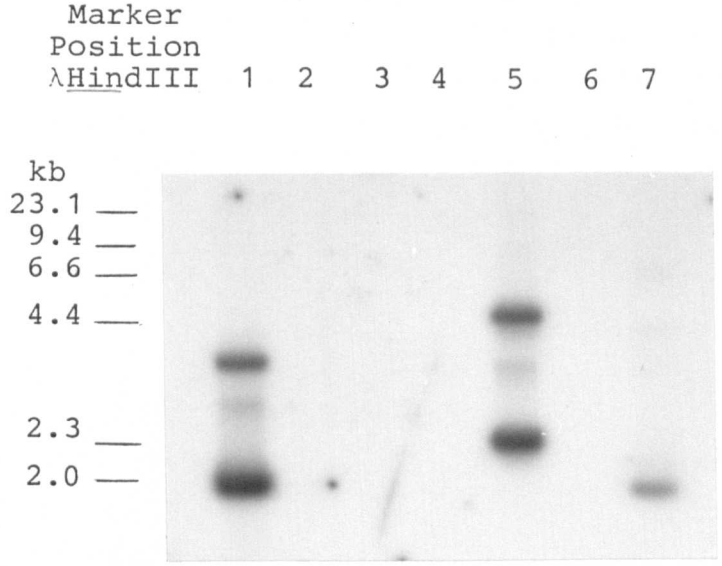
Comparison of HaeIII digests of pAT153 and the clones confirmed this expectation.

The 2 probe-positive clones were identical except that one had, in addition, an extra cloned HaeIII fragment of approx.1 kb, and it was not examined further.

The other plasmid, pMJ1, was analysed with reference to the chromosomal restriction map (Figure 5c), with enzymes for which sites are found in the appropriate area, and with those for which sites do not occur. For cleavage with EclI, the plasmid was transferred to a dam$^-$ strain, MFJ92. A number of other en-

FIGURE 6

Autoradiograph of Southern blot of plasmid DNA extracted from colonies hybridising to subtilisin-specific 14-mer probe.



lanes 2,3,4,6: DNA from negative clones picked at random.
lane 1: DNA from clone containing pMJ1.
lane 5: DNA from second positive clone.
lane 7: MFJ29 chromosomal DNA, HindIII digest.

zymes for which computer predictions suggested only one possible site in the coding sequence were also tested on pMJ1. A few digests are shown in Fig. 7.

Surprisingly, pMJ1 was linearized neither by HindIII (used in cloning), nor ClaI, for which a site lies just upstream from the unique HindIII site on pAT153. The map of pMJ1 arrived at is shown in Figure 8a. It is entirely consistent with the chromosomal data (Figure 5c), but for the "missing" HindIII site.

III.3.3.2 Confirmation that the Cloned Fragment is Part of Subtilisin Carlsberg Gene by Sequence Analysis

The EcoRI-BalI fragment of approx. 650 bp was isolated from pMJ1 and used directly for sequencing . The two 14-mers comprising probe pair no.8 (Figure 5a) were synthesized individually, and used as primers in Sanger sequencing reactions (II.14.2). They were used at a 20-fold molar excess over the double-stranded template.

With one primer, 5'-GCGATGAAACAGGC-3', a readable sequence of 35 bases was obtained, and was consistent with the coding sequence for codons +145 to 155 of the published amino-acid sequence of subtilisin Carlsberg (Smith et al, 1968). Further sequencing was thus undertaken (see III.6, Figure 14).

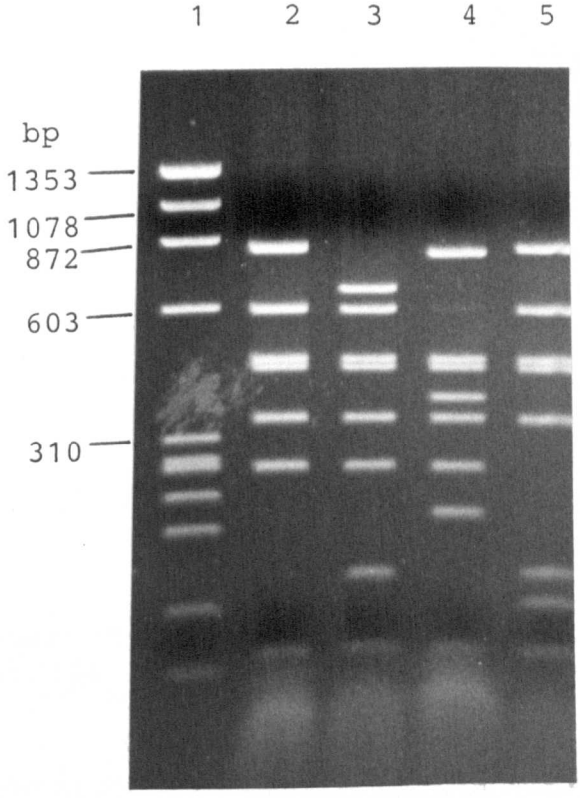
To resolve the question of the "missing" HindIII site, the small EcoRI fragment, shown at about 12 o'clock in Figure 8, was recloned in M13 (II.14.2) and sequenced by the Sanger method.

The sequence revealed that the subtilisin sequence at the HindIII site is identical to a nearby vector sequence, such that a direct tandem 7 bp repeat must have been generated in the primary ligation product. Figure 8b shows the probable structure of this recombinant plasmid about the HindIII cloning site, as well as the actual sequence found in pMJ1.

From these two short stretches of sequence, the cloned insert

FIGURE 7: RESTRICTION ANALYSIS OF pMJ1

Restriction digests were analysed on 1.5% agarose gels.



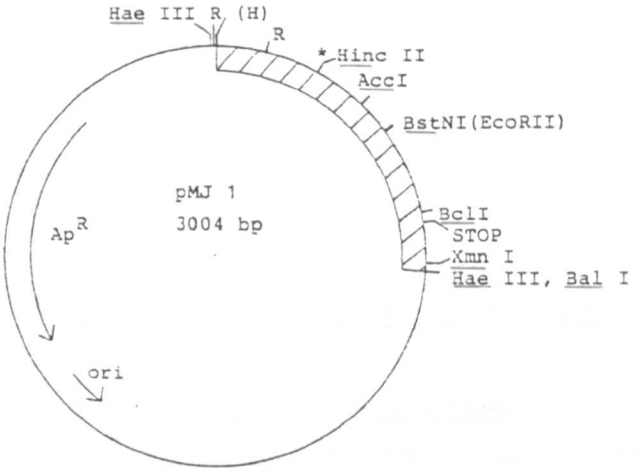
lane 1: ØX174/HaeIII
lane 2: pMJ1/HaeIII
lane 3: pMJ1/HaeIII + EcoRI
lane 4: pMJ1/HaeIII + XmnI
lane 5: pMJ1/HaeIII

FIGURE 8

Structural Analysis of pMJ1

(a)

Restriction Map of pMJ1

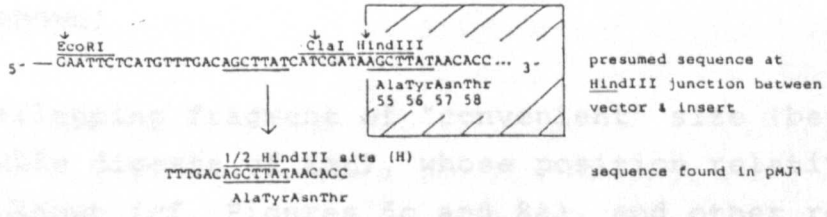


——	pAT153 sequences
▨	B. licheniformis DNA
*	oligomeric probe target site
STOP	translation termination codon
R	EcoRI
(H)	Spoilt vector-insert HindIII junction

Note: Apart from the insert-proximal EcoRI, HaeIII sites, no vector restriction sites are shown.

(b)

Putative Direct Tandem Repeats Leading to Deletion & Loss of HindIII Cloning Site from pMJ1



overlined	restriction enzyme recognition sites
underlined	direct 7bp repeat
hatched area	subtilisin sequences; codon number in mature enzyme indicated
small vertical arrows	enzyme cleavage site

could be oriented: The (missing) HindIII site lies in codon +54, the EcoRI site in +96. It was later established by sequencing that the BclI site is very near the end of the coding region in codon +266, as anticipated from computer predictions (Figure 14).

It was therefore clear that the entire 3' terminus of the gene had been cloned.

III.4 Cloning a Fragment Overlapping 5' end of Coding Sequence

III.4.1 Identification of a Suitable Fragment

An EcoRI digest of chromosomal DNA of MFJ29 was probed by Southern blotting to determine the size of the upstream EcoRI fragment, using as probes nick-translated pMJ1 (II.7.1) and the pMJ1/EcoRI large fragment only (Figure 8a), and comparing the pattern obtained with each on duplicate blots.

Only one EcoRI band of c.3.7 kb hybridized in each case, whereas detection of a second was expected, by homology to the 120 bp of subtilisin sequences (on pMJ1) that lie upstream from the cloned EcoRI site (in codon +96) (Figure 8a).

This was interpreted as implying the presence of another chromosomal EcoRI site so close to the one in codon +96, that the resultant fragment could not bind sufficiently well to the filters, and so escaped detection in blotting experiments. It was however later established, using a larger upstream probe, that the up- and downstream EcoRI fragments are of precisely the same size (not shown).

To find a 5' overlapping fragment of "convenient" size (between 1 and 2 kb), double digests of XmnI, whose position relative to the gene was known (cf. Figures 5c and 8a), and other restriction enzymes were analysed by Southern blotting. Enzymes tested were those which could readily be used for cloning in available vectors. Single digests served as controls. The enz-

ymes included BamHI, SalI, BglI, SmaI, BalI, PstI, SacI, XbaI, StuI & EcoRV.

The sites lying closest to the gene were BclI and StuI. These lie resp. 1.9 and 0.7 kb upstream from EcoRI (codon +96) (see Figure 11).

BclI thus seemed to be the only site useful for the purpose of splicing a new overlapping clone to the pMJ1 sequences.

In the first instance, therefore, it was decided to clone the 2.4 kb BclI fragment: it was large enough to carry all conceivable 5' cis-acting regulatory sequences, and unlike StuI had "sticky" ends, making ligation and cloning more efficient. Moreover, subtilisin-specific fragments could be enriched by size-fractionation prior to cloning, thereby reducing the number of clones that have to be screened. This is not the case with the alternative to cloning specific fragments viz. using partially digested donor fragments. Here one also risks cloning non-contiguous fragments simultaneously.

III.4.2 Construction of Cloning Vectors for BclI Fragment

Problems associated with the cloning of this fragment might immediately be anticipated, should expression be possible in the cloning host.

Only the 8 C-terminal codons of the protein, as well as the transcriptional, and translational, termination signals are missing. Should the promoter have powerful activity in the host, clonal stability might well be compromised by the resulting high level of transcription per se (see I.10.3.1). Moreover, the translation product (an almost complete subtilisin molecule), might well have a deleterious proteolytic activity. Inspection of the crystal structure (Wright et al, 1969) shows that the C-terminal residues occur as a tail well away from the active site. If translation should extend beyond the BclI site in codon + 266 in a plasmid clone, an active protease molecule

is not inconceivable, provided that the resulting fused "nonsense" C-terminal tail does not interfere with protein folding.

In an attempt to overcome these potential difficulties, 2 plasmids were constructed, in which a BclI cloning site, translational stop codon(s) and a transcriptional terminator were provided.

The construction scheme is shown in Figure 9. An intermediate, pMJ3, was made by incorporation of a synthetic linker KFN 19/20 between the EcoRI and HindIII sites of pBR322. This linker carries a BclI site and an umber codon in all 3 reading frames.

In pMJ19, the transcriptional terminators were provided by the E. coli 5S RNA gene from ptrc160, whose structure is not shown.

In pMJ7, the cloned subtilisin sequences themselves (from BclI downstream) were used. It has been established that this short BclI-HaeIII segment of pMJ1 (Figure 8a) does indeed carry the transcriptional terminator for the gene (see III.8.3).

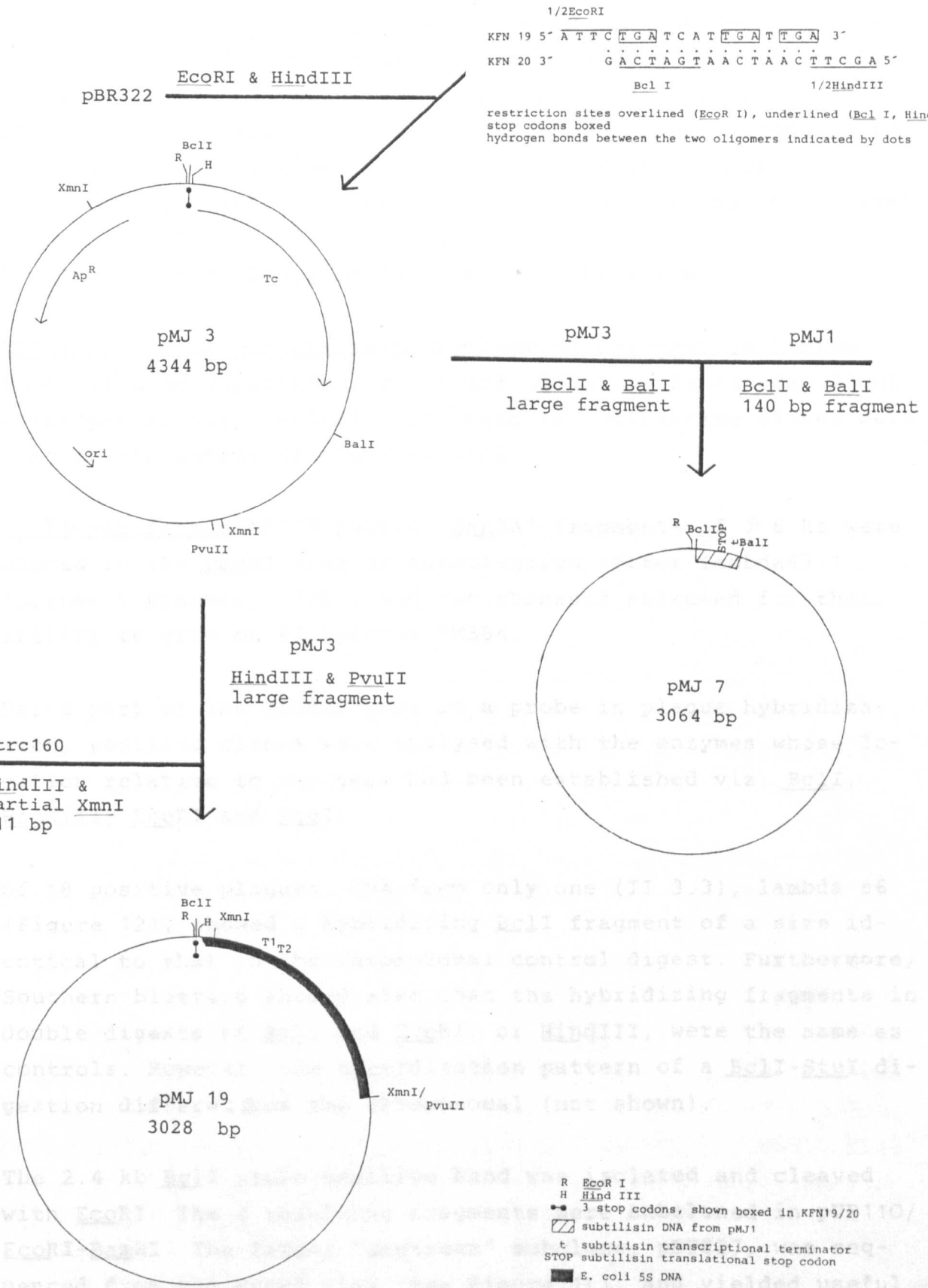
These plasmids were utilized in attempts to clone the BclI fragment in MFJ92. None of the colonies which were positive in colony hybridization, using part of the cloned gene as probe, contained recombinant plasmid with the desired subtilisin BclI fragment, and were not investigated further.

In experiments carried out by Jan-Ingmar Flock, attempts to clone the same BclI fragment, or one larger, in pUB110, in a B. subtilis host were similarly unsuccessful, whether starting with size-fractionated BclI DNA, or with partial Sau3A1 digests. As in the experiments described above, restriction analysis of hybridization-positive plasmids did not conform to expectations. These clones were not analysed further.

Two other approaches were thus pursued. Since the use of transcriptional terminators in E. coli had not proved useful, the

FIGURE 9

CONSTRUCTION OF CLONING VECTORS FOR THE BclI FRAGMENT



application of a low copy number plasmid cloning vector was explored in an effort to minimize unfavourable effects. Cloning in lambda phage was also attempted. Here host cell survival after eventual expression of "undesirable" foreign products encoded by the clone might not be quite so critical for clonal propagation, since the host is killed anyway during the "burst".

Results of these 2 approaches are detailed below.

III.4.3 Cloning an Upstream Overlapping Fragment in Lambda

These cloning experiments were carried out by Jan-Ingmar Flock. After preliminary analysis, potentially interesting clones were sent to the author for further work.

B. licheniformis MFJ29 partial Sau3A1 fragments of $\geq$ 6 kb were cloned in the BamHI arms of substitution vector lambda47.1 (Loenen & Brammar, 1980), and recombinants selected for their ability to grow on P2 lysogen NM364.

Using part of the cloned gene as a probe in plaque hybridization, positive clones were analysed with the enzymes whose location relative to the gene had been established viz. BclI, HindIII, EcoRI and StuI.

Of 18 positive plaques, DNA from only one (II.3.3), lambda s6 (Figure 12), showed a hybridizing BclI fragment of a size identical to that in the chromosomal control digest. Furthermore, Southern blotting showed also that the hybridizing fragments in double digests of BclI and EcoRI, or HindIII, were the same as controls. However, the hybridization pattern of a BclI-StuI digestion differed from the chromosomal (not shown).

The 2.4 kb BclI probe-positive band was isolated and cleaved with EcoRI. The 2 resulting fragments were subcloned in pUB110/ EcoRI-BamHI. The larger "upstream" subclone, pBF827, was sequenced from the EcoRI site (see Figure 13), and yielded useful results.

It was clear that there were 3 HindIII sites within a 60 bp stretch just upstream from EcoRI. The site that had previously been mapped (Figure 5; in codon +54, and missing in pMJ1) was the most 3' of them. Further sequencing, not detailed here, revealed that recognizable subtilisin sequences extended upstream as far as codon -11. There were a further 15 codons between these sequences and the lambda BamHI cloning site/Sau3A1 junction (all in frame). The origin of these sequences is not known. They were not connected to the subtilisin gene by a Sau3A1 recognition site, and it is therefore fair to conclude that a deletion involving B. licheniformis DNA had occurred.

The hybridizing BclI 2.4 kb lambda s6 fragment, though coincidentally of the same size as the B. licheniformis chromosomal subtilisin fragment, was in fact a hybrid of lambda and bacterial sequences.This could not have been ascertained simply by calculation of the sizes of the BclI fragments expected. There are several closely-spaced BclI sites about the BamHI sites on lambda that were used to create the cloning site in 47.1. This 'phage differs moreover in many uncharacterized positions from the lambda reference virus.

Only one clone, lambda s7, had a StuI site in the correct position ie. identical to the chromosomal control, relative to the EcoRI and BclI sites.

Lambda s7 was retained for further analysis.

Thus, despite the capacity of lambda 47.1 to incorporate large fragments (its cloning capacity ranges from 4.7 to 19.6 kb), only one of the 18 positives carried a site (StuI), which lies not more than 400 bp upstream from the deduced mature subtilisin coding sequences.

III.4.4 Cloning BclI Fragment on a Plasmid Under Low Copy Number Conditions

III.4.4.1 B. subtilis System

Although low copy number analogues of some Gram-positive cloning vectors are available eg. p900 (Figure 20), in my hands colony hybridization was not sufficiently sensitive to detect them in B. subtilis hosts.

The simple technique utilizing Whatman filters, which was used for E. coli plasmid hosts (II.8.2) was not applicable to B. subtilis (not shown).

Nylon filters (II.8.3) gave satisfactory results in a model system with a "normal" copy number plasmid, but with its low copy analogue the signal obtained with labeled probe was not sufficiently strong.

The method described in Maas (1983), which involves steaming the colonies to improve the efficiency of lysis was not useful in this instance either.

Attention thus focused on E. coli hosts once more.

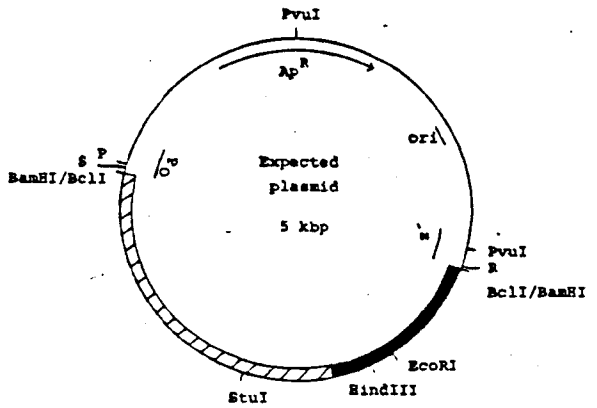
II.4.4.2 E. coli System

A strain, MFJ55, was made available by W. Tacon. It has an uncharacterized lesion that impairs ColE1-based plasmid replication, such that < 5 copies per chromosome are produced by plasmids such as pAT153, which are normally present at many times that number.

Plasmid pUR222 (Figure 10a) was suggested as cloning vector (Ruther et al, 1981). The unique restriction sites about the BamHI cloning site located in the polylinker, (PstI - SalI-BamHI - EcoRI), situated N-terminally in lacz, would assist both isolation and restriction analysis of the cloned insert, and allow rapid sequencing (Ruther et al, 1981). As MFJ55 has the chromosomal deletion lacX74 = Δ(lacI POZY) one of the

FIGURE 10

(a) Structure of Plasmid Expected From BclI Cloning



Only one of the 2 possible orientations of the BclI insert is shown

pUR222 vector sequences

B. licheniformis subtilisin (mature enzyme)

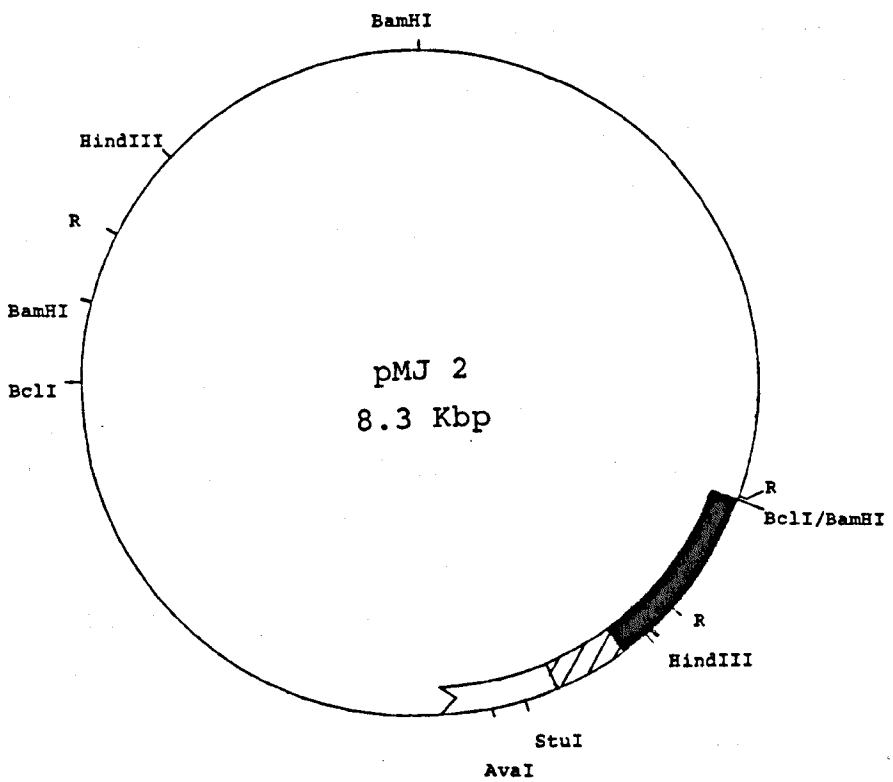
sequences upstream of mature coding region

PO lac operon promoter, operator

z' lac z from 5th codon

P,S,R PstI, SalI, EcoRI cloning sites in polylinker

(b) Structure of pMJ2



NOTE: Only sites that could be placed with certainty are indicated

B. licheniformis subtilisin coding region prepro mature enzyme

upstream sequences, 5'limit uncertain

major potential advantages of using a plasmid such as pUR222 viz. the ability to screen directly for recombinants (white colonies) on X-gal plates in an "α-complementing" host (II.14. 2.1) could not be exploited.

More pertinently as far as subtilisin cloning is concerned, the mutant strain lacks the lac repressor gene, which leaves the lac promoter on pUR222 constitutively active. Any cloned gene would come under its control. If lac transcription through the subtilisin gene were detrimental, it would make the cloning of the BclI fragment in only one orientation more likely. MFJ29 BclI fragments of approx. 2.4 kb were consequently cloned into pUR222/BamHI, which had been phosphatase-treated (II.16). There were an estimated 500 recombinants (50% recombinants) among the Ap^R transformants of MFJ55 that were screened, as judged by restriction analysis of plasmid from randomly picked clones.

Total plasmid was extracted from 21 pools of 50 colonies each, and tested in Southern blotting with nick-translated subtilisin DNA from pMJ1 as probe. pMJ1 and pUR222 served resp. as positive and negative hybridization controls. The plasmids were nicked by UV irradiation of the gel prior to denaturation and transfer, as described in III.3.3. Only one plasmid was positive. It was isolated from the pool and called pMJ2.

III.4.4.2.1 Structural Analysis of pMJ2

Figure 10a depicts the expected structure of pMJ2: a simple insertion of the chromosomal 2.4 kb BclI fragment into the unique BamHI cloning site of pUR222.

Figure 10b shows the restriction map of pMJ2, in so far as it could be ascertained.

With all enzymes tested, the total length of pMJ2 was c. 8.3 kb.

BamHI, SalI or PstI sites occur uniquely in the cloning polylinker of pUR222 (see Figure 10a). It had been established that these sites do not occur on the desired subtilisin fragment. Restriction analysis of pMJ2 revealed a minimum of 2 SalI sites, 2 BamHI sites and 6 PstI and EcoRI sites. The complexity of double digestions with these enzymes, which gave many small fragments, precluded the deduction of an unambiguous map for the entire plasmid. It was not possible either to locate the lac promoter-proximal part of the polylinker.

HindIII digestion however resulted in only 2 large fragments of c.3.7 and 4.5 kb. These were isolated individually. By a combination of restriction analysis, hybridization, and sequencing, the location of the subtilisin gene on the HindIII fragments was deduced to be as shown in Figure 10b. Note the cluster of HindIII sites within the gene coding sequence, as had been established from the sequence of lambda s6 (III.4.3). From the map of pMJ2 it can be seen that if the gene were expressed, it would form a fusion protein with LacZ, having 52 AAs encoded by the linker and lacZ (out of frame), and ending at an umber codon in codons 54-55 of the lacZ sequence (Kalnins et al, 1983).

Since there was about 3.3 kb of" extra" DNA in pMJ2, two immediately obvious possible explanations for these sequences suggested themselves: an "Insertion Sequence" (IS) could have transferred to the BclI subtilisin insert, or else two BclI fragments from the "2.4 kb" size-fraction might have been cloned simultaneously. An attempt was made to obtain some experimental support for either of these possibilities.

III. 4.4.2.2 Does pMJ2 Carry An IS Sequence?

So-called IS sequences, small transposable elements, have long been observed in E. coli, and have recently also been found in a Bacillus sp. (Mahillon et al, 1985). They can have polar pos-

itive or negative effects on the transcriptional activity of the sequences they interrupt.

The possibility of IS insertion needs always to be considered when cloning in E. coli, especially where activity of the clone may be detrimental to the cell (Saunders et al, 1984a; Rood et al, 1980). In such instances, transcriptional stop signals within the IS serve to counter the unfavourable activity.

Typically, IS insertion is detected when the same rearrangement of already well-characterized cloned material is found repeatedly upon recloning under circumstances where activity of the clone is elevated.

The cloning host, MFJ55, carried Tn10, and its right hand repeated sequence, IS10R, is capable of independent transposition (Roberts et al, 1985). With only one clone (pMJ2) it is rather difficult to draw more than tentative conclusions from the data.

Arguing against the incorporation of an IS on pMJ2 is the length of the "extra" DNA: 3.3 kb vs. the 750-1600 bp, typically found for most IS sequences identified to date (Calos & Miller, 1980; Kleckner, 1981).

"Composite elements"/transposons on the other hand, where a non-cryptic gene is bordered by IS sequences, can be of 2-3 kb eg. Tn681, Tn9, Tn903, which resp. encode enterotoxin, Cm^R and Km^R (Kleckner, 1981).

pMJ2 was checked for the presence of above-mentioned antibiotic-resistance markers, as well as Tet^R and Emy^R and found to be negative. MFJ55 was first cured of Tn10 by selecting Tet^S on special plates (II.10). An "improved" recipe (Maloy & Nunn, 1981) was tried, but no colonies grew up. On the original Bochner plates, 24 fast-growing colonies were picked after 24h, and after repurification all proved to be Tet^S. pMJ2 transformants of one isolate, MFJ233, did not grow in the presence of Emy,

Tet, Km or Cm . Enterotoxin production was not checked.

IS insertion is often associated with duplication or deletion of flanking markers (same refs as above). There is no evidence for duplication of probe-hybridizing sequences on pMJ2. Certainly the StuI site is unique.

However, as will be described below, some of the other enzymes that could be used to analyse the subtilisin sequences show strong site preference and thus give multiple hybridizing bands of larger size than the minimal fragment.

The inability to map the entire plasmid satisfactorily precluded a direct comparison of the restriction pattern with those of characterized IS sequences. Some rearrangement of sequences on pMJ2 had undoubtedly occcured (see next section). But whether this was due to an IS must remain an open question.

It appears that if any E. coli-derived DNA was involved in the rearrangement, it is no longer on pMJ2. Southern blotting of DNA from MFJ55 and MFJ233 did not reveal detectable homology to the plasmid.

III.4.4.2.3 Were 2 BclI Fragments Cloned Into the pUR222 BamHI Site? Instability of pMJ2 in dam⁻ Strain

Size-fractionated DNA is always slightly heterogeneous. A fragment of 3.3 kb could well have been in the "2.4 kb" BclI fraction used in cloning pMJ2.

Southern blotting of MFJ29 DNA indeed showed that pMJ2 hybridizes to two BclI fragments: the 2.4kb subtilisin fragment, and another of c.3kb. This larger band did not appear to be a partial digestion product as it did not hybridize to the 500bp StuI-HindIII fragment of pMJ2. The plasmid also hybridizes to additional, smaller fragments in an EcoRI digest, besides the up- and downstream 3.7kb subtilisin fragments. However, in view of the subtilisin gene's homology to other related genes, eg.

ISP-I (Fig. 2), these results are insufficient in themselves to draw a firm conclusion.

BclI/BamHI junctions are cleaved by neither enzyme. If indeed 2 inserts are present on pMJ2, there will be a BclI site at their junction. The ligation mixture had not been cleaved with BclI prior to transformation of MFJ55. This would have linearized recombinants with multiple inserts and hence eliminated their ability to transform.

As BclI is sensitive to dam methylation, pMJ2 was transformed into MFJ92 to expose such a postulated site, and render it cleavable.

Analysis of plasmid recovered from MFJ92 (as well as MC1061=the MFJ55 parent strain, and HB101) revealed that it had undergone extensive deletion to a size of 3-4 kb. Each isolate was different. Upon repeated retransformation into the low copy host MFJ55, or MFJ233 (=MFJ55 cured of Tn10) no alteration in the plasmid has been observed.

Evidently, on transformation into strains where the copy number of pMJ2 returns to a level typical of pUR222, the plasmid becomes unstable. It is clear too from these experiments that it cannot be lac promoter activity through the clone that accounts for the instability, since in HB101 the promoter will be repressed. Its transcriptional activity would hence be lower than in MFJ55.

The instability at high plasmid copy number was precisely the behaviour expected of a subtilisin BclI clone, after the numerous unsuccessful cloning experiments, and had indeed been the rationale behind the use of strain MFJ55 as host.

Had IS insertion and/or rearrangement taken place on pMJ2 to counteract any adverse effects the clone may have had on the host, the resulting cointegrate, or otherwise altered plasmid,

would be expected to be stable to a subsequent increase in plasmid copy number. It seemed therefore probable that the subtilisin promoter activity was intact on pMJ2.

Since pMJ2 could not be transferred without deletion to a *dam*$^-$ background, or to any which did not keep the plasmid copy number low, an attempt was made to P1-transduce MFJ55 to *dam*$^-$ (II.11) by a strategy suggested by N. Fiil viz. by co-transduction with *malT*. This locus is separated from *dam* by less than one minute on the *E. coli* chromosome (Bachmann & Low, 1980).

In brief, lambda^{R}Mal$^-$ MFJ233 mutants were selected directly on McConkey plates containing 1% maltose after infection with lambda$_{vir}$ (10^6-10^9 pfu per 10 µl indicator cells gave confluent plaques). Mucoid lambda-resistant colonies were ignored.The majority of the remaining lambda^R clones were red. Seventeen white (MalT$^-$) colonies, otherwise morphologically identical to the red ones, were reisolated repeatedly on the selective plates. Of 9 which showed no red revertants, isolate 7 showed the most feeble growth on minimal AB plates containing maltose, and was used as the transduction recipient.

A P1 lysate of the Mal$^+$*dam*$^-$ strain GM33 was prepared (II.11) and the recipient transduced with it (II.11). After 48h the vigorously growing Mal$^+$ transductants were picked from the fine background lawn. In all 28 were restreaked once on AB, and then several times on McConkey-maltose.

To find a *dam*$^-$Mal$^+$ co-transductant, *Mbo*I digestions of chromosomal DNA from the 28 Mal$^+$ colonies were inspected and compared with undigested and control *Hin*dIII digests. *Mbo*I is sensitive to *dam* methylation, as is *Bcl*I, but as it recognizes a 4 bp sequence, it gives smaller average fragment size, and makes digestion more obvious.

DNA from GM33 and MFJ233 served resp. as positive and negative *Mbo*I-sensitive controls. None of the samples from the transduc-

tants could be cleaved with MboI (or BclI), though they all were cleaved by HindIII.

The same experiment carried out on another E. coli strain had given 1/12 dam$^-$Mal$^+$ co-transductants (N. Fiil, personal communication). Selection against acquisition of the dam$^-$ allele may operate, since none of the 28 MFJ233 Mal$^+$ transductants were dam$^-$. dam$^-$ mutants have a pleiotropic phenotype similar to strains defective in DNA repair. The nature of the mutation in MFJ55 leading to the lower copy number of ColE1 plasmids is not known. It could, for example, involve the polA locus. DNA polymerase I is normally required for ColE1 replication (Davison, 1984). The enzyme also has a role in DNA repair (Lewin, 1974). Double mutants of dam, and recB, recC, which are required for repair, are not viable (Bale et al, 1979). The uncharacterized mutation in MFJ55 may well compromise the host, in combination with the dam allele, to such an extent that double mutants are either inviable or strongly selected against.This experiment was not repeated.

However, on later subcloning from pMJ2 into B. subtilis, (Fig. 26), a BclI site was revealed, and is shown in Figure 10b.

It is found about 1 kb further away from the StuI site than on the chromosomal map. Hence it could not represent the junction of the 2.4 kb subtilisin BclI fragment, and another such from B. licheniformis without some rearrangement having taken place.

The 2 clones pMJ2 and lambda s7 were compared and investigated further to determine
a) how far the chromosomal gene-proximal sequences extended upstream on both clones (III.5)
b) whether any promoter activity was present immediately adjacent to the clones (III.7) and
c) how far the transcriptional start site was located from the translational start codon (III.8).

III.5 Defining the 5' Limit of the Clones

III.5.1 Fine Restriction Mapping of the Subtilisin Locus

In order to facilitate analysis of the clones to hand, a more detailed restriction map of the chromosomal locus was essential.

Preliminary studies had not revealed any site useful for cloning purposes lying 5' to the gene between BclI (at least 1.6 kb from the gene) (see Figure 5c) and StuI, which is situated very near the start of the coding sequences (see Figures 13 and 14).

Using either the StuI-EcoRI 700 bp or HindIII-EcoRI 120 bp fragment of pMJ2 as hybridization probe, Southern blotting of MFJ29 DNA was carried out. Figure 11a lists the enzymes used to digest EcoRI-digested chromosomal DNA. Figure 11b depicts the map deduced. Note that for some enzymes only the closest site to the gene has been mapped: there are definitely others further upstream on the EcoRI fragment.

It is noteworthy that with the exception of DdeI, enzymes for which sites were found clustered between BclI and StuI have GC-rich recognition sequences (see Figure 11c), and suggest that the entire region is GC-rich. Enzymes of this type tend to show strong site preferences (New England Biolabs catalogue 1985/86, p. 111). NarI and BanI probably cleave at the same site. There are 3 other restriction enzymes whose recognition sites are composed entirely of G and C residues, viz. SmaI and XmaI (same sequence), and ApaI, but cleavage by these in the region was not seen.

The sequence data obtained from the clones (Figure 14) showed an extremely AT-rich region just upstream from the subtilisin gene, as one might expect from what is known of Gram-positive promoter regions (I.3.4, I.3.5.3, I.3.9).

It is thus probably significant that a GC-rich region lies just beyond. Such stretches are frequently found in intergenic regions in the Gram-positive Streptomyces (R. Hütter, personal

FIGURE 11

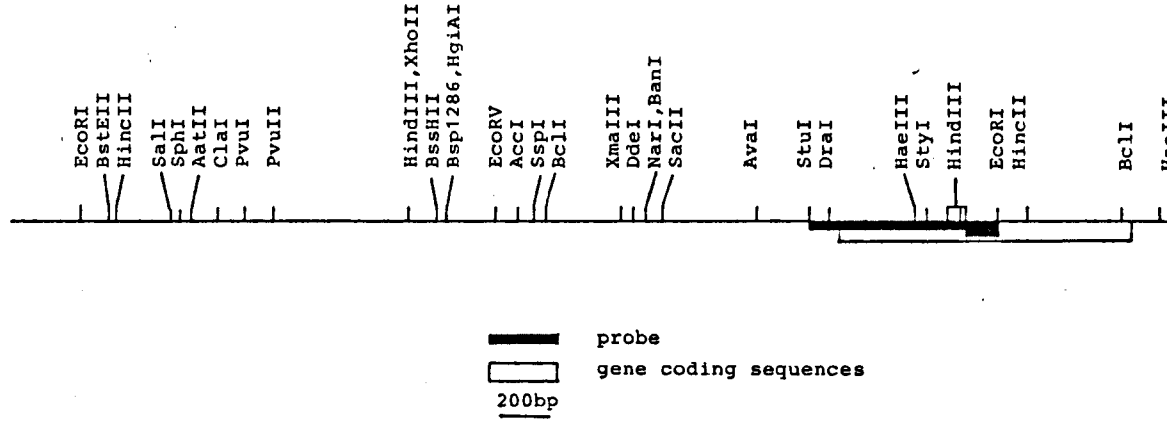
FINE CHROMOSOMAL MAPPING OF THE REGION UPSTREAM FROM THE SUBTILISIN CARLSBERG GENE IN MFJ29

(A) ENZYMES USED IN DOUBLE DIGESTION WITH EcoRI

AatII*, ApaI, AvaI, AvaII, AccI, BalI, BanI, BamHI*, BclI, BglI*, BglII*, Bsp1286, BssHII, BstEII, ClaI*, DdeI, DraI, EcoRI, EcoRV, HaeII, HaeIII*, HgiAI, HincII, HindIII, HpaI, KpnI*, MluI, NaeI, NarI, NcoI, NdeI, NheI, NotI, NruI, PstI*, PvuI*, PvuII*, SacI*, SacII, SalI*, ScaI, SmaI*, SpeI, SphI*, SspI, StuI, StyI, XbaI*, XhaI, XmaI, XmaIII, XmnI.

* Shorter probe used in blotting experiment (see Figure 11b)

(B) FINE RESTRICTION MAP



(C) RECOGNITION SITES OF ENZYMES WHICH CLEAVE BETWEEN BclI and AvaI

SacII	5'	CCGC•GG	3'
NarI		GG•CGCC	
BanI		$GG^{\bullet TA}_{CG}CC$	
XmaIII		C•GGCCG	
DdeI		C•TNAG	

communication), and presumably serve to eliminate undesirable RNA polymerase "read-through" activity coming from further upstream by making the "melting" of the GC-rich region energetically more difficult.

III.5.2 Comparing Lambda s7 and pMJ2

The integrity of clones pMJ2 and lambda s7 was compared with chromosomal DNA in Southern blotting of double digests of EcoRI and the nearest mapped sites lying 5' to the gene viz. StuI, AvaI, SacII, NarI and DdeI, listed in order of increasing distance from the translational start codon (cf.Figs.11b and 14).

Both clones had StuI and AvaI sites correctly placed with respect to EcoRI, but not the remaining sites lying further upstream (not shown).

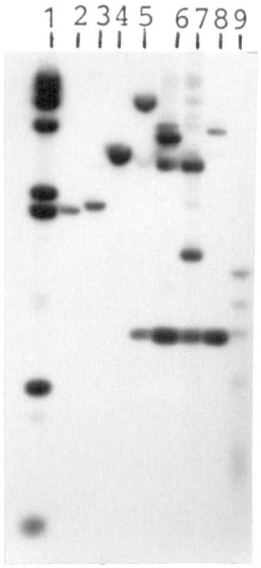
Figure 12 shows the result of probing blots of EcoRI-AvaI digests with progressively shorter probes, whose relationship to the gene is shown in Fig. 12d. Note that the AvaI digestions are incomplete, probably due to the sensitivity of the enzyme to C-methylation.

It is clear that the 2 clones are identical as far upstream as AvaI, since the same fragment in control MFJ29 DNA and the clones hybridizes to a 15-mer, NOR193, which corresponds to the sequence on pMJ2 just 3' from the AvaI site (Figure 14). (The "non-specific" binding is more pronounced with the shorter probes in all cases. The incompleteness of the AvaI digest certainly contributes to the multiplicity of bands seen in the appropriate samples. It is not unlikely though that the sequences carried on the shorter probes, being typical "upstream" sequences, occur several times in the genome).

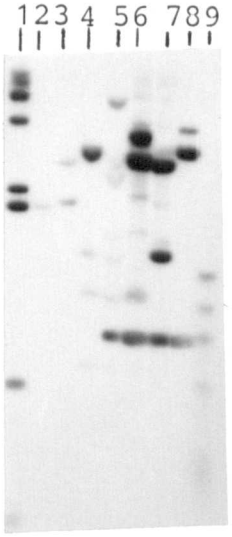
Thus regardless of the mechanism, in both lamba s7 and pMJ2, sequences upstream from the subtilisin gene have been lost, translocated or otherwise altered at a position between 230 and 650 bp from the translational start codon (cf.Figs.11c & 14).

FIGURE 12: COMPARISON of pMJ2 & λs7 by SOUTHERN ANALYSIS

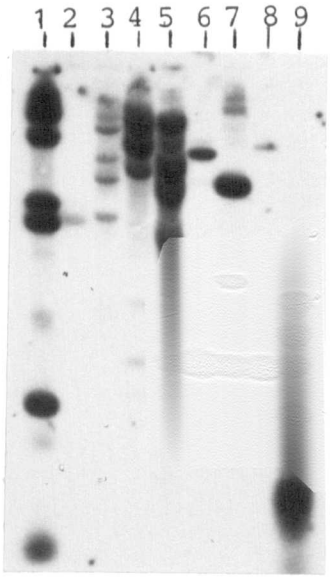
(A) PROBE= StuI-HindIII 500 bp



(B) PROBE= AvaI-StuI 120 bp

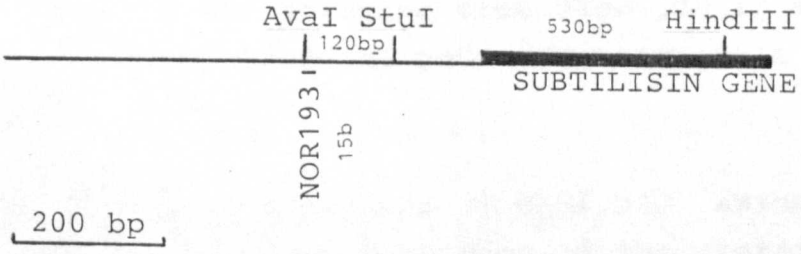


(C) PROBE= 15-mer NOR 193



Lane 1:λ/HindIII labeled markers
Lane 2:MFJ29 DNA/ DdeI
Lane 3:MFJ29 DNA/ HindIII
Lane 4:MFJ29 DNA/ EcoRI
Lane 5:MFJ29 DNA/ AvaI & EcoRI
Lane 6:pMJ2 / AvaI & EcoRI
Lane 7:λs7 / AvaI & EcoRI
Lane 8:pMJ31A / AvaI & EcoRI'
Lane 9:ØX 174/HaeIII labeled markers

(D) RELATIVE LOCATIONS of PROBE TARGETS



For sequencing purposes, and expression studies, material derived from pMJ2 was used exclusively.

No steps were taken to establish whether a deletion had occured in lambda s7 (as was the case in lambda s6) or whether the cloned fragment is intact, ie.extends from a Sau3A1 site between SacII and AvaI.

III.6 Sequence of the Subtilisin Carlsberg Gene and Flanking Sequences

III.6.1 The Sequencing Strategy

The sequencing of the subtilisin gene was undertaken in collaboration with Jan-Ingmar Flock. Almost all the dideoxy sequencing (II.14.2) was carried out in Dr. Flock's or M. Uhlèn's laboratory, whereas the chemical sequencing (II.14.1) was performed by the author.

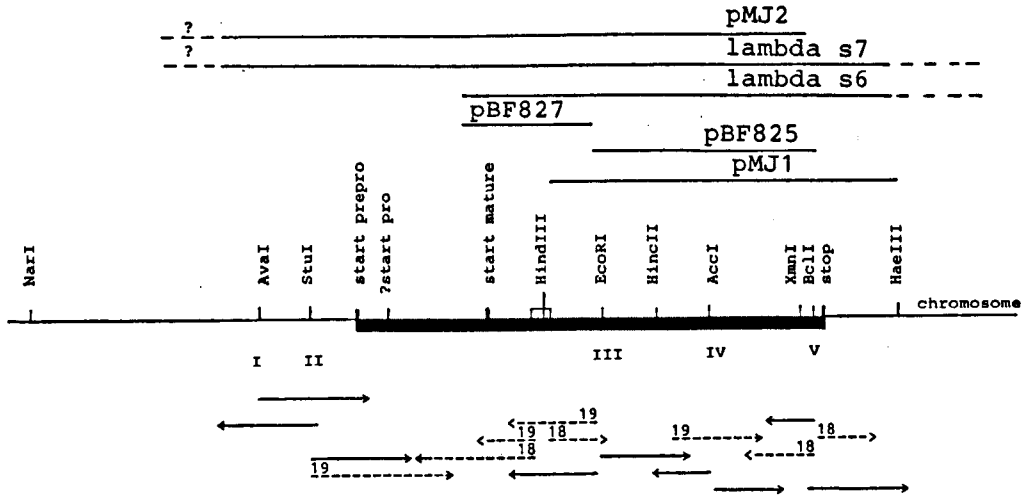
Figure 13 shows the strategy used for sequencing the subtilisin locus. The relationship of the various clones discussed in the text to the gene is shown clearly above the restriction map. Also indicated are the extent of sequence read in each direction, as well as the fragments used to perform the chemical sequencing. In the two short regions where the sequence was obtained in only one direction or on one strand only, it appeared to be completely unambiguous on repeated gels. Figure 14 shows the sequence obtained. Appendix 2 shows diagrammatically the restriction sites found in the sequence.

III.6.2 Open Reading Frame (ORF)

The sequence shows an ORF starting from base 370 in Fig.14, which comprises 1137 nucleotides and terminates in a TAA stop codon at position 1507.

This encodes a 379 residue protein of M_r 38,910, assuming that both the initial Met residues form part of the protein product. It may be that the first Met codon is not used (but see below), and the initiator f-Met may be removed post-translationally (Lewin, 1974).

FIGURE 13: SEQUENCING STRATEGY FOR CARLSBERG GENE



100 bp

gene coding sequences

broken arrows : sequencing by dideoxy method in mp18 or 19

solid arrows : sequencing by chemical method

(direction and extent of sequence indicated)

? exact endpoint of clone uncertain

Fragments used for sequencing from positions I to V:

I	pMJ31A Asp718*-EcoRI	→
II	pMJ42 ClaI*-BamHI	←
	pMJ24 PstIlinker*-EcoRI	→
III	pBF827 EcoRI*-ClaI	←
	pMJ2 EcoRI*-StuI	←
	pMJ1 EcoRI*-BclI	→
IV	pMJ1 AccI*-XmnI 800bp	←
	400bp	→
V	pMJ1 BclI*-EcoRI 500bp	←
	large fragment	→
	pMJ7 BclI*-PvuII 700bp	→

* denotes labelled site

FIGURE 14: SEQUENCE OF SUBTILISIN CARLSBERG GENE AND FLANKING REGIONS

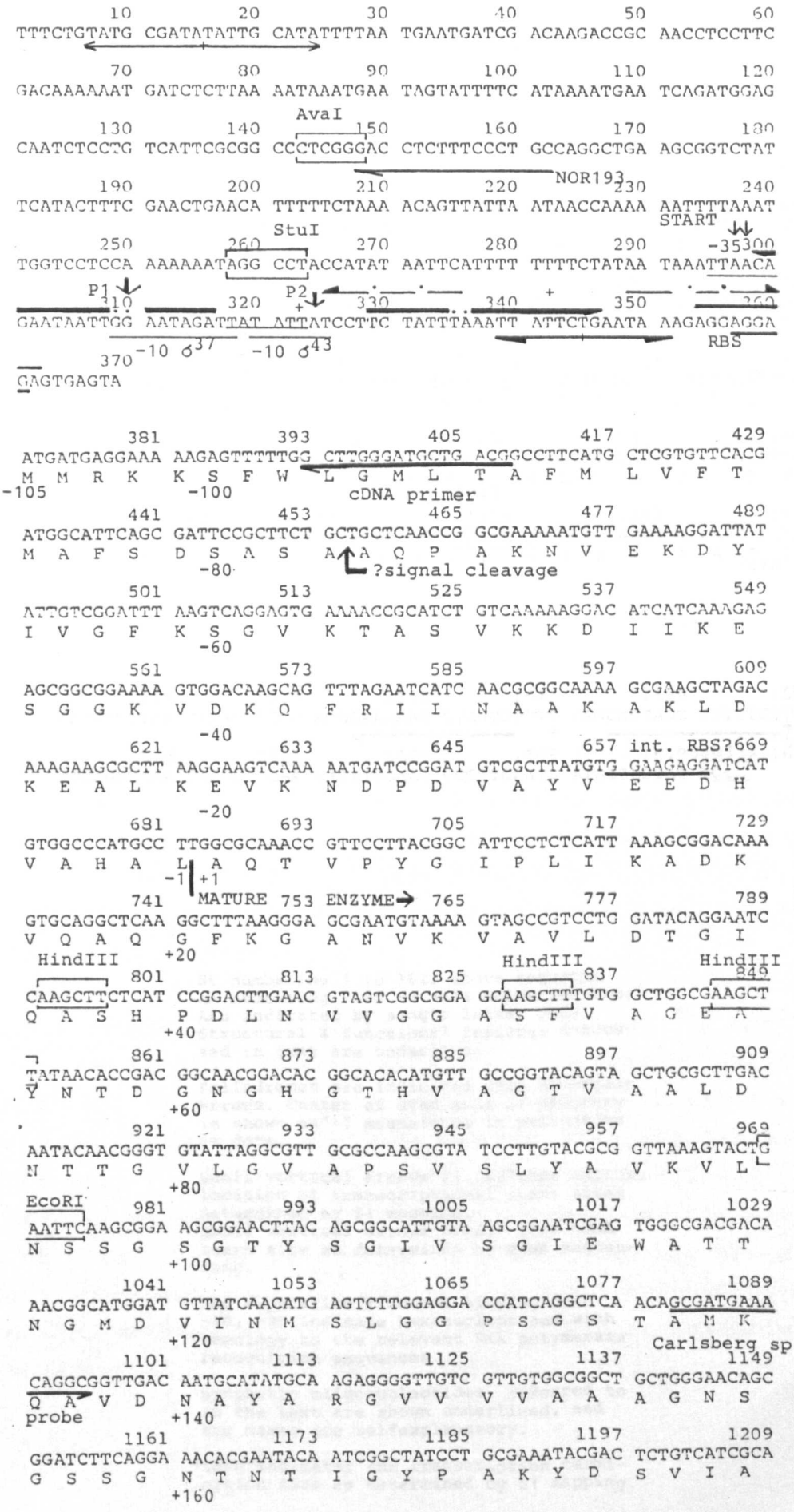
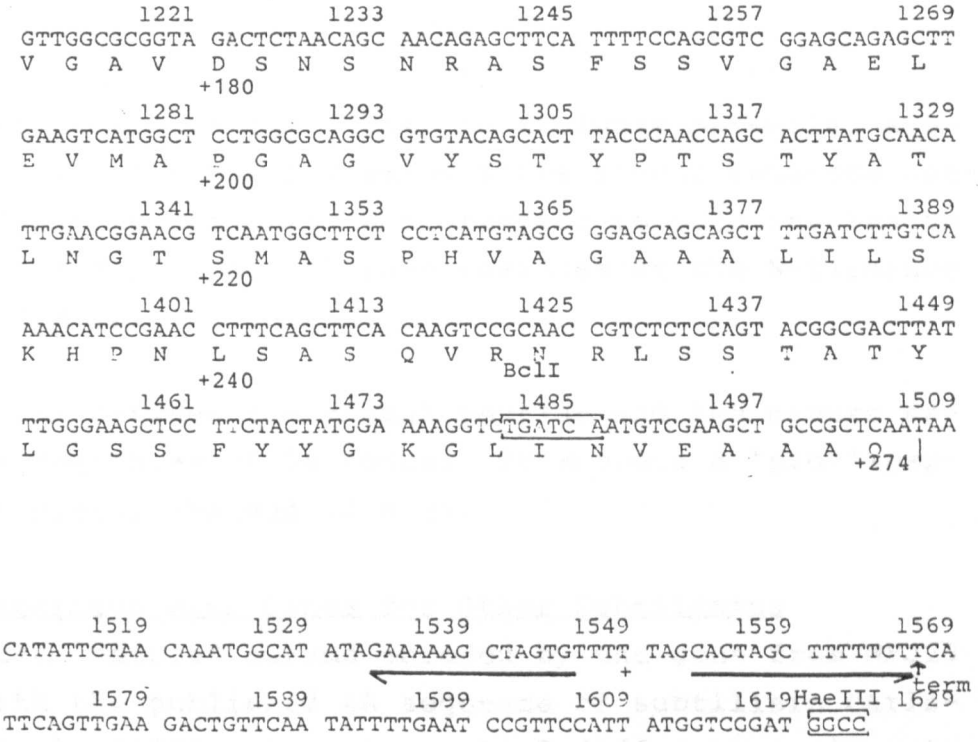


FIGURE 14: SEQUENCE OF SUBTILISIN CARLSBERG GENE AND FLANKING REGIONS (cont.)



Nt numbering 1 to 1623 above sequence
AA numbering -105 to +274 below sequence
AAs indicated by single letter code.
Structural & functional features discussed in text are underlined.

Palindromes are indicated with divergent arrows. Center of dyad axis of symmetry is shown as"+", mismatches in palindrome as dots.

Small vertical arrows"P1, P2"show approx. location of transcriptional start sites determined by S1 mapping.
Small vertical arrows"START" show mRNA start site as determined by cDNA sequencing.

RBS = putative ribosome binding site
-10, -35 indicate hexanucleotides with homology to the relevant RNA polymerase recognition sequences.

Synthetic oligonucleotides referred to in the text are shown underlined, and the names are selfexplanatory.

"term"indicates the transcription termination site as determined by S1 mapping.

The amino acid sequence deduced from the nucleotide sequence is in keeping with expectations when compared with the structure of the B. amyloliquefaciens and B. subtilis subtilisin genes which both encode a BPN' type enzyme (Wells et al,1983; Stahl & Ferrari, 1984 resp.; I.2.2.2;I.8.2).

Using the rules of von Heijne (1984) to evaluate potential signal cleavage sites, a typical Gram-positive signal sequence constitutes the first 29 residues, with respect to both its length and the number of positively charged residues at the N-terminus (viz. 3) (see I.9.1).

There is a region between the signal peptide and the mature exoenzyme coding sequences of 76 codons. It encodes a "pro-" peptide, which is highly charged (I.8.2).

III.6.2.1 Comparison with Genes for Other Subtilisins

When the deduced "mature" enzyme encoded by the gene from MFJ29 is compared with the published AA sequence of subtilisin Carlsberg (Smith et al, 1968), there are only 5 differences viz. at residues +102, 128, 157, 160 and 211. (Fig.1, lines 1 & 2; but note that residue +56 of the BPN' sequence (line 5) is not present in the Carlsberg enzyme, so the numbering for the latter must be adjusted accordingly).

The apparent discrepancies at positions 157, 160 and 211 may be attributed either to the reported amino-acid sequencing ambiguities (Smith et al, 1968) or else to strain variation, since they can be accounted for by a single base change. The last explanation probably also accounts for the differences at +102 and +128. The primary sequence of another Carlsberg-like subtilisin, subtilisin DY (line 3 in Fig.1; Nedkov et al, 1985) has the same amino acids at positons 102, 128 and 157 as those encoded by the gene from MFJ29, whereas at positions 160 and 211 it resembles the protein sequenced by Smith et al (1968).

The deduced prepro- AA sequence of subtilisin Carlsberg and those published for two BPN'-type enzymes were also compared (Figure 15). While there are a number of blocks of conserved residues (underlined), notably before the proposed signal-pro, and pro-mature processing sites, the degree of divergence at the N-terminus (signal sequence) is quite marked (see I.9.1). This difference was highlighted when the corresponding nucleotide sequences of these genes were compared by the "dot matrix" method (Maizel & Levik, 1981; Deveraux et al, 1984). The two sequences are arranged along X and Y axes respectively, and each position (or stretch of residues) tested for identity (or a chosen degree of homology). Where the requirement is met, a dot is placed at the cross-over point. Over similar regions the dots coalesce to form a line of identity. This analysis was carried out by M. Uhlèn at the EMBO laboratories in Heidelberg, and the result shown in Fig.3 of the publication in Appendix 1.

III.6.2.2 Secondary Structure Predictions

In spite of the considerable variation in the signal sequence and the signal-proximal part of the pro-region, secondary structure predictions for the BPN' and Carlsberg enzymes show that their respective hydrophobicity plots (Kyte & Doolittle, 1982), and α-helix/β-sheet predictions (Chou & Fasman, 1978). are almost superimposable, ie. a similar tertiary structure over the entire precursor protein may be expected for both subtilisin types (Fig.4 in Appendix 1). Similar results were obtained with a hydrophilicity plotting program based on Hopp & Woods (1981) (not shown).

III.6.3 Upstream Gene Flanking Region

Just upstream from the ORF there are two overlapping, possible ribosome binding sites ie. sequences complementary to the 3'-end of the B. subtilis 16S rRNA, which has the sequence 3'-OH-UCUUUCCUCCACUA (Shine-Dalgarno, I.7.1; Murray & Rabinowitz, 1982). These are found between nucleotides 353 and 361 in Fig.14.

FIGURE 15: DEDUCED SUBTILISIN PREPRO- PEPTIDES

```
                -100        -90        -80        -70        -60        -50        -40        -30        -20        -10
1)   VRGKKVW ISLLFALALI FTMAFGSTSS AQAAGKSNGE KKYIVGFKQT MSTMSAAKKK DVISEKGGKV QKQFKYVDAA SATLNEKAVK ELKKDPSVAY VEEDHVAHAY
2)     S  L         T        - NM         ST               A  S                    N   A   D                    I  E

3)   MMRKKSFW LGMLTAFMLV FTMAF-SDSA SAAQPAKNVE KDYIVGFKSG VKT--ASVKK DIIKESGGKV DKQFRIINAA KAKLDKEALK  EVKNDPDVAY VEEDHVAHAL
```

Conserved amino acids have been underlined.

1) Deduced prepro- sequence of B. amyloliquefaciens subtilisin (Wells et al 1983) ┐
2) Differences from 1) in deduced prepro- sequence of B. subtilis subtilisin (Stahl and Ferrari 1984) ┘ — BPN´type
3) Deduced prepro-sequence of B. licheniformis subtilisin (Jacobs et al 1985). — Carlsberg type

A number of gaps have been introduced to improve the homology, and are indicated thus : (-)

When calculated by the rules of Tinoco et al (1973), the free energies of formation of mRNA-16S rRNA hybrids for the first possibility (viz.AGGAGGAGA, overlined) is -16.8 kcal, and -11.6 kcal for the second (viz.AGGAG, underlined). Both values are above the average for E. coli RBSs (I.7.4), and the higher one is typical of those found in Gram-positive systems (Hager & Rabinowitz, 1985a).

This suggests that the first probably serves as the RBS, but both possibilities are situated favourably with respect to the first ATG codon: When the spacing between the Shine-Dalgarno and the initiating codon is measured from the base that can pair with the U in the -CCUCC- of the 16S rRNA, the average spacing in E. coli is 10 bases (Stormo et al, 1982). In Gram-positives the value is 10.4 bases (Hager & Rabinowitz, 1985a). As the two possible SD sequences are located 12 and 9 nucleotides resp. from the first ATG, they argue for it indeed being the initiator codon, though this was not shown experimentally.

Note also the potential RBS (ΔG= -10.4 kcal/mol) starting at nucleotide 657 viz. GGAaGaGG, just before the start of the mature coding sequence. An RBS of this type would certainly function in E. coli. In fact it is identical to the putative RBS for the T7 early gene 1 (Table II in Hager & Rabinowitz, 1985a).

While ATG has been shown to be the translation initiation codon in almost all genes from E. coli studied to date, there are a number of examples where GTG has this function (I.7.2; Stormo et al, 1982). It is thus interesting to note that a GTG codon is found just after this postulated "internal" RBS, at codon -5 in the sequence. Such a juxtaposition of possible translational signals does not occur in the BPN' genes, even though the area concerned is highly conserved. It is not inconceivable that translation initiation/reinitiation could occur at this position in E. coli and lead to a detrimental accumulation of an active intracellular form of "mature" subtilisin that has only 5 additional AA residues at its N-terminus. This could be one expla-

nation for the difficulties in cloning the subtilisin Carlsberg gene - in E. coli, at any rate.

As far upstream as the sequence was determined, the region is predominantly AT-rich, and contains runs of A and T residues typical of Bacillus promoter regions (I.3.4; I.3.5.3; I.3.9).

Just upstream from the putative ribosome binding site(s), and centred about position 321, are two overlapping hexamers having a 5 out of 6 match with the consensus sigma-43 -10 recognition sequence viz. TATAAT. At the usual spacing of 17 nucleotides from this, (about position 300), a hexamer having the same degree of similarity to the canonical sigma—43 -35 region (viz. TTGACA) is situated. Note though that the "invariant G" in this stretch is not present (Table 1; I.3.4).

At 309-318, a possible sigma-37 -10 region abutts the one for sigma -43 mentioned above. It has an 8 out of 10 match with the consensus (Table 1,& Wang & Doi, 1984). No corresponding -35 region can however be found.

Inspection of the sequence further upstream shows several possibly significant features. Because of the predominance of A and T residues, many closely related sequences can be found as direct or indirect repeats. A perfect direct 7 bp repeat (at 29-35, 85-91 and 104-111) is seen. Several palindromes are present: one of 9 bp centred between bases 5 and 6, and one of 6 bp centred on base 345. Whether these short sequences have any functional significance is questionable.

There are however two palindromes whose positions relative to the approximate transcriptional start sites/promoter consensus hexanucleotides, and RBS suggest a likely regulatory role.

There is an 18 bp palindrome (with only 2 mismatches) extending from base 299 to base 346, and centred between bases 322 &323. The approximate transcriptional start sites identified by

S1-mapping (III.8.2) certainly lie in the region encompassed by this palindrome. A similar region of dyad symmetry, though less marked than this one, covers the two promoter regions identified for the B. subtilis subtilisin gene (Wong et al, 1984). Such features are typical of operator regions (Rosenberg & Court, 1979), and constitute recognition sites for regulatory proteins.

Most interestingly, the RBS sequence is found on one arm of a 12bp palindrome (325-336 & 349-360). mRNA initiated upstream from position 325 would thus be able to adopt a 2^{o} structure in which the SD is sequestered in the stem, and translation could therefore be made less efficient (see I.7.1).

Such a 2^{o} structure has been postulated to account for the poor ribosome-binding ability of transcripts from the minor glycerol-stimulated P3 gal promoter. In contrast, in the shorter P2 transcript, which is efficiently translated, the SD is placed in the loop (Queen & Rosenberg, 1981). Induction of the cat-86 gene by chloramphenicol is thought to involve destabilization by the inducer of a stem-loop mRNA structure that sequesters the RBS (Ambulos et al, 1986). In a duplication mutant, where a "free" RBS is also available, expression is constitutive (Ambulos et al, 1985).

These speculations would receive some support if exact localization of the transcriptional start sites placed the one upstream from the left-most arm of the palindrome involving the RBS, and the other such that the postulated stem-loop formation was not possible.

If one considers transcripts starting at a position 5' from base 299 in Fig. 14, it is not difficult to imagine two mutually exclusive, alternative conformations for mRNA molecules, where the one involves palindrome 299-346, the other the RBS palindrome 325-360. The transcriptional start site of the subtilisin Carlsberg gene identified by cDNA synthesis and sequencing (Appendix 3) in fact lies around base 298, so mRNA

secondary structures may very likely be important in the control of this gene.

Demonstration of promoter activity in these flanking sequences is described in III.7.

III.6.4 Downstream Gene Flanking Region

A palindrome follows the Open Reading Frame in Figure 14. The calculated free energy of the hairpin structure is -21.1 kcal /mol (Devereaux et al, 1984) which is similar to other Gram-positive termination signals (Lehtovaara et al, 1984; Paddon & Hartley, 1986; Wells et al, 1983, Vasantha et al, 1984).

Evidence is presented in III.8.3 to show that this structure indeed serves as the transcriptional terminator for the subtilisin gene in B. licheniformis, from which it originates.

III.7. Expression of the Subtilisin Carlsberg Gene.

III.7.1 Assembling the Gene

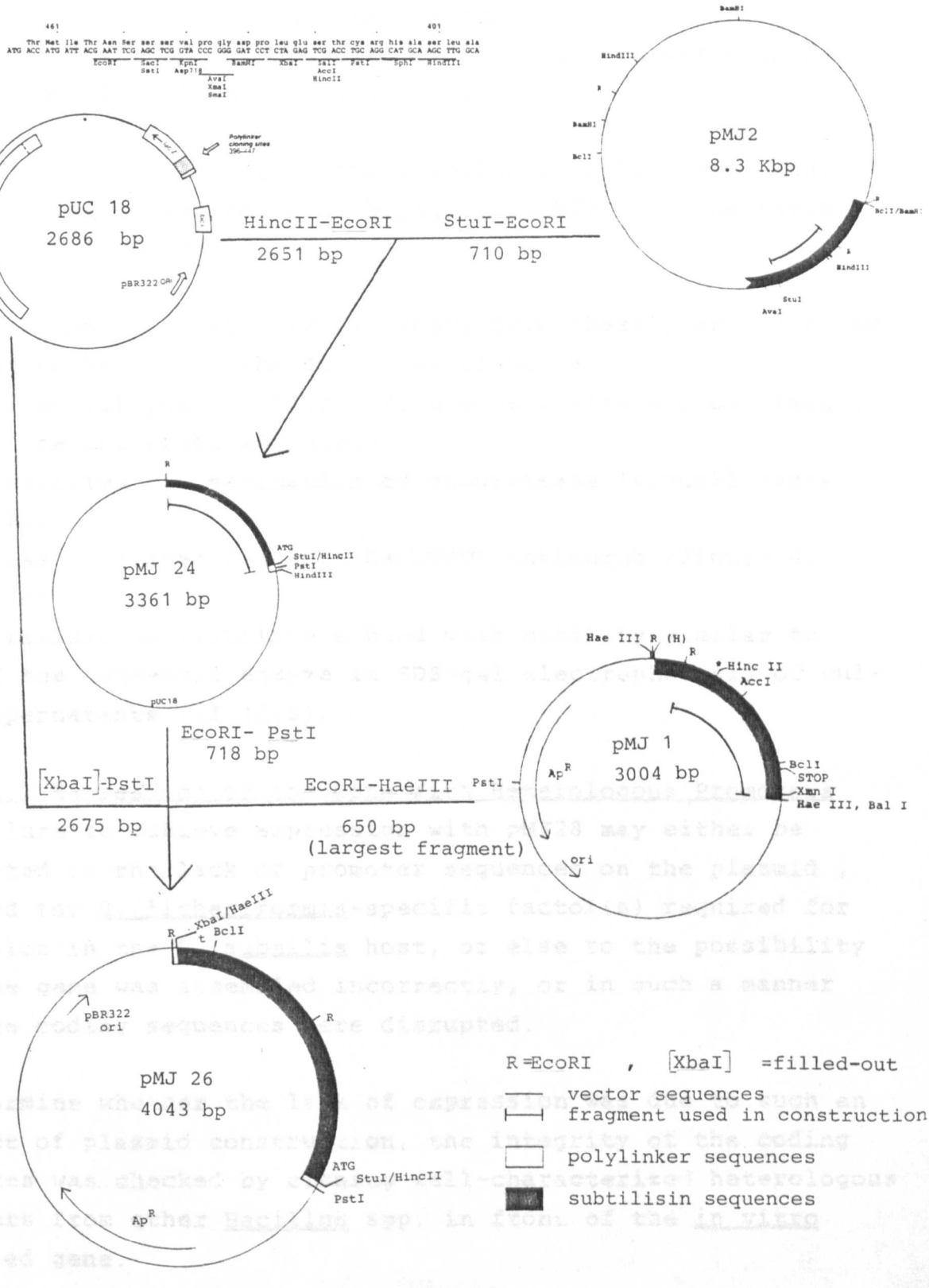
Since an attractive sigma-43 promoter had been found just upstream from the gene coding sequences (Figure 14, III.6.3), the gene was assembled in E. coli from the partial clones pMJ1 and pMJ2 , and then transferred to B. subtilis to determine whether promoter activity could be detected.

Figure 16 depicts schematically how the 5' portion (from pMJ2) and the 3' part of the gene (from pMJ1) were spliced together at the EcoRI site. Note: The 5' limit of the fragment taken from pMJ2 was StuI. At the time these experiments were carried out, the presence of the AvaI site had not yet been established.

These steps were carried out in the low copy mutant host MFJ233 to avoid the deletions seen with pMJ2 at normal plasmid copy number. The final plasmid pMJ26, carrying the complete reconstituted gene, proved to be stable upon retransformation into MFJ92.

FIGURE 16

ASSEMBLING THE SUBTILISIN GENE



III.7.2.1 Transferring the Gene to a Gram-Positive Replicon for Expression

Figure 17 shows how the complete coding sequence was transferred from pMJ26 to Gram-positive vectors.

Both p820 and its low copy number analogue, p900, were used. The recombinant was stable in B. subtilis MFJ93 in the higher copy number form, pMJ28.

No expression of subtilisin Carlsberg from these plasmids could be detected by any of the following criteria:
(a) on skim-milk plates (II.9.1.1; the halo size was not larger than in the untransformed host),
(b) by colorimetric estimation of exoprotease (Azocoll assay, II.9.1.2),
(c) by use of a specific anti-Carlsberg antiserum (Figure 4, well 3) or
(d) by failure to vizualize a band with mobility similar to that of the authentic enzyme in SDS-gel electrophoresis of culture supernatants (II.12.3).

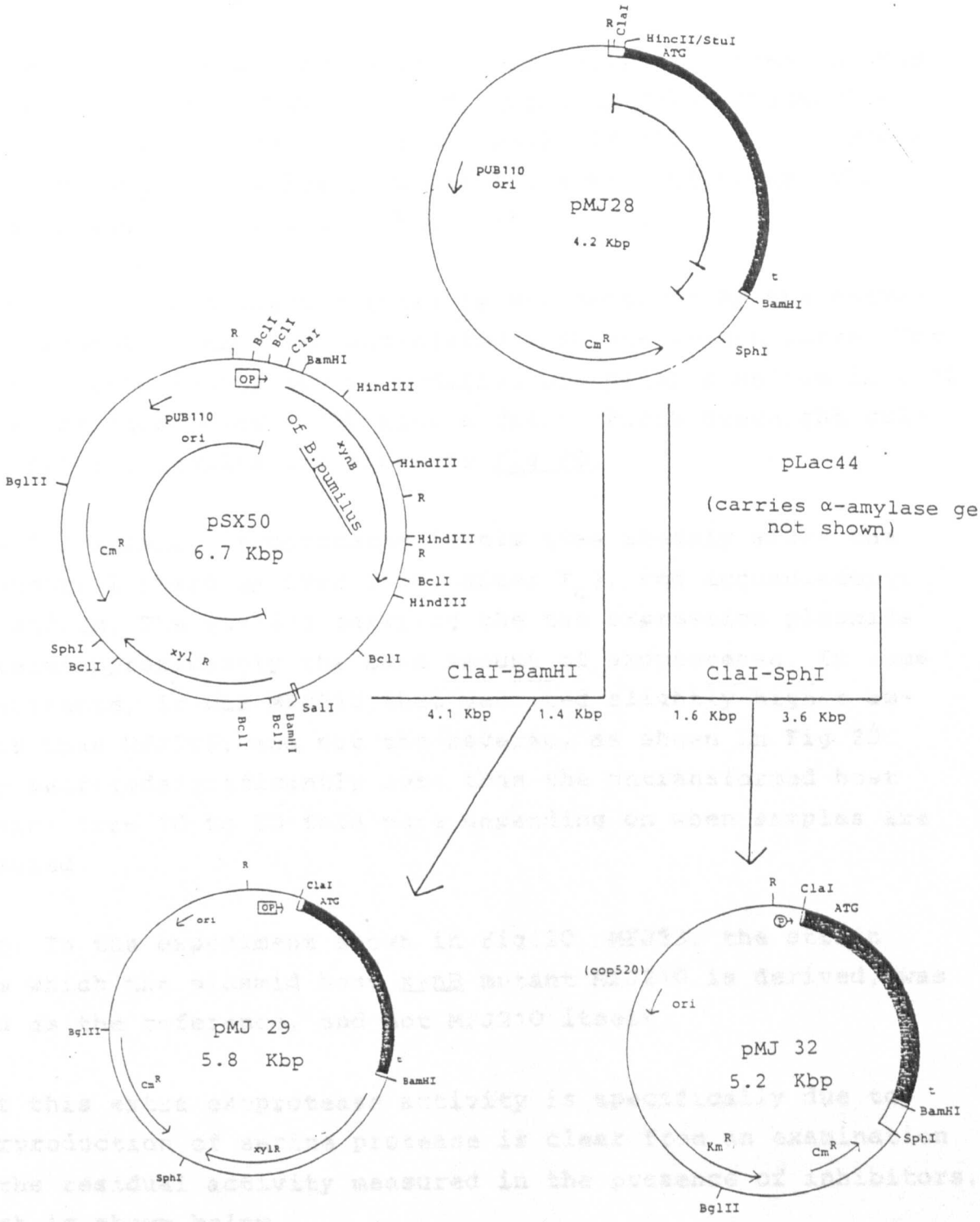
III.7.2.2 Expression of the Gene With Heterologous Promoters

The failure to achieve expression with pMJ28 may either be attributed to the lack of promoter sequences on the plasmid , the need for B. licheniformis-specific factor(s) required for expression in the B. subtilis host, or else to the possibility that the gene was assembled incorrectly, or in such a manner that the coding sequences were disrupted.

To determine whether the lack of expression was due to such an artefact of plasmid construction, the integrity of the coding sequences was checked by cloning well-characterized heterologous promoters from other Bacillus spp. in front of the in vitro assembled gene.

Figure 18 shows the steps involved in constructing pMJ29 and pMJ32, where respectively the xylose-inducible B. subtilis

FIGURE 18: EXPRESSION OF SUBTILISIN GENE WITH HETEROLOGOUS BACILLUS PROMOTERS



xynB promoter, and B. stearothermophilus α-amylase promoters were utilized.

Figure 19 shows the large halos of proteolysis formed on skim-milk plates by supernatants of B. subtilis cells containing plasmid pMJ32 (MFJ213) and pMJ29 (MFJ205 , grown in the presence of xylose). Under the same conditions, the untransformed host shows no halo (MFJ93).

Extracellular proteolytic activity was measured by the convenient Azocoll method and correlated with the growth curve. For all such experiments 100 ml modified Schaeffer's medium in a 1L flask was inoculated 1:100 with a fresh LBroth overnight culture (II.2). Results are shown in Fig.20.

In all 3 strains, exoprotease levels rose steeply after the exponential phase was over (i.e. after T_o), and accumulated in the medium. The strains carrying the two expression plasmids secreted approximately the same amount of exoprotease. In some experiments, it was MFJ213 that secreted slightly higher amounts than MFJ205, and not the reverse, as shown in Fig.20. They secreted significantly more than the untransformed host strain: from 10 to 20-fold more depending on when samples are measured.

Note: In the experiment shown in Fig.20, MFJ93, the strain from which the plasmid host xynB mutant MFJ210 is derived, was used as the reference, and not MFJ210 itself.

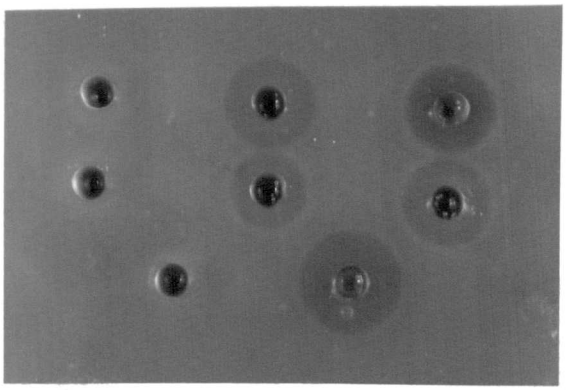
That this extra exoprotease activity is specifically due to overproduction of serine protease is clear from an examination of the residual activity measured in the presence of inhibitors, which is shown below.

FIGURE 19

OVERPRODUCTION of SUBTILISIN from CLONED GENES

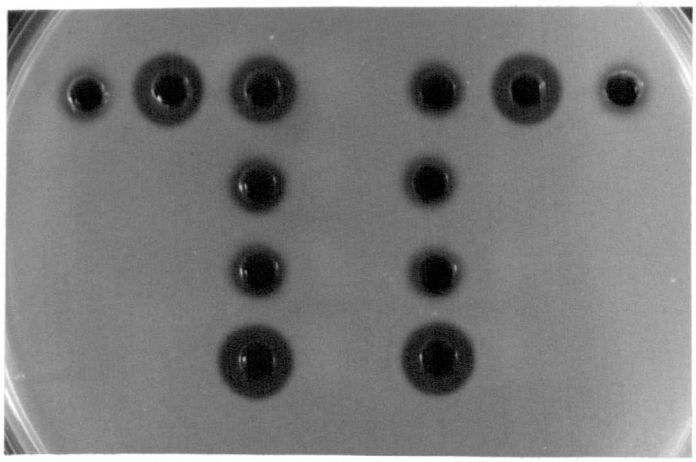
Samples are 20h LBroth supernatants, unless otherwise stated. "*" designates 4-day BP-X samples.

Skim-milk /Gelatin Plate



MFJ93 135 93*
93 213 205(+xylose)
29 29*

Skim-milk Plate



MFJ195 246 210 157 250 249
196 248
195 249
246 250

REDUCED SUBTILISIN PRODUCTION by apr DELETION MUTANTS

Skim-milk Plate

MFJ 93 157 168wt MFJ231

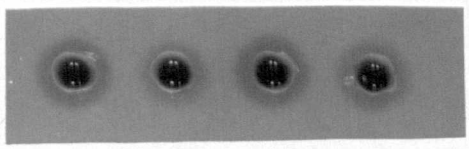


FIGURE 20: EXPRESSION OF SUBTILISIN CARLSBERG BY HETEROLOGOUS PROMOTERS

Growth in Schaeffer's medium was monitored by optical density at 450 nm. Glucose was omitted from, and 0.2% xylose (w/v) added to, the medium in the case of MFJ205. Exoprotease activity was measured by the Azocoll assay (II.9.1), and results recorded as A_{520}.

MFJ 93 = "wt"
MFJ213 = MFJ210 (pMJ32), α-amylase promoter
MFJ205 = MFJ210 (pMJ29), xylosidase promoter

GROWTH A_{450}		PROTEASE ACTIVITY A_{520}
X	MFJ 93	*
Δ	MFJ213	□
▲	MFJ205	■

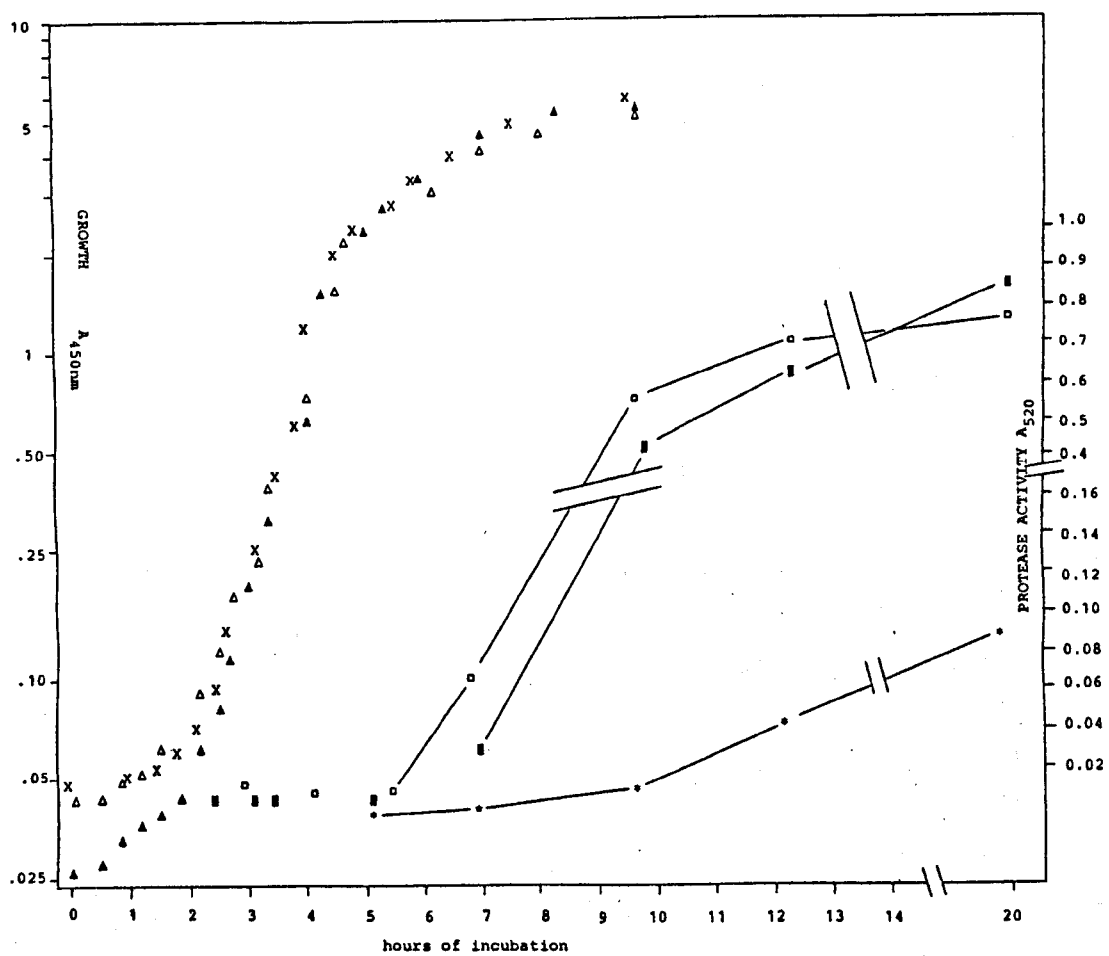


Table 7: Overproduction of Carlsberg Enzyme Using Heterologous Promoters

Exoprotease Production in MFJ93, 205 & 213

A_{520} 20hr supernatants; normalized for sample volume of 0.45 ml.

	Total	+EDTA 10mM	+PMSF 1 mM
MFJ 93	0.083	0.016	0.044
MFJ205	0.900	0.814	0.020
MFJ213	0.805	0.790	0.013

Note: It was consistently found that the sum of the activities measured individually in the presence of the inhibitors usually failed to reach 100%. This is discussed more fully in IV.6.4.4.

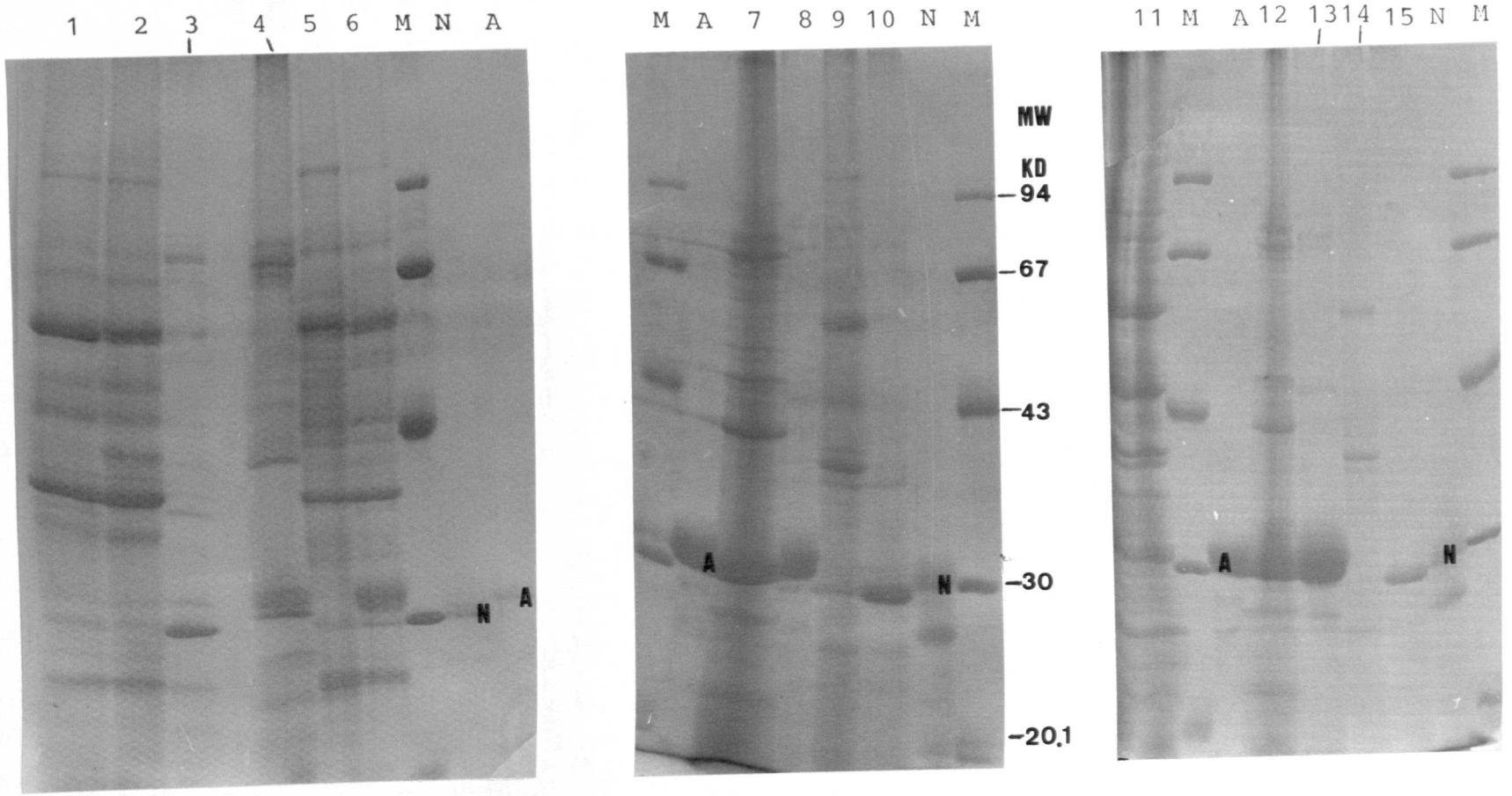
Almost all the activity disappears when the serine protease inhibitor PMSF is preincubated with the sample. In contrast, the metalloprotease inhibitor EDTA has little effect. The lower than wild-type residual activity in supernatants of overproducers MFJ 205 and 213 in the presence of PMSF is most probably due to degradation of the neutral protease by the excess subtilisin secreted into the medium (Dancer and Mandelstam, 1975a).

Proof that the "extra" serine protease secreted by MFJ 205 and 213 is of the subtilisin Carlsberg type was obtained by SDS gel electrophoresis and immunodiffusion experiments.

Figure 21 shows that MFJ 205 (lane 6, grown in the presence of xylose) and MFJ 213 (lanes 8 & 13) secrete a protein that co-electrophoreses with a prominent diffuse band in the supernatant from MFJ29 (lanes 4,7 & 12), the strain from which the subtilisin Carlsberg gene was cloned. This band has a mobility similar to, but not identical with, that of authentic subtilisin Carlsberg (lanes A). It is not present in samples from the untransformed host MFJ 210 (lanes 9 & 14), and is distinct from both authentic subtilisin NOVO (lanes N) and B. subtilis's own

FIGURE 21: SDS GEL ELECTROPHORESIS of SUPERNATANTS from EXOPROTEASE-PRODUCING STRAINS

Samples are TCA precipitates of culture supernatants. Lanes 1-6, LBroth overnight cultures; lanes 7-15, 4-day BP-X cultures.
Standards: M, molecular weight standards; A, "Alcalase" (commercial subtilisin Carlsberg); N, commercial subtilisin NOVO. Samples: lane1, MFJ93; 2, MFJ157; 3, 10, 15, MFJ135; 4,7,12, MFJ29; 5, MFJ198; 6, MFJ205 (+xylose); 8,13, MFJ213; 9,14, MFJ210; 11, MFJ246.



subtilisin, seen as the most prominent band of overproducer MFJ135 (lanes 3,10 & 15). The small differences between the mobility of the cloned gene products and the markers is most probably due to strain variation ie. small amino acid/charge differences in the enzymes.

Figure 4a shows the immunoprecipitin band of identity between subtilisin Carlsberg ("Alcalase", wells 1 & 4) and a component of the culture supernatants of MFJ29 (well 5) and 213 (well 6).

The failure to detect expression with pMJ28 (MFJ198, Fig.4A, well 3; Fig.21, lane 5) was thus not due to some error in reassembling the coding sequences for subtilisin. It probably indicates the lack of a functional promoter in front of the gene.

III.7.3 Processing of Prepro-Subtilisin

An interesting point arises regarding the timing of subtilisin secretion seen in the two strains MFJ 205 and 213.

While one might well expect a post-exponential appearance of subtilisin when it is under its own regulatory elements (I.2.1; I.6.5.3, MFJ93 in Fig.20), or under an α-amylase promoter (I.5.1.1, Nicholson et al, 1985), this is certainly not the case with the induced xylosidase promoter used in pMJ29.

Moreover, in a construction identical to pMJ32, but where the B. stearothermophilus α-amylase promoter drives transcription of a β-lactamase gene instead of the subtilisin gene in a B. subtilis host, it was found that β-lactamase was expressed constitutively from the plasmid (L. Christiansen, personal communication). In this particular instance therefore, an earlier appearance of subtilisin may also have been expected. The multiple copies of the promoter region evidently disturb the temporal activation normally seen with the single chromsomal copy in the strain from which the promoter derives. Normal gene regulation is often impaired by amplification of target sequ-

ences (see I.6.5.2 for other examples).

The timing of expression in MFJ205, grown in the presence of 0.2% xylose, was examined more closely to ensure that low levels of exoprotease at an early time in growth had not escaped detection.

Figure 22 shows the temporal pattern of both induced xylosidase expression (intracellular) (measured with PNPX as described in II.9.2), and exoprotease activity in the reference strain MFJ 93. In the absence of xylose, MFJ93 shows no xylosidase activity (data not shown). The cells were grown in the presence of 0.2% xylose (w/v).
Note: As the MFJ93-derived pMJ29 host MFJ210 is Xyl^-, MFJ93 was used as a control. Xylosidase promoter activity in MFJ210 can only be monitored by following synthesis of the mutant xynB mRNA. It has however been established that the kinetics of xylosidase induction in MFJ210(pSX50), ie.in the Xyl^- mutant carrying a plasmid with the intact gene, are identical to those of the wild-type chromosomal gene (Sven Hastrup, personal communication).

Xylosidase is detectable fully 2 hours prior to T_o in MFJ93. When the early samples were assayed for a short period, the initial rate of reaction was directly proportional to cell density over the range A_{450} 0.3 to >2. (See inset in Fig.22). Exoprotease is first detected in stationary phase.

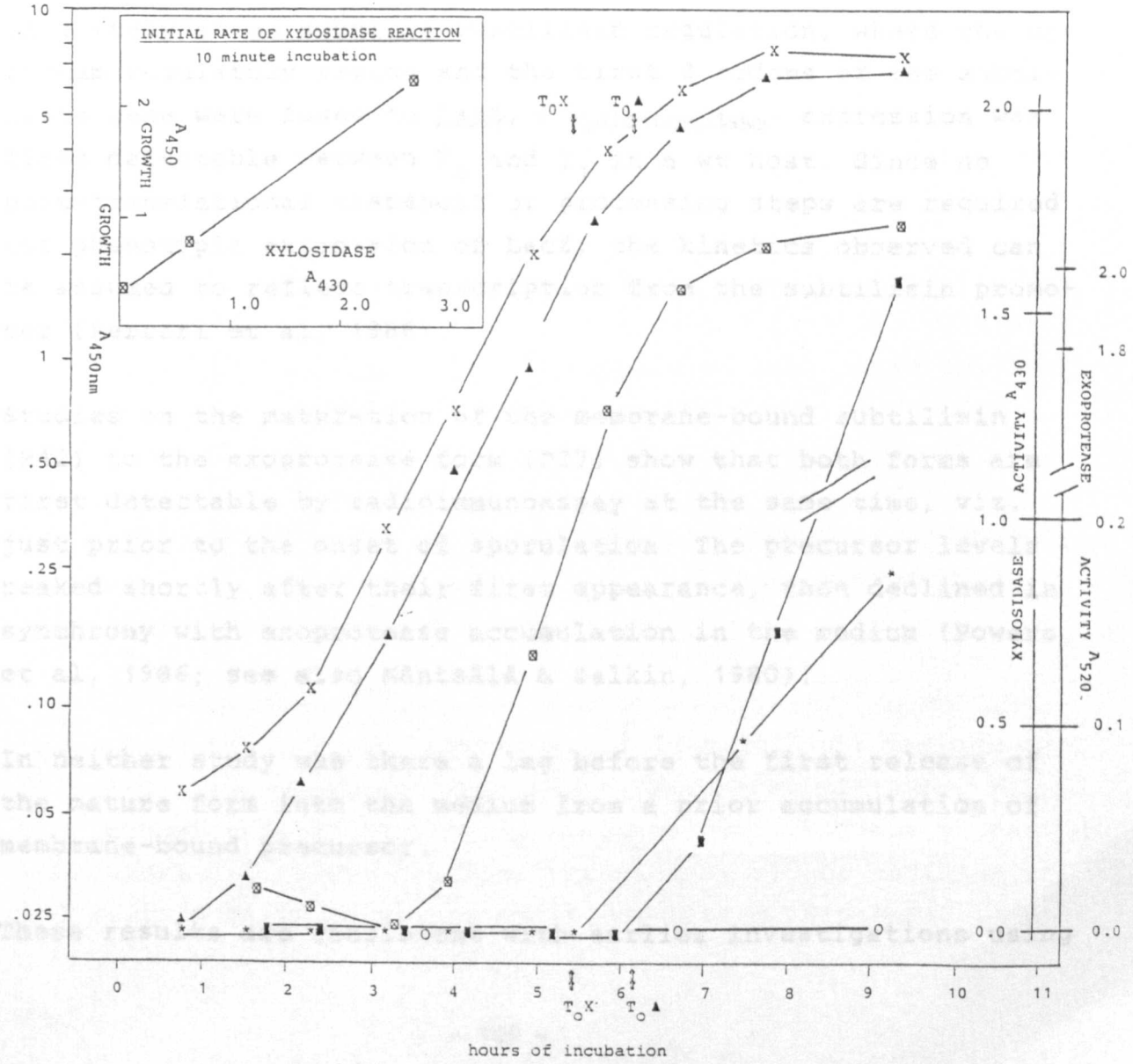
In MFJ 205 too, grown under identical conditions, exoproteolytic activity is first detected after T_o. The value obtained in the Azocoll protease assay soon after T_o, where cell density is 5 (viz. A_{520}=0.04), is sufficiently high to have permitted reliable measurement of exoprotease down to densities of 0.5 or lower, assuming that subtilisin secretion had been linearly related to cell density. However, no such activity was seen immediately prior to T_o, at OD=3.

FIGURE 22: DELAYED APPEARANCE OF SUBTILISIN IN THE MEDIUM WHEN EXPRESSED BY AN INDUCED PROMOTER ACTIVE IN EXPONENTIAL PHASE

Cells were grown in Schaeffer's medium containing 0.2% xylose, as described in the text. As indicated in the figure, samples were removed for determination of exoprotease activity (Azocoll assay, II. 9.1.2, A_{520}), and intracellular xylosidase activity (II.9.2, A_{430}), at different times of growth.

MFJ93 = "wt" ; MFJ205 = MFJ210(pMJ29)

	GROWTH	PROTEASE ACTIVITY	XYLOSIDASE ACTIVITY
	A_{450}	A_{520}	A_{430}
MFJ 93	X	*	⊠
MFJ205	▲	▮	O



One must thus conclude that no extracellular subtilisin was present at the earlier times.

Although it was not shown directly (by the measurement of subtilisin-Carlsberg specific mRNA in MFJ205, or by following a transcriptional fusion with xynB), that the induced xynB promoter indeed directs subtilisin transcription in exponential phase (where it is demonstrably active in the reference strain MFJ93, Figure 22), there is not much reason to doubt that it does so. MFJ210 was derived from MFJ93 by substitution of the wild-type xynB allele with the same gene carrying a 4 bp deletion in the xylosidase coding sequence. This leads to the Xyl^- phenotype, but the regulatory elements are unaltered (see Note above).

In a study of B. subtilis subtilisin regulation, where the upstream regulatory region and the first 8 codons of the subtilisin gene were fused to lacZ, β-galactosidase expression was first detectable between T_0 and T_1 in a wt host. Since no post-translational transport or processing steps are required for phenotypic expression of LacZ, the kinetics observed can be assumed to reflect transcription from the subtilisin promoter (Ferrari et al, 1986).

Studies on the maturation of the membrane-bound subtilisin (P42) to the exoprotease form (P27) show that both forms are first detectable by radioimmunoassay at the same time, viz. just prior to the onset of sporulation. The precursor levels peaked shortly after their first appearance, then declined in synchrony with exoprotease accumulation in the medium (Powers et al, 1986; see also Mäntsälä & Zalkin, 1980).

In neither study was there a lag before the first release of the mature form into the medium from a prior accumulation of membrane-bound precursor.

These results are consistent with earlier investigations using

B. amyloliquefaciens, in which it was demonstrated that newly synthesized α-amylase and protease molecules enter an intracellular pool. After translation is complete, up to 15 minutes can elapse before there is release of enzyme into the medium (Gould et al, 1975). Vasantha & Thompson (1986) have similarly observed an (unspecified) lag between the synthesis of apr or npr presequence-spa fusions, and emergence of protein A into the medium.

As protoplasts neither displayed a pool, nor the lag between synthesis and secretion (Gould et al, 1975), it seems likely that the delay reflects passage through the cell wall. Pulse-chase experiments revealed that the delay was however not obligatory, and that individual, newly-synthesized molecules could emerge in the shortest time measured in above-mentioned study, viz. 5 mins.

It seems therefore that transcription and first appearance of subtilisin in the medium are closely linked, and normally take place around To.

Why there should be a considerable delay (several hours) after the presumed time of transcription (deduced from the xylosidase expression in MFJ93), before extracellular subtilisin appears in the medium of MFJ205 is not clear, but deserves further study.

It is notable that the pMJ29-encoded exoprotease is first detected at the usual time of appearance of the endogenous subtilisin in stationary phase.

Based on the observation that maturation (prepro- to mature enzyme) of heterologous mutant inactive subtilisin precursors failed to take place in Δ apr hosts, Powers et al (1986) concluded that this processing is autocatalysed. Release of mature (inactive) P27 enzyme could be effected by either addition of authentic subtilisin to the cells, or else by prolonged incu-

bation, until almost all cells have sporulated, lysed , and released their intracellular proteases. The authors mention the possibility that it may be these proteases which initiate subtilisin processing *in vivo*, by having sporulated early around To. This seems highly plausible since Beppu (personal communication) also measured increased *exo*protease production in ISP overproducers (I.2.3.3.3). Significantly for the autoactivation hypothesis, preprosubtilisin produced *in vitro* remains in the P42 precursor form. This suggests that *in vivo* factors eg. membrane location, precursor folding, processing enzyme, other co-factor) are necessary to give rise to the P27 mature enzyme.

It would therefore be interesting to examine subtilisin-specific mRNA production and subtilisin antigen associated with membranes in exponential cells of MFJ205 to determine whether the late release observed is attributable to transcriptional or post-transcriptional factors.

It remains to be established firmly whether the *xynB* promoter on pMJ29 is active in that period, and that initiation takes place at that promoter. It is unlikely, but not inconceivable, that the *B. licheniformis* sequences immediately 5' to the subtilisin gene could hinder early transcription in some way. For example, they could bind a negatively-acting factor that prevents passage of RNA polymerase past it early in growth. Alternatively, transcription from the *xynB* promoter through the palindrome that includes the RBS (Fig.14, base 325-360) could lead to an mRNA that is not translatable (as discussed in III. 6.3), possibly until some other factor is present in the cell.

Should promoter activity be as expected, it should be possible to demonstrate accumulation of P42 in the membrane until active enzyme is released after To. This would indicate that whatever triggers subtilisin maturation in *B. subtilis* is not present earlier in growth.

III.7.4 Testing "Upstream" Sequences From pMJ2 for Promoter Activity

Since no promoter activity had been found downstream from the StuI site (pMJ28; Figure 17), the adjacent 5' sequences were isolated from pMJ2 and examined for their ability to direct transcription of the subtilisin Carlsberg gene (pMJ42) or a promoterless xylosidase gene (pMJ33).

Figure 23 shows how the 2.9 kb EcoRI-ClaI "upstream" fragment was put together with the above-mentioned genes. To facilitate the constructions, a ClaI linker was inserted in the unique StuI site in pMJ2. The resultant plasmid is called pMJ2-ClaI. To ensure that the correct pMJ2 sequences were used, the HindIII-ClaI band, which runs well separated from the other large HindIII band on agarose gels was first isolated and then cut with EcoRI.

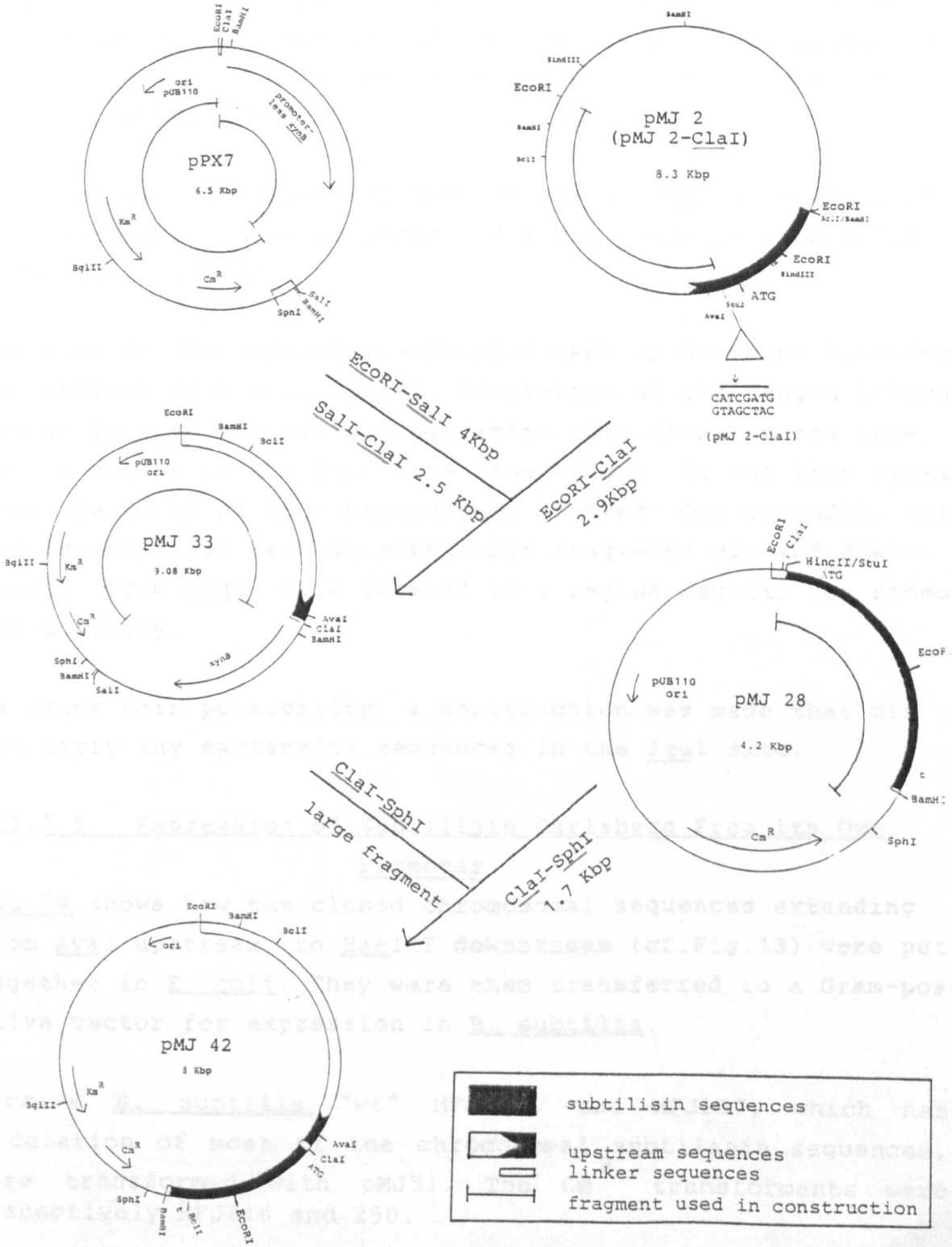
MFJ210 was transformed with pMJ33 (resulting in MFJ219) and pMJ42 (to give strain MFJ221).

The former strain showed no sign of xylosidase synthesis using the sensitive fluorescent substrate MUX (II.9.2), and the latter no sign of subtilisin Carlsberg synthesis either by immunodiffusion (II.15) or SDS gel electrophoresis (II.12.3) (data not shown).

This was surprising as the instability of pMJ2 at high copy number (III.4.4.2.3) had been interpreted as indicating some activity detrimental to the host. Despite the signs of rearrangement upstream from the subtilisin gene (III.4.5.2), it was anticipated that some subtilisin promoter activity was present.

It therefore appeared that the subtilisin Carlsberg promoter

FIGURE 23: TESTING CLONED "UPSTREAM" SEQUENCES
(5´from StuI site) FOR PROMOTER ACTIVITY



either might be situated relatively far from the gene itself, ie. beyond the 5' limit of the gene-proximal sequences which were unaffected by rearrangement. Alternatively, B. licheniformis-specific factor(s) might be required for utilization of the promoter. Since transformation of B. licheniformis was not successful in my hands, this latter possibility could not readily be tested.

To ascertain approximately how far the in vivo promoter lies from the gene coding sequences, RNA was isolated from MFJ29 (see below, III.8).

The size of the subtilisin-specific mRNA in Northern blotting was shorter than anticipated. Estimation of its length indicated an in vivo transcriptional start site close to the gene, in the region of the StuI site (see below). It was thus likely that the 30 bp of non-chromosomal, "linker" DNA on pMJ42, which are interspersed between subtilisin fragments up- and downstream from StuI, were located in a region crucial for promoter activity.

To check this possibility, a construction was made that did not carry any extraneous sequences in the StuI site.

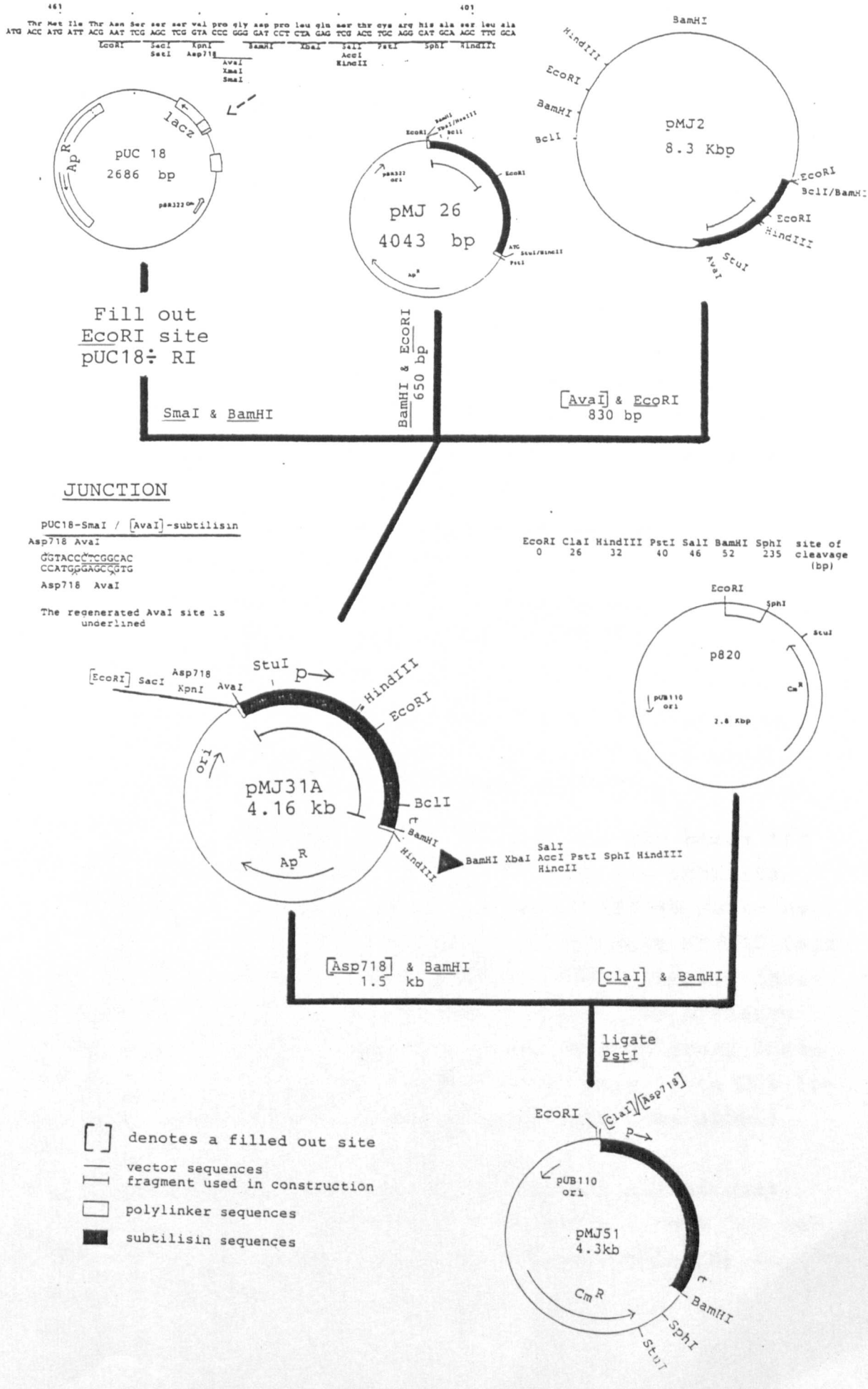
III.7.5 Expression of Subtilisin Carlsberg From its Own Promoter

Fig.24 shows how the cloned chromosomal sequences extending from AvaI upstream, to HaeIII downstream (cf.Fig.13) were put together in E. coli. They were then transferred to a Gram-positive vector for expression in B. subtilis.

Strains B. subtilis "wt" MFJ210, and MFJ157, which has a deletion of most of the chromosomal subtilisin sequences, were transformed with pMJ51. The Cm^R transformants were respectively MFJ246 and 250.

On L Broth plates containing skim-milk, exoprotease overproduc-

FIGURE 24: ASSEMBLING THE SUBTILISIN GENE FOR EXPRESSION FROM ITS OWN PROMOTER



tion (ie.halo size) relative to the untransformed hosts was barely discernable with the wt, but was obvious in the case of MFJ250. These differences are more readily demonstrated by testing 20h L Broth culture supernatants on agarose plates (II.9.1.1). Fig.19 shows the dramatic difference between MFJ157 and its pMJ51 transformant, MFJ250 . In contrast, MFJ210 and its pMJ51 transformant MFJ246 have more or less the same exoprotease activity.

The extent of overproduction of 4-day BP-X cultures was quantified by the OPA assay, with casein substrate (method not described here; H. Outtrup, personal communication). Shown below are absolute exoprotease levels for "wt" MFJ210 and a series of Cm^R transformants.

Table 8: Overproduction of Carlsberg Enzyme Using Own Promoter

	OPA	strain
MFJ210	2.9	wt with respect to protease
MFJ196	1.2	MFJ210(p820)
MFJ195	1.3	MFJ210(pMJ28)
MFJ246	10.1	MFJ210(pMJ51)
MFJ157	0.8	apr deletant

It is obvious that the vector p820, which forms the basis for the expression plasmids in this study, in some way inhibits exoprotease production. The culture medium of MFJ196 contains less than half the proteolytic activity of the host MFJ210 (see also Fig.19). This cannot be accounted for solely by any inhibitory effect on growth exerted by the plasmid. The presence of p820, pMJ28, or pMJ51 in the cell leads to stationary phase cell densities in the presence of Cm 6 µg/ml only 15 to 20% lower than that reached by the untransformed host (not shown).

The difference between MFJ195 and MFJ196 is not significant, though it is not impossible that pMJ28 supports a very low level of subtilisin Carlsberg synthesis, not detectable by any of the means tried (III.7.2.1).

pMJ51 transformant MFJ246 secretes 3x more exoprotease than does the host, MFJ210. Due, however, to the depressive effect of the vector on exoprotease levels (seen in both MFJ195 & 196), the Carlsberg promoter may actually be directing the production of 7-8x more protease. This represents a promoter activity rather lower than that shown by the cloned endogenous B. subtilis subtilisin promoter under similar conditions, and on a vector closely related to p820 (see IV.5.2).

The vector effect was also obvious when p820 and pMJ28 were transferred to the subtilisin deletant (Fig.19, MFJ248 and 249 respectively, cf. with MFJ157).

An homologous series of p820,pMJ28 and pMJ51 transformants of MFJ157 were also tested in the OPA assay. L Broth 20h supernatants did not contain protease levels high enough to be detected in that assay. Protease accumulates to at least 50-100x higher levels during a 4 day incubation in BP-X, than can be achieved in L Broth (H. Outtrup, personal communication). In the 4 day incubation in BP-X, it was clear that the plasmids had been lost from MFJ157 (all transformants had the same activity as the host). A Cm concentration of 15 µg/ml is needed for plasmid maintenance in MFJ157, and levels had evidently fallen too low during the incubation.

That the overproduction of exoprotease by MFJ246 was due specifically to synthesis of the Carlsberg-type subtilisin, was demonstrated by immunodiffusion, and by direct visualization of the protein by SDS gel electrophoresis. Fig.4B illustrates the immunological identity of authentic subtilisin Carlsberg ("Alcalase"), and a component of the supernatant of MFJ246 (wells 10 & 12). In contrast, a supernatant containing large amounts of the BPN'/NOVO type enzyme, did not cross-react with the serum (overproducer MFJ135, well 9).

A band comigrating with authentic Alcalase was seen in 4 day

supernatants of MFJ246 (Fig. 21, lane 11). This band was not present in samples from the host MFJ210 (lane 9). In similar samples from strains which overproduce subtilisin from a cloned gene (BPN' type: MFJ135, lanes 10,15; Carlsberg type: MFJ213, lanes 8,13), the subtilisin band was by far the most prominent. This is presumably due to degradation of other components by the protease. This was not the case with MFJ246, and probably reflects the very moderate increase in exoprotease production.

III.8 Mapping the Transcriptional Start & Terminator Sites

Promoter activity had been demonstrated on pMJ51 (Fig.24), but not on pMJ28 (Fig.17) (III.7). These plasmids differ only in that pMJ51 carries an additional 120 bp which lie upstream from the StuI site derived from the gene-proximal sequences that were cloned on pMJ2.

No promoter activity had been detected when a linker of 30 bp was present in the StuI site (pMJ42, Fig.23), or when cloned sequences upstream from StuI were tested on their own (pMJ33, Fig.23). It was therefore important to ascertain where the transcriptional start site lay relative to the StuI site, as it seemed that sequences to either side of the site were required for promoter activity. These apparently had to be correctly juxtaposed, since the presence of linker sequences in pMJ42 had eliminated the functional ability of the promoter.

The S1 mapping technique (Berk & Sharp,1977; principle shown in Fig.26a&b) determines where complementarity to specific RNA transcripts cease on a cloned DNA. It does not a priori reveal a transcriptional start site or terminator.

III.8.1 Isolation of RNA from B. licheniformis Expressing Subtilisin Carlsberg

To establish when the B. licheniformis cells began secreting subtilisin, MFJ29 was grown in modified Schaeffer's medium,

and the medium monitored for exoprotease production.

In the skim-milk plate assay (II.9.1.1) samples taken after T3 gave halos of proteolysis upon overnight incubation, with a large increase at T5. In the standard Azocoll assay however (II.9.12), the T6 samples showed no activity. At 20h of incubation (approx. = T15), a value of 0.09 A_{520} units was obtained per ml supernatant. This was only c. half as high as the exoproteolytic activity of B. subtilis reference strain MFJ93 cells grown in parallel.

Cells were thus harvested for RNA extraction during exponential growth (at A_{450}=1), at T6, and at T12.

Several methods of RNA extraction were tested. The convenient guanidinium-CsCl method described in Maniatis et al (1982) p. 196, (but using mechanical abrasion of frozen cells in a mortar with quartz grains) gave a poor yield, presumably due to inadequate cell lysis. Most recently, adaptations of this method to Bacillus have been published. In one procedure, a prolonged incubation with a high concentration of lysozyme was used (Paddon & Hartley, 1986). In another, the cells were opened by passaging twice through a French press (Ray & Haldenwang, 1986).

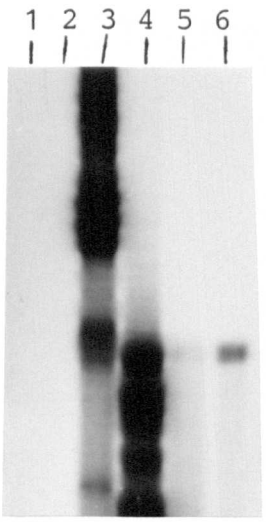
The method of Gilman & Chamberlin (1983) (II.6) gave high yields (c. 500 µg from 20 ml cells). Figure 25 shows the result of a Northern Blot (II.8.4), where the RNA preparations were probed with the 500 bp StuI-HindIII pMJ2 fragment (Figure 10b).

Subtilisin-specific mRNA was detectable only in the samples from T12 (lanes 5 & 6). The transcripts in lane 5 were only clearly visible after a longer exposure, and showed signs of degradation, presumably due to the poor DNAaseI preparation that had been used. From the protease plate assay results, specific transcripts had been anticipated at T6, but perhaps mRNA levels were too low at that time to be detected by the Northern blot method.

FIGURE 25

NORTHERN BLOT

MFJ29 total RNA preparations were probed with nick-translated pMJ2/ StuI & HindIII 500 bp fragment. RNA was extracted from cells harvested at different stages of growth as described in the text.
lane 1: exponential cells
lane 2: T6 cells
lane 3: λ/ BstEII labelled markers
lane 4: pBR322/ MspI labelled markers
lane 5: T12 cells, RNA treated with crude DNAase I
lane 6: T12 cells, RNA treated with Worthington RNAase-free DNAase I.



<- subtilisin-specific transcripts

The failure to obtain expression from plasmids pMJ33 and pMJ42 (III.7.4), had first led to expectations that the start site for transcription would be situated relatively far from the gene, and near the location of the rearrangements in pMJ2. By comparison of the mobility of the subtilisin-specific mRNA band with the DNA markers (Figure 25), which have been calibrated with other well-characterized mRNAs, a size of >1.2 but <1.4 kb was estimated (Esper Boel, personal communication).

Since the transcription terminator could be guessed at (Figure 14; III.6.4), the initiation site seemed to be in the region of the StuI site. The transcriptional start (and stop) sites were consequently ascertained.

III.8.2 Locating the Transcriptional Start Site

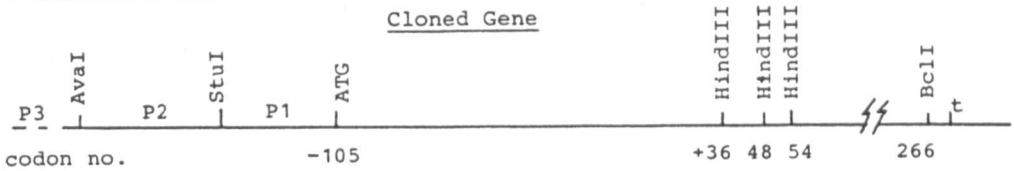
The subtilisin Carlsberg transcriptional start was thought to lie around the StuI site, from an estimation of the size of subtilisin-specific transcripts found in B. licheniformis (III.8.1).

The principle of the S1-nuclease mapping technique is shown in Figure 26a&b. The 3 possible locations of a start site relative to StuI and AvaI are considered. The formal possibility of an in vivo transcriptional initiation site lying further upstream than AvaI is included in Fig.26. This possibility had been excluded by the failure to bind to the Northern blot of 15-mer NOR193, which corresponds to the sequence on the coding strand adjacent to the AvaI site (see Fig. 14). The fairly discrete width of the RNA band which hybridized to the StuI-HindIII probe (Figure 25) also argues against several distantly spaced initiation sites.

In Figure 26(b) the broken line in each of the 3 cases depicts the size of the 5'-labeled probe fragment of the coding strand that will be protected from digestion by S1-nuclease in a mRNA -probe DNA hybrid. It can then be visualized on denaturing gels

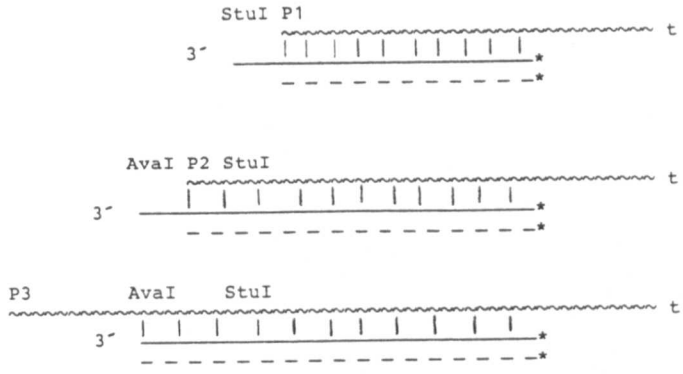
FIGURE 26: MAPPING TRANSCRIPTIONAL START SITE BY S1-NUCLEASE METHOD

A) POSSIBLE PROMOTER LOCATIONS



P1,P2,P3 = possible promoter locations; t = terminator

B) PRINCIPLE of S1 MAPPING



anneal labelled probe & mRNA together

S1 treatment of hybrids removes single-stranded regions

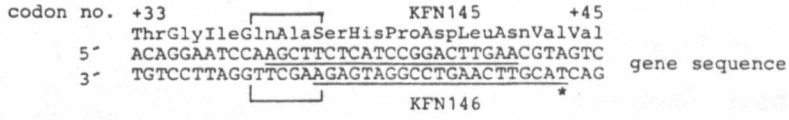
labelled products analysed on denaturing gel

~~~ mRNA |||| hydrogen bonds

——* labelled coding strand

– –* labelled protected strand after S1 nuclease digestion of RNA-DNA hybrid

C) LABELLING PROMOTER-PROXIMAL HindIII SITE (coding strand) with LINKERS SYNTHESIZED ACCORDING TO SEQUENCE

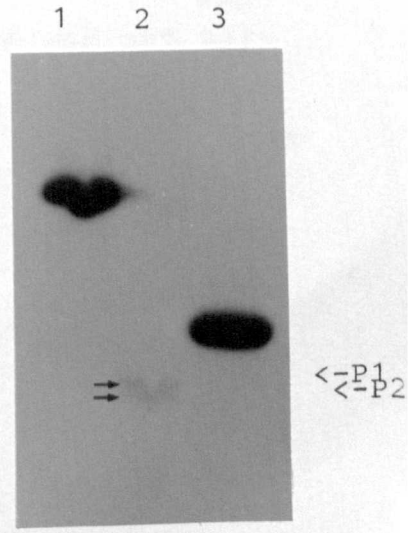


KFN145,146 =synthetic 21-mers,underlined
* indicates the labelled nucleotide
HindIII site boxed

D) RESULTS OF S1 MAPPING

Products of the S1 digestion were analysed on a 40cm. 8% acrylamide-urea gel, run until the XCFF dye had travelled 60cm.
lane 1: AvaI-HindIII-linker* probe
lane 2: probe*& T12RNA/ S1 digest
lane 3: StuI-HindIII-linker* size marker

Arrows P1 and P2 indicate the two protected fragments.


~~~

by autoradiography and its size compared with that of control fragments.

Apart from MaeI, which was not available when this experiment was started, for the purposes of labelling a suitable probe, the HindIII site in codon +36 was the closest, fully characterized, and convenient restriction site downstream from StuI.

pMJ31A (Figure 24) was used to prepare the probe. HindIII-digested pMJ31A was labeled by 2 standard methods. Kinase labelling (II.7.3) after phosphatase treatment (II.16), and labelling with T4 DNA polymerase (II.7.2.2) led to insufficient incorporation of counts into the DNA.

Satisfactorily high specific activity after labelling at the 3' end of the fragment was achieved by a different route, suggested by S. Hastrup (Figure 26c). Oligomers corresponding to the sequence including, and immediately after the HindIII site were synthesized. The 5' end of the oligomer representing the coding strand (KFN146) was kinased to high specific activity with γ-^{32}P-ATP, the other (KFN 145) with cold ATP. The complementary oligomers were annealed together (II.7.3) and ligated to the HindIII digested pMJ31A.

Labelled fragments were isolated from gels (II.17) after StuI or Asp718 digestion. The pUC18 Asp718 site was used in preference to AvaI as complete digestion (and hence good yields) could not be obtained with AvaI. Asp718 digestion of pMJ31A leads to an Asp718-HindIII fragment that is 6 bases longer on the bottom strand than is the corresponding chromosomal AvaI-HindIII fragment (Figure 24). Strand separation was not performed. Hybridization to the T12 RNA, and S1-digestion were carried out as described in (II.6). The hybridization temperature was 45^{o}C; at higher temperatures no "protection" of probe against S1 treatment was seen (not shown).

It is usual when "mapping" transcriptional start points to

determine the site with an accuracy of a few bases. This is done by running the products of chemical sequencing (II.14.1) of the labelled probe beside the S1-digestion products on an analytical denaturing gel (II.12.2, see next section). As the StuI-HindIII, and Asp718-HindIII fragments were respectively 555 and 674 bases long on the coding strand after ligation of the labelled oligomer, this was not practicable. An attempt to achieve a sequence "ladder" that extends so far up was a failure.

Figure 26d shows the result of the S1 analysis. From the distance travelled by the marker fragments, which differ in length by 119 bases, the size of the 2 strong protected bands (shown with arrows) can be approximated. Mapped sites were estimated to lie respectively 49 and 63 bases downstream from StuI, but it must be stressed that this is at best an approximation. The fact that these large protected bands show up discretely suggests that their end-points do indeed lie further apart than the S1-generated "spread" of a few bases caused by "nibbling", that is often seen (Hentschel et al, 1980).This suggests that there are two transcripts of different lengths.

The estimated positions of P1, P2 are shown in Figure 14, and lie near the promoter -10 regions postulated in III.6.3. They lie about 60 and 46 bp resp. upstream from the first ATG codon (Figure 14).

This result is in line for those obtained for other Bacillus genes: the B. subtilis ISP-I, α-amylase, and B. stearothermophilus β-galactosidase transcriptional start sites are located 44, 30,and 28 bases from their respective translation initiation codons (Koide et al, 1986; Lehtovaara et al, 1984; Hirata et al,1986). In B. licheniformis β-lactamase, the start site is found rather further away, viz. 166 bases from the first codon (McLaughlin et al, 1982).

In the case of the B. subtilis subtilisin gene, the two trans-

cription start sites mapped 15 bases apart, and lie at the end of the two "typical" sigma-37 -10 regions that had been identified by sequence inspection (Fig 33; Wong & Doi,1984). They lie respectively 19 and 33 bases upstream from the initiating GTG codon.

Thus the subtilisin Carlsberg gene, in respect of multiple promoter sites, their close apposition to the initiation codon, and in including a region of dyad symmetry about the promoter (II.6.3, Figure 14), seems to resemble the homologous gene from B.subtilis (Fig. 33). However, the determination of the transcriptional start site by another method gave results that conflict somewhat with these (see below, Appendix 3, and Conclusions).

It seems to be quite a common feature of highly regulated genes to be transcribed from two different start sites. This was discussed fully in I.3.

Three options are available for determining the exact subtilisin Carlsberg start points, now that the approximate locations of P1 and P2 are known:

(1) the NlaIII site lying 157 bp downstream from StuI (Appendix 2; Fig.14, positions 414-417) could be labeled with synthetic linkers as performed above. The S1 mapping could thus be repeated and chemical sequencing ladders would reveal where protection against S1 digestion by mRNA ceases.

(2) Complementary DNA could be made with reverse transcriptase on the mRNA starting with a 5'-labelled oligomeric primer synthesized according to the coding strand sequence (Ghosh et al, 1978). Two labelled strands, differing in length probably by about 10 to 15 bases, should be seen. Their lengths could be determined by comparison with chemical sequencing ladders of a known sequence. Alternatively, the cDNA could be isolated and sequenced directly (This has in fact been carried out. The re-

sults are shown in Appendix 3).

(3) The T4 DNA polymerase method of Hu & Davidson (1986) could be applied. A single-stranded probe covering the coding strand in the region upstream from the promoter is hybridized to mRNA. Primed T4 DNA polymerase synthesis extends from a site upstream only as far as the hybrid; it is blocked at the first nucleotide of the mRNA. The length, or sequence, of the newly synthesized strand is determined.

It is especially necessary to locate the transcriptional initiation sites precisely, in order to help assign roles to the palindromes in the promoter region (II.6.3), and to extend our knowlege of Gram-positive promoters.

III.8.3 Mapping the Transcriptional Terminator

Terminator mapping is based on the same principle as promoter mapping (Figure 26b), but the problem is reversed. The coding strand is labeled at its 3' end instead, and the size of the labeled S1-protected fragment in the mRNA-DNA hybrid measured. This is obviously determined by how far the mRNA extends in the 3' direction.

A 152 bp BclI-BamHI fragment, which contains the 8 C-terminal codons of the subtilisin Carlsberg gene was labeled at the BclI site by" filling-out" with Klenow enzyme (II.7.2.1) and isolated from pMJ26 after BamHI digestion (Fig. 17).

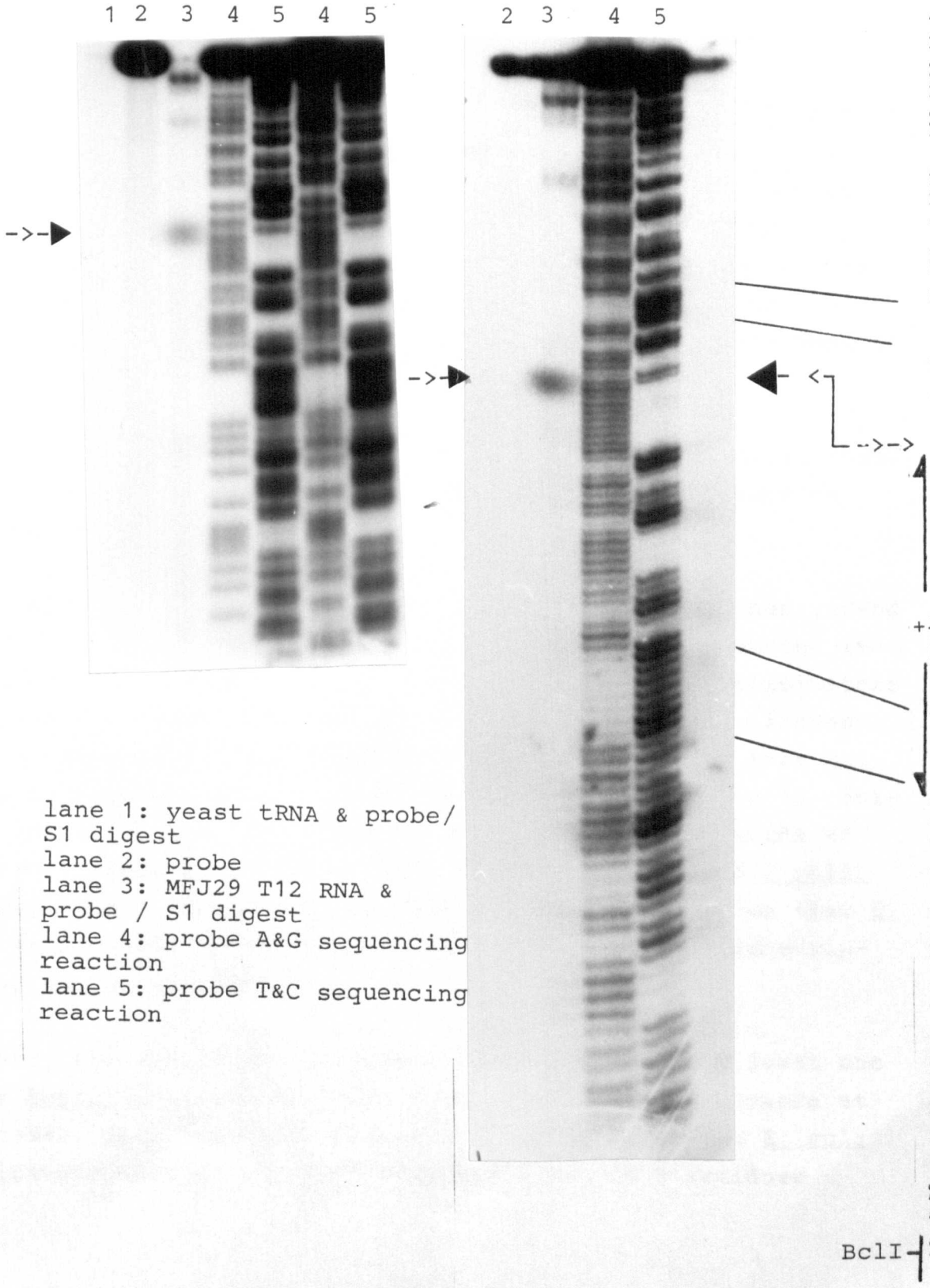
Hybridization at 66°C to MFJ29 T12 RNA, and S1 digestion were performed as described in II.6, and the protected fragment was examined on an 8% acrylamide-urea gel (II.12.2).

Figure 27 shows the result of this analysis on a 20 cm. (Fig. 27a) and a 40 cm. gel (Fig. 27b). There was slight cross-contamination from one track to another in the latter. A strong diffuse protected band is seen (Figure 27, heavy arrow).

FIGURE 27: LOCATION OF THE TRANSCRIPTIONAL TERMINATION SITE

Autoradiographs of S1 digestion products on 8% denaturing gels. The probe was a 152 bp BclI-BamHI fragment from pMJ26. The BclI site was labelled. The sequence of this fragment is shown alongside 27(B). The terminator palindrome is shown with divergent arrows, its axis with a cross. The termination site is shown with small arrows.

(A) 20 cm. Gel (B) 40 cm. Gel



lane 1: yeast tRNA & probe/ S1 digest
lane 2: probe
lane 3: MFJ29 T12 RNA & probe / S1 digest
lane 4: probe A&G sequencing reaction
lane 5: probe T&C sequencing reaction

By comparison with the (A+G) and (T+C) chemical sequencing ladders obtained with the probe (II.14.1), it has a mobility between those of the labelled fragments ending just before T and G as indicated in the known sequence (see Figure 14), which is displayed beside the gel . The width of this band is about 2 nt, and is probably due to the ragged ends ("nibbling") commonly produced by S1-nuclease (Hentschel et al, 1980).

The mobility of S1-digestion products (which have 3'-OH ends) must be corrected when compared with products of chemical cleavage (which have 3'-P ends). The absence of a phosphate group reduces mobility by about 1/2 nt. The chemical products also migrate faster relative to their enzymic counterparts as they lack both the base and the sugar moeity of the base that is "read" in the sequence. Thus an adjustment of 1.5 bases must be made (Hentschel et al, 1980; Brosius et al, 1982).

When this is done, the limit of the protected fragment is found to lie just at the distal end of the palindrome remarked on previously (II.6.4).

The "hairpin" structure that the DNA/mRNA may adopt thus indeed functions as a transcriptional terminator. Such hyphenated regions of dyad symmetry are typical of Gram-negative terminators (Rosenberg & Court,1979; Holmes et al, 1983). From an inspection of several 3' gene flanking sequences, it seems that Bacillus terminators typically have a palindrome of 11 to 15 residues (Wells et al, 1983; Yang et al, 1984, 1986; Vasantha et al, 1984; Stahl & Ferrari, 1984), which is similar to E.coli. In keeping with these observations is the demonstration that E. coli terminators can be recognized by the Bacillus transcriptional machinery (Peschke et al, 1985).

However, the subtilisin Carlsberg terminator, and at least one other Bacillus terminator that has been mapped (Lehtovaara et al, 1984), lack the other features of rho-independent E. coli terminators, such as a G-rich stem and a run of T residues

immediately after the palindrome.

In contrast, transcripts do end in a position corresponding to a run of Ts immediately following a palindrome in one B.stearothermophilus terminator that has been studied (Waye & Winter, 1986), and in a stretch of Ts forming the distal end of a palindrome from B. amyloliquefaciens (Kallio, 1986).

III.9 Alternative Strategy for Cloning Genes with Strong Promoters

Problems have been experienced in various organisms in the cloning of genes that are either very actively transcribed, or whose product is lethal to the host.

In the former category, vector stability is presumably compromised by the high level of transcription (Brosius, 1984). Examples of such strongly transcribed genes include some from *S. aureus*, which were successfully cloned in *B. subtilis* by chromosomal integration, but not on multicopy plasmids (*blaZ*, Saunders et al, 1984 a&b; *spa*, Fahnestock et al, 1986). The *B. licheniformis* β-lactamase from an overproducer could be cloned in *B. subtilis* on a low copy plasmid, but not on pUB110 (Imanaka et al, 1981).

An alternative to keeping the copy number of the cloned gene low is to use a promoter weaker than the gene's own to stabilize plasmid clones. Examples of successful application of such a strategy in *B. subtilis* are *S. aureus spa* (Fahestock & Fisher, 1986), and *B. licheniformis* α-amylase (Sibakov & Palva, 1984; Sibakov, 1986a). A promoter down-mutation can also render stable cloning possible where it previously had not been eg. *S. pneumoniae mal* genes in *E. coli* (Stassi & Lacks, 1982) and in *B. subtilis* (Espinosa et al, 1984).

Placement of strong terminators downstream from the highly expressed gene can enhance stability in *E. coli* (Gentz et al, 1981).

Where a gene product is lethal to the host, various elegant approaches have been successfully applied to either regulating, or eliminating, phenotypic expression, and so making cloning possible.

For example, the lambda cII gene could be cloned in E. coli only when it was under the control of the repressed P_L promoter in a lambda lysogen (Shimatake & Rosenberg, 1981). The ability to regulate expression was also useful in the case of a cloned vesicular stomatitis virus glycoprotein gene, which was under trpE control. The clone was stable in $trpR^+$, but not in derepressed $trpR^-$ cells (Rose & Shafferman, 1981).

Potentially toxic Bacillus spp.-derived genes have been cloned instead as the suppressor-sensitive mutant allele (B. subtilis spoIIAC in E. coli sup^o strain: Yudkin, 1986) or as a Tn917-insertionally inactivated copy (from an RNAase-negative transposition product)(B. amyloliquefaciens extracellular RNAase gene in E. coli: Paddon & Hartley, 1986).

In the case of the B. licheniformis subtilisin gene it was unclear why stable cloning of "upstream" sequences in E. coli was problematical. The level of expression of the endogenous subtilisin seen in the wt B. licheniformis host (III.8.1), and in cloned, multicopy form in B. subtilis MFJ246 (III.7.5) suggest that toxicity factors other than excessive promoter activity may play a role. However, promoter-proximal sequences had undoubtedly been lost in the cloning process (see Conclusions).

It would in any case have been desirable to be able to either insertionally inactivate the gene in B. licheniformis, or otherwise alter it in vitro, prior to (re)-cloning in E. coli, or B. subtilis, possibly with the help of strong terminators (see IV.6.1,-6.2 & -6.7 for a full discussion of the chromosomal manipulations involved).

As the gene coding sequences were available, and the mechanisms involved are well understood for B. subtilis, these options were in theory open, assuming that interaction between chromosomal and incoming homologous sequences takes place in B. licheniformis in similar fashion.

Hybrid plasmids carrying intact or *in vitro* modified subtilisin sequences, but which are unable to replicate in *B.licheniformis*, could be targeted to the chromosomal locus, and integrated there (Niaudet et al,1982; Ferrari et al,1983). Appropriate *in vivo* recombinations, or *in vitro* excision of sequences including the integrated plasmid, followed by propagation in *E. coli*, would respectively yield a strain with an inactive subtilisin gene or a plasmid clone carrying the desired sequences.

The targetting, integration excision approach has been successfully applied to the cloning of the *gyrA* gene, *terC* and *B.subtilis* subtilisin flanking sequences (Lampe & Bott 1984; Smith et al, 1985; this thesis, IV - resp.).

However, since in my hands transformation of *B. licheniformis* was not successful (see *Appendix 4*), these approaches were abandoned in favour of the development of an integrable, sensitive, promoter-cloning vehicle for *B. subtilis*, pMJ50. This is described below, and is of general applicability provided that species-specific factors are not required for detectable promoter activity.

A recent report describes the construction of a similar system, but using phage Ø105 as the vector, and a promoterless amylase gene (Seki et al, 1986).

III.9.1 Construction of pMJ50, the Promoter Cloning Vector

This approach involved the synthesis of a plasmid carrying a Gram-negative origin of replication, a selective marker which could be expressed in both *E. coli* and *B. subtilis*, a promoterless gene preceded by a unique cloning site, whose gene product is easily screened for in *B. subtilis*, whose expression is dependent only on the activity of a promoter which can be inserted at the unique site and, a *B. subtilis*-derived gene was included to provide homology with the chromosome and thus target the nonreplicable plasmid to an integration site. As a precaution strong *E. coli* transcriptional terminators were cloned

behind the promoterless gene.

The resulting plasmid, pMJ50, is depicted in Figure 28. Note the stop codons in all 3 reading frames before the xynB RBS. (Not all restriction sites are shown).

The component parts of the plasmid were derived as follows. The ori was derived from pBR322. The CAT gene served as selective marker for B. subtilis transformation. The xylosidase gene of B. pumilus from which the promoter had been removed leaving a unique ClaI cloning site was kindly provided by Sven Hastrup. The homologous subtilisin gene was used to allow integration into the chromosome. The E. coli 5SRNA, which is followed by 2 strong terminators, which also function in B. subtilis, was utilized.

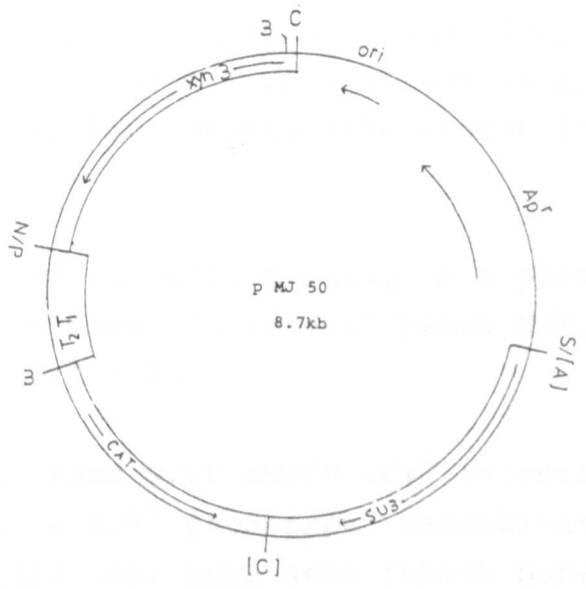
To facilitate the construction, a pUC18 derivative (pMJ43) with (inter alia) ClaI and StuI sites in the polylinker was used. After linearization with ClaI and StuI, the vector was treated with phosphatase to reduce recircularization (II.16).

Plasmid pMJ15 (IV, Fig.36a) was modified by filling out the unique ClaI site located between subtilisin and CAT genes to permit isolation of a BamHI-AccI fragment from the modified pMJ15[ClaI].

After ligation of the 4 fragments indicated in the diagram, E. coli competent cells were transformed and Ap^R transformants selected. These were replica-plated onto fresh Ap and Cm plates containing repectively 100, and 6 µg/ml of antibiotic. A high proportion of transformants were Cm^R.

Plasmid was extracted from several such colonies and the construction verified with 5 different restriction enzymes as that indicated in Fig.28.

FIGURE 28: STRUCTURE OF pMJ50, AN INTEGRATIVE, SINGLE COPY PROMOTER CLONING VECTOR



xynB	B.pumilus xylosidase gene (promoterless) on 2kb ClaI-NsiI fragment from pPX7
T1T2	E.coli 5S RNA gene and terminators on 650bp PstI-BamHI fragment from pSX68
CAT	chloramphenicol transacetylase gene
SUB	B.subtilis subtilisin gene CAT & SUB on 3.6kb [AccI] -BamHI fragment of pMJ15 [ClaI]
———	pMJ43 moeity= pUC18 derivative with altered polylinker. Linearized with StuI and ClaI.
C	ClaI unique promoter cloning site
B	BamHI
[X]	filled-out site X
N/P	NsiI/PstI junction
S/[A]	StuI/ [AccI] junction

SEQUENCE DOWNSTREAM FROM UNIQUE ClaI SITE SHOWING STOP CODONS BEFORE xynB INITIATION CODON (S. Hastrup)

ClaI

ATCGATCATATGGATGATTTTAAAATGTAGGTCGTGAAAGAGGAGGAAGTTGAATG

XynB
+1

III.9.2 Testing pMJ50 by Cloning a Characterized Promoter Fragment into the ClaI Site

Xylosidase is produced by B. subtilis wt only in response to the presence of its substrate xylose. A B. subtilis mutant has been constructed which has a 4 bp deletion in xynB, and in which xylosidase activity consequently cannot be induced (SHa28 = MFJ210).

Xylosidase production in cells growing on a plate can be detected with a very sensitive fluorescent probe MUX, which is sprayed on to the cells (II.9.2).

MFJ210 cells transformed with pMJ50 (Cm^R selection) should not, and do not, display a Xyl^+ phenotype. Recombination between the defective B. subtilis host xynB gene (which nonetheless has a functional promoter) and the plasmid-borne B. pumilus xynB, which lacks a promoter, is unlikely to occur since insufficient homology exists between these genes. No conversion of the Xyl^- MFJ210 to a Xyl^+ phenotype has been observed among the several thousand pMJ50 transformants screened.

The copy number of pMJ50 should be one, or close to one, per chromosome, especially if the antibiotic pressure is removed after establishment of the transformants. Loss of the integrated plasmid should not occur (Yang et al, 1986; Janniere et al, 1985). If a promoter is cloned into the ClaI site of pMJ50, expression of xylosidase will depend on the characteristics of that promoter.

The α-amylase promoter was used to test the pMJ50 system. Non-induced expression of xylosidase would thus be expected in stationary phase (III.7.3).

The c. 180bp B. stearothermophilus α-amylase promoter was isolated from pMJ32 (see Fig. 18). The plasmid was linearized at the EcoRI site, the ends filled out and ClaI linkers ligated on. Digestion with ClaI thus freed the promoter from the rest

of the plasmid on a small ClaI-ended fragment. The digestion mixture was ligated to pMJ50 that had been linearized with ClaI. E. coli MFJ233 cells were transformed, and plasmid was prepared from total Ap^R transformants (this passage through E. coli was performed in order to generate covalently closed, circular, amplified, DNA for the relatively inefficient integration step). This DNA was used to transform MFJ210. Cm^{6R} transformants were sprayed with MUX and visualized under UV light. B. subtilis wt (on plates +/- xylose) and MFJ210(pMJ50) served as controls. Negatives (wt; MFJ239 = MFJ210(pMJ50); most of the transformants) appeared yellow. The positives (wt on xylose, a few MFJ210 transformants) appeared as blue fluorescent colonies. One such Xyl^+ transformant (MFJ240) was reisolated (in the absence of Cm, in order to lower copy number to c.1).

MFJ240 was checked to ascertain whether the xylosidase activity detected in the MUX assay was due to recombination of pMJ50 with the defective xynB locus, which could result in a functional xynB gene, or due to the cloning of the α-amylase promoter in front of the xynB promoterless gene, as intended.

If recombination within xynB had taken place, the xylosidase activity would be dependent on xylose induction, and be detectable during vegetative growth, as it is in the wild-type (cf. Fig. 22; kinetics of xylosidase induction in MFJ93). If, on the other hand, the cloned amylase promoter was responsible for the Xyl^+ phenotype of MFJ240, the xylosidase activity would be expressed (only) in stationary phase, even in the absence of xylose.

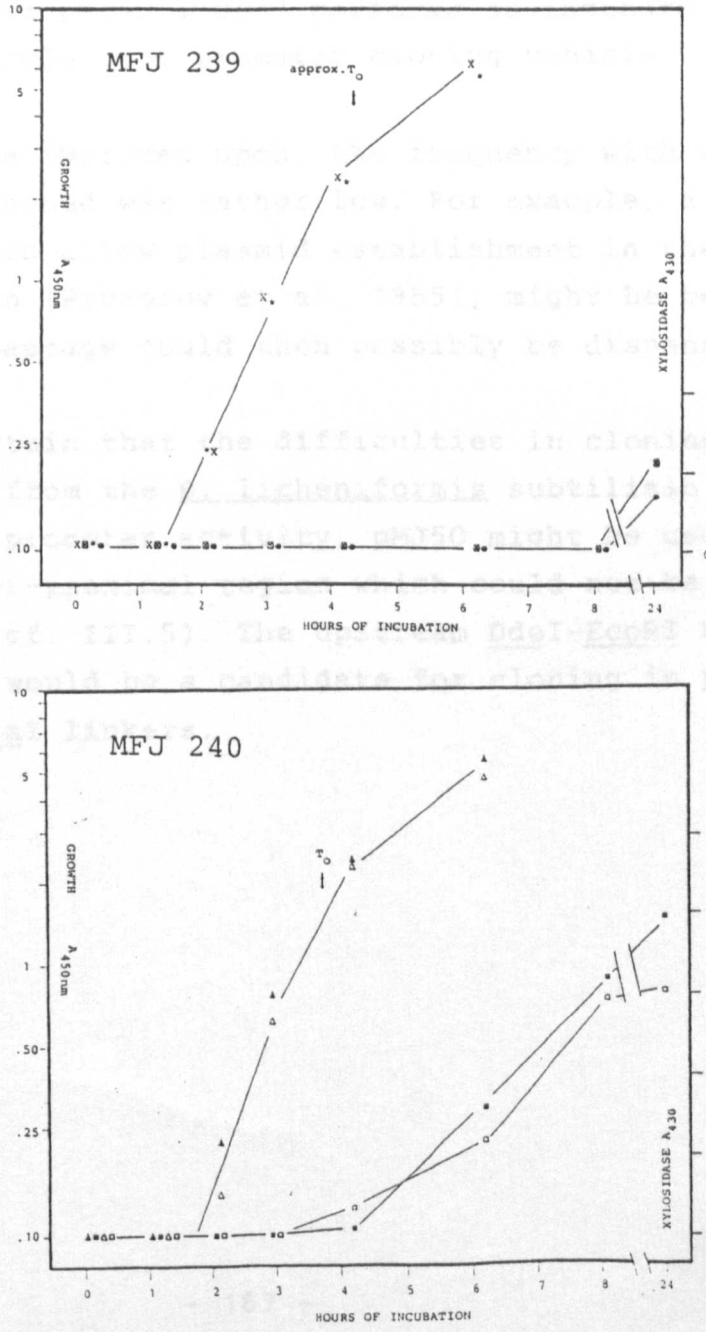
The Xyl^+ MFJ240 transformant was consequently grown in L Broth in the presence of, and absence of 0.2% (w/v) xylose. Samples of culture were taken regularly and assayed for xylosidase activity with PNPX (II.9.2). As a control, the Xyl^- MFJ239 = MFJ210(pMJ50) was treated in parallel. The results (Fig. 29) show unequivocally that xylosidase activity, regard-

FIGURE 29: XYLOSIDASE GENE COMES UNDER CONTROL OF PROMOTER CLONED INTO UNIQUE ClaI SITE OF INTEGRATIVE PROMOTER CLONING VECTOR pMJ50

The cells were inoculated 1:100 from a fresh LBroth overnight culture and grown in LBroth with or without 0.2% xylose. Samples were taken and assayed for xylosidase activity with PNPX (see II.9.2), for 1.5 hours at 37°C, at various times as indicated.

	- XYLOSE		+ XYLOSE	
	GROWTH	XYLOSIDASE	GROWTH	XYLOSIDASE
	A_{450}	A_{430}	A_{450}	A_{430}
MFJ239	X	⊠	*	●
MFJ240	Δ	□	▲	■

MFJ239 = MFJ210::pMJ50; MFJ240 = MFJ210::pMJ50+ P_{amy}



less of the presence or absence of xylose, is observed only in stationary phase with MFJ240. The control MFJ239 displays no xylosidase activity under any of the conditions tested. The small amount of colour formation in the 24 hour samples may be attributed to degradation of PNPX by other enzymes, or by spontaneous breakdown of the substrate.

It has not yet been confirmed by Southern hybridization that integration of pMJ50 was at the subtilisin locus, and that copy number was close to 1, but experience with similar plasmids make it most likely (IV).

It was concluded that pMJ50 indeed performs as intended, and can be used as a single copy promoter cloning vehicle.

The plasmid could be improved upon: the frequency with which integrants were obtained was rather low. For example, a ts origin of replication to allow plasmid establishment in the cell prior to integration (Prozorov et al, 1985), might be beneficial. The E. coli passage could then possibly be dispensed with.

While it is not certain that the difficulties in cloning the sequences upstream from the B. licheniformis subtilisin gene were due to strong promoter activity, pMJ50 might be useful for cloning the promoter-proximal region which could not be isolated in this study (cf. III.5). The upstream DdeI-EcoRI 0.9 kb fragment (Fig.11b) would be a candidate for cloning in pMJ50 with the help of ClaI linkers.

Chapter IV: Cloning a Gene for Subtilisin NOVO/BPN' and Generation of a B. subtilis Subtilisin Deletant

IV.1 Why the Gene for Subtilisin BPN' was Cloned

The cloning of this gene was undertaken for several reasons. There was the then outstanding question of subtilisin's role in sporulation to be resolved (see I.2.2.4). Modification of the properties of the enzyme has been a goal for many academic and commercial groups (see I.2.2.3 & -4).

However, the main motivation here was to use the cloned gene, altered *in vitro*, to generate a defined subtilisin deletion mutant of *B. subtilis* (specifically, of the Novo Laboratories' reference strain DN497 = MFJ93). This would be one step towards creating a well-defined low-protease producing host which could be used for the synthesis of foreign gene products. An attempt to make such a mutant had been carried out in these laboratories using the cloned subtilisin gene from *B. amyloliquefaciens* but had failed (personal communication, B. Diderichsen). The 85% homology that exists between the subtilisin coding sequences from the 2 strains (Wells et al, 1983) apparently was insufficient for the required recombinational events to occur. The fully homologous gene had thus to be cloned.

At the time the work was embarked upon, reports of the cloning of the *B. amyloliquefaciens* gene were available (Wells et al, 1983; B.Diderichsen, personal communication), and there had been mention of the cloning of the *B. subtilis* gene, though its sequence had not yet been published (Wells et al, 1983). It was becoming clear that the cloning of the subtilisin gene from *B. licheniformis* was not following the straightforward progress in my hands that others had experienced with other *Bacillus* spp., so it was decided to clone the gene from *B. subtilis*.

Note: While it is clear that *B. subtilis* 168 synthesizes an Amylosacchariticus enzyme (Fig.1), this will be referred to here as NOVO/BPN' type. The enzyme belongs to the BPN' immu-

nological group, and has one additional AA at position +56, compared with the Carlsberg-type enzymes (see also I.2.2.1).

IV.2 Outline of Cloning Strategy

As described below, specific fragments were identified in Southern blotting experiments, and cloned on a high copy number plasmid in E. coli. Colonies were screened with a specific probe. The cloned fragment was transferred to a Gram-positive vector, then transformed into B. subtilis, where the cloned subtilisin gene was expressed.

IV.3 Chromosomal Mapping: Selection of Subtilisin-Specific Fragment for Cloning

Chromosomal DNA was prepared from B. subtilis MFJ93 and digested with enzymes for which cloning sites existed in the Gram-positive and -negative vectors that were readily available .

IV.3.1 Use of Heterologous Probes

In an initial blotting experiment, DNA from B. subtilis and B. licheniformis was probed with a fragment from the cloned B. amyloliquefaciens subtilisin gene (kindly donated by Børge Diderichsen). As the hybridizing fragments in the B. licheniformis digests were of similar size to those previously identified using the Carlsberg-specific oligomers, it was clear that the homology between the 3 strains was great enough to allow the use of either a B. amyloliquefaciens- or a B. licheniformis-derived subtilisin fragment to probe for the B. subtilis subtilisin. In later experiments these fragments were used interchangeably (see Fig. 30a).

B. amyloliquefaciens is more closely related to B. subtilis than is B. licheniformis: at any rate, with respect to GC content and the type of subtilisin the strain secretes (Marmur et al, 1963; Welker & Campbell, 1967; Keay & Moser, 1969). The situation is, however, reversed with respect to α-amylases. amyL and amyA are homologous, but unrelated to the amylase of B. subtilis (Gray et al,1986). It had been mentioned in Wells et al

(1983) that the B. subtilis and B. amyloliquefaciens subtilisin genes show c. 85% homology throughout the coding sequences, including the pre-pro region.

An upstream fragment from the B. amyloliquefaciens clone was used as a "5' probe". This ClaI 1.5 kb fragment is indicated in Fig.30a and covers a large region upstream, the prepro-, and mature sequences up to the codon for AA +35. It thus carries approximately 26% subtilisin coding sequences. However, when the fragment was tested in hybridization against similar amounts of B. amyloliquefaciens and B. subtilis DNA, the signal from the latter was disproportionately weak, and the use of this probe was abandoned.

This was surprising in the light of the information given above, and probably accounts for the failure to obtain a subtilisin deletion mutant in B. subtilis using material from this B. amyloliquefaciens clone. Southern blotting revealed that this clone differed in many respects from that reported by Genentech (Wells et al, 1983), including the relative disposition of EcoRI, BamHI, ClaI and PvuII sites around the gene (data not shown). These differences must be attributed to B. amyloliquefaciens strain variation. The cloned fragments utilized here were shown to correspond to the chromosomal sequences by restriction mapping.

IV.3.2 Preliminary Chromosomal Map

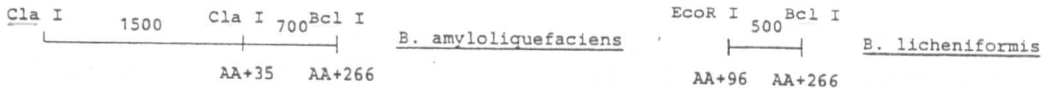
With the first blot it was established that while EcoRI and ClaI alone gave probe-specific fragments of 6.9 and 3.5 kb respectively, a double digest showed only a 2.6 kb fragment, thereby establishing the close proximity of EcoRI and ClaI to a gene of estimated size around 1.5 kb.

Attention was therefore focussed on double digests of each of these 2 enzymes in combination with another, and a provisional restriction map was built up. It was then tested with the appropriate combinations of other enzymes. A schematic repre-

FIGURE 30: RESTRICTION MAPPING THE B. subtilis SUBTILISIN LOCUS

(A) HETEROLOGOUS PROBES

The size, in bp, of each fragment is shown, as well as the region of the polypeptide it encodes (AA number).



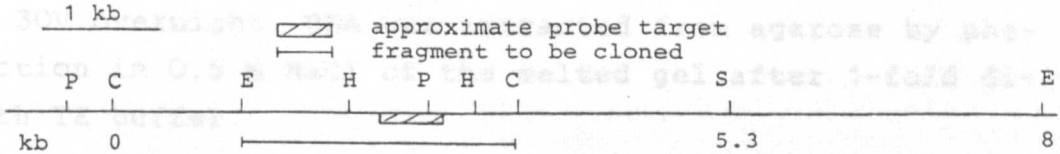
(B) RESULTS OF SOUTHERN ANALYSIS

ENZYMES	Size of specifically hybridizing fragments (kb)
Sph I	very large
EcoR I	6.9
Cla I	3.5
Bcl I	4.2
Pvu II	3.1
Hind III	1.1
EcoR I + Sph I	4.2
EcoR I + Cla I	2.6
Cla I + Pvu II	2.5
EcoR I + Pvu II	1.7
Hind III + Pvu II	0.7 + 0.4 (faint)

Furthermore,
the EcoR I fragment was not cut by BamH I, Pst I or Sal I
the Cla I fragment was not cut by BamH I, Pst I, Sal I or BclI
the Hind III fragment was not cut by EcoR I, Cla I, or Bcl I

(C) RESTRICTION MAP OF SUBTILISIN LOCUS

These results were combined to give the following coherent map



Note: The Bcl I sites were not placed. The relative positions of the Hind III sites with respect to Pvu II were not certain, but the relative intensity of the small Pvu II-Hind III bands as well as the fact that only one band is seen with Cla I + Pvu II was suggestive.

sentation of all the results, together with the deduced map is shown in Figs.30b & c resp.

The fact that PvuII evidently cut within the area covered by the probe fragments, and the distance from ClaI, and EcoRI, to PvuII meant that the ClaI-EcoRI 2.6 kb fragment chosen to be cloned could carry sequences long enough to include the whole gene regardless of whether the larger portion of the probe annealed to the right of the PvuII site, rather than to the left as it has been drawn in Fig. 30c. The relative intensities of the 2 PvuII-HindIII fragments, and the fact that in the ClaI & PvuII digestion only a 2.7 kb band was seen, account for the positions of the HindIII sites, and the probe target area relative to PvuII being assigned as in Fig. 30c.

IV.3.3 Details of the Cloning Procedure

The procedure is shown diagrammatically in Fig. 31a.

DNA fragments: B. subtilis MFJ93/ClaI&EcoRI 2.6 kb fraction.
Vector : pBR322/ClaI & EcoRI (ClaI-PstI-EcoRI)

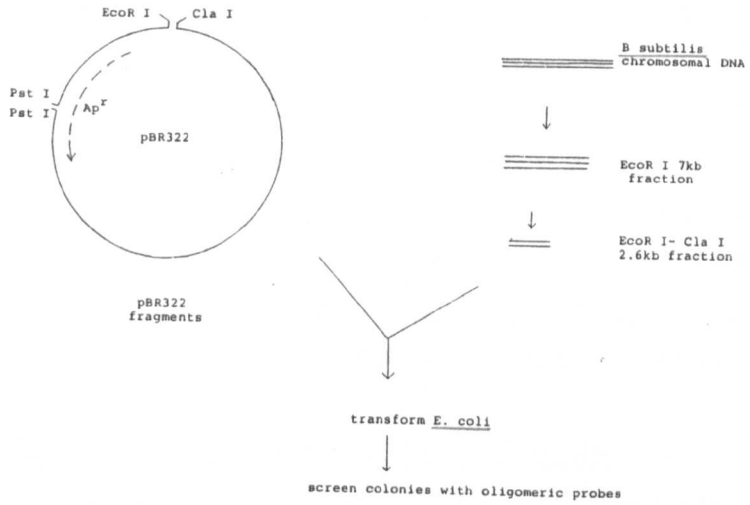
The chromosomal DNA was size-fractionated twice in order to enrich for the desired sequences: 200 µg DNA was digested with EcoRI after which a fraction around 7 kb was isolated. Once these fragments had been purified and digested with ClaI, fragments of approximately 2.6 kb were isolated. The size-fractionation was performed on 1% low melting temperature agarose gels. About 20 µg of DNA was loaded per slot, and the gels run slowly at 30V overnight. DNA was extracted from agarose by phenol extraction in 0.5 M NaCl of the melted gel after 1-fold dilution with TE buffer.

The survival of probe-hybridising sequences was monitored throughout the work-up and the yield estimated to be about 3 µg DNA, of fragment size 2.3-3.5 kb.

To eliminate background transformation by non-recombinant plasmid, the pBR322 vector DNA was prepared as two fragments, which

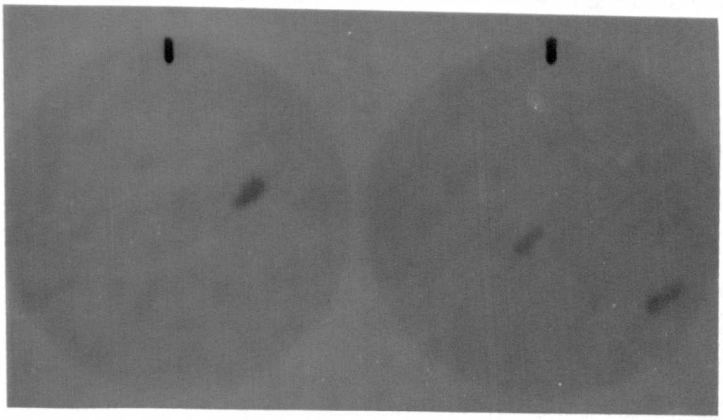
FIGURE 31: CLONING OF B. subtilis SUBTILISIN GENE

(A) CLONING PROCEDURE

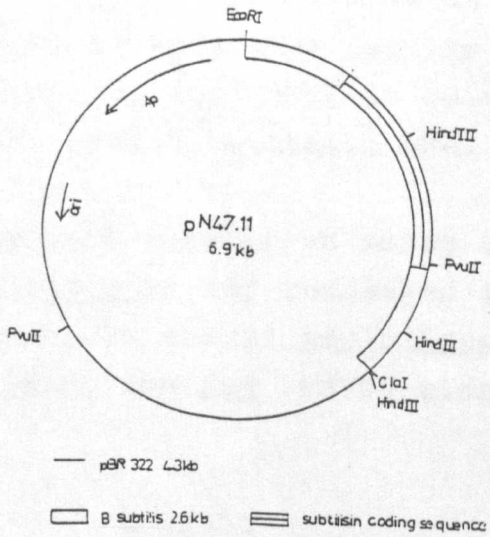


(B) HYBRIDIZATION -POSITIVE COLONIES

All the colonies that hybridized to probe were found to contain the cloned 2.6kb subtilisin fragment.



(C) SUBTILISIN CLONE pN47.11



would only be able to circularize during ligation by the inclusion of a donor fragment (Fig. 31a). Consequently, 2 equal portions of pBR322 DNA were digested with PstI & EcoRI, PstI & ClaI and the 752 bp (PstI & EcoRI) and 3586 bp (PstI & ClaI) fragments isolated on agarose gels and electroeluted.

Approximately equal amounts of donor and vector DNA were ligated, and MFJ92 competent cells transformed. 10 of the 2600 transformants obtained were picked at random. Plasmid DNA was prepared from them and restriction digestion with PstI, using pBR322 as a control, showed that at least 9 were recombinants.

IV.3.4 Results of Screening for Subtilisin Clones

The transformants were screened by colony hybridization. Approximately 1000 colonies were picked out in a grid on fresh plates, 100 per plate, and replica-plated. Colonies from one plate of each pair were transferred to Whatman nr. 540 filters and processed as described in II.8.2. A satisfactorily low background could not be achieved in my hands using the heterologous subtilisin gene fragments that had been labelled by nick-translation as probes. Instead, kinased oligomers that had been used to screen for the B.amyloliquefaciens gene (by B.Diderichsen), were used. The oligomer mix was as shown below and corresponds to all permutations for the coding strand from a position corresponding to AA+117 to +121:

$$5'- AC^{A}_{G}TCCAT^{A}_{G}TT^{A}_{G}TT - 3'$$

This would have a T_D of 34-40°C. An incubation and wash temperature of 28°C was shown to give good results both using the B. amyloliquefaciens clone as a control in colony hybridization, and in Southern blotting with B. subtilis DNA.

Seven positive colonies were identified among the transformants with the oligomer (e.g.Fig.31b) and rechecked in colony hyridization . As positive controls the B. amyloliquefaciens clone in both a dam$^+$ E. coli strain, and dam$^-$ MFJ92 cloning host was

used. The B. licheniformis clone in the dam$^-$ MFJ92 host served as a negative control.

This showed clearly that the signal from plasmid in the dam$^-$ host was poorer than from the dam$^+$ wild-type (not shown). 24h amplification on chloramphenicol was essential for any signal to be seen at all from the mutant. Had I been aware of this at the outset such a strain would not have been used as the cloning host. Presumably the poor hybridization is due to inadequate lysis. There are anecdotal reports of poor plasmid yields from dam$^-$ strains which could also be attributed to poor lysis. In my hands though, plasmid yields from dam$^+$ and dam$^-$ strains are comparable. Upon rechecking, the seven transformants were again positive. Plasmid was prepared from all seven and analysed in restriction digests.

Five of the seven were identical, and restriction patterns conformed exactly to what had been deduced from Southern blotting experiments. The two remaining clones contained, in addition to the desired subtilisin fragment, other fragments.

The location of the 2 HindIII sites relative to PvuII was determined by excision of an EcoRI-PvuII band followed by HindIII digestion and determination of the sizes of the PvuI-HindIII products. Neither of the 2 BclI sites that could not be located on the chromosomal map was found to lie between EcoRI and ClaI (data not shown). One clone, containing pN47.11, was chosen for further experiments. A map of this plasmid is shown in Fig.31c.

IV.4 Sequencing the B. subtilis MFJ93 Subtilisin Gene

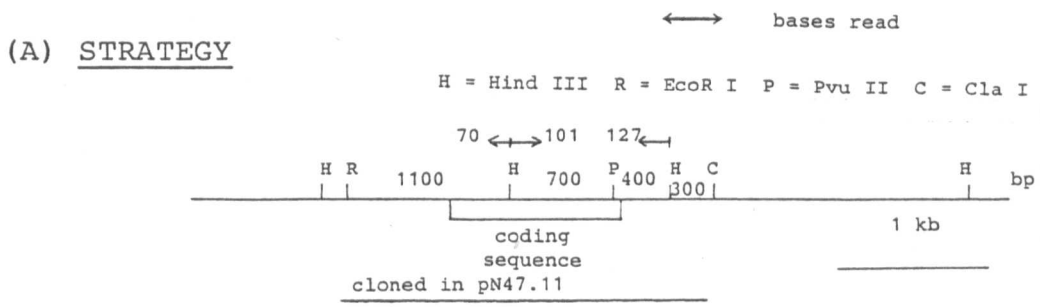
IV.4.1 Sequencing Strategy & Results

The sequencing strategy is illustrated in Fig. 32a.

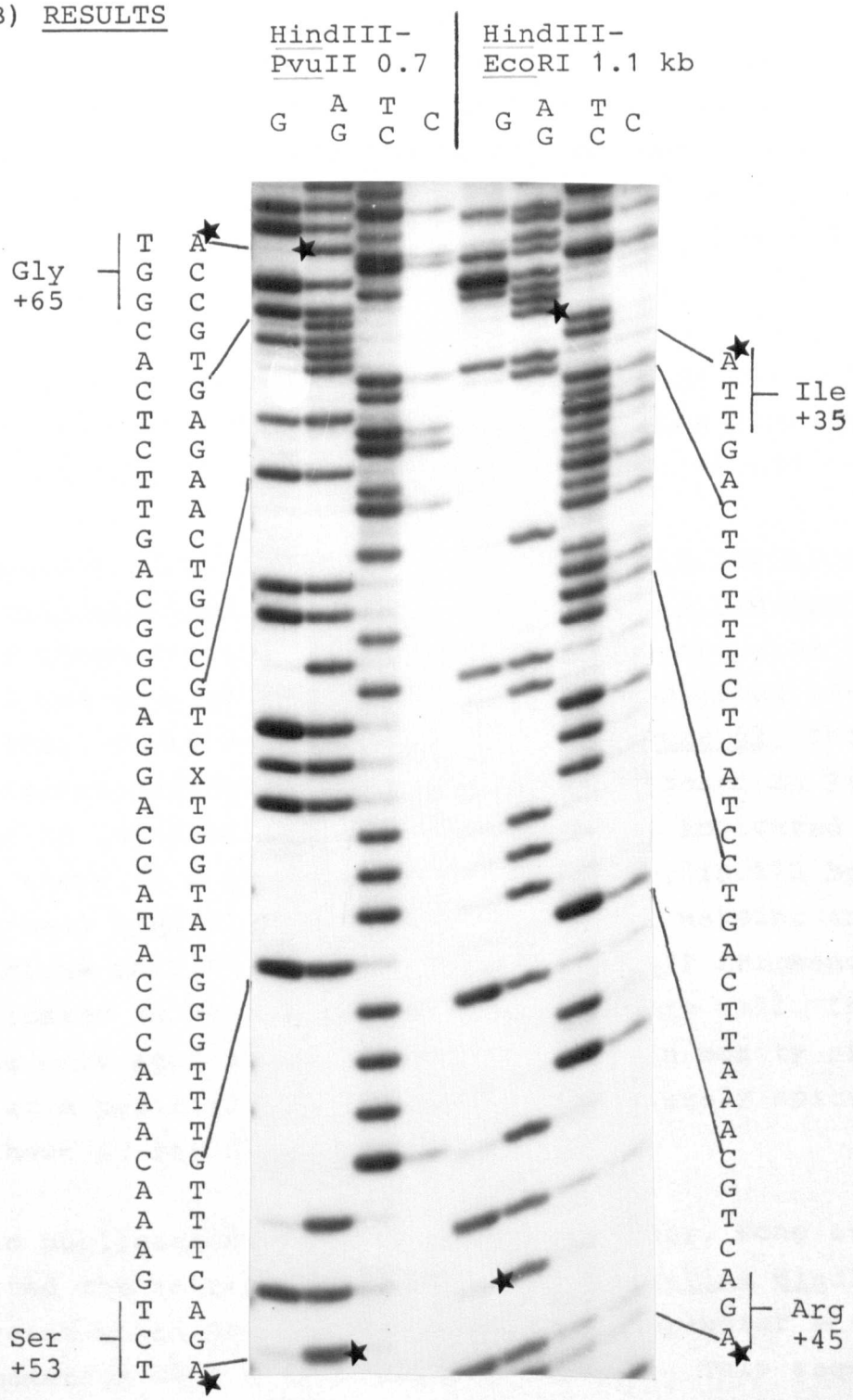
The plasmid pN47.11 was cut with HindIII and the ends labelled by filling out with Klenow enzyme as described in II.7.2. After EcoRI and PvuII digestions respectively, the 1100 bp HindIII-EcoRI, and 700 and 400 bp HindIII-PvuII fragments were isolated

FIGURE 32
PARTIAL SEQUENCE OF SUBTILISIN BPN´GENE

(A) STRATEGY



(B) RESULTS



on 1.5% agarose gels.

Maxam and Gilbert sequencing reactions were carried out as described in II.14.1, and the products analysed on denaturing 8% and 20% gels (II.12.2).

In Figure 32a the number of bases read with certainty from the gels is indicated. Figure 32b shows part of the autoradiograph of a 20% gel covering the region of the 1100 bp fragment coding for Ile+35 to Arg+45. Read from top to bottom, the sequence on the non-coding strand is obtained. Also shown in the figure is the region of the 700 bp fragment corresponding to the coding strand for Ser+53 to Gly+65 (from bottom to top).

These assignations were made by comparison with the published amino acid sequence (Markland and Smith, 1971). The proximal HindIII site was thus shown to lie between codon +48 and +49.

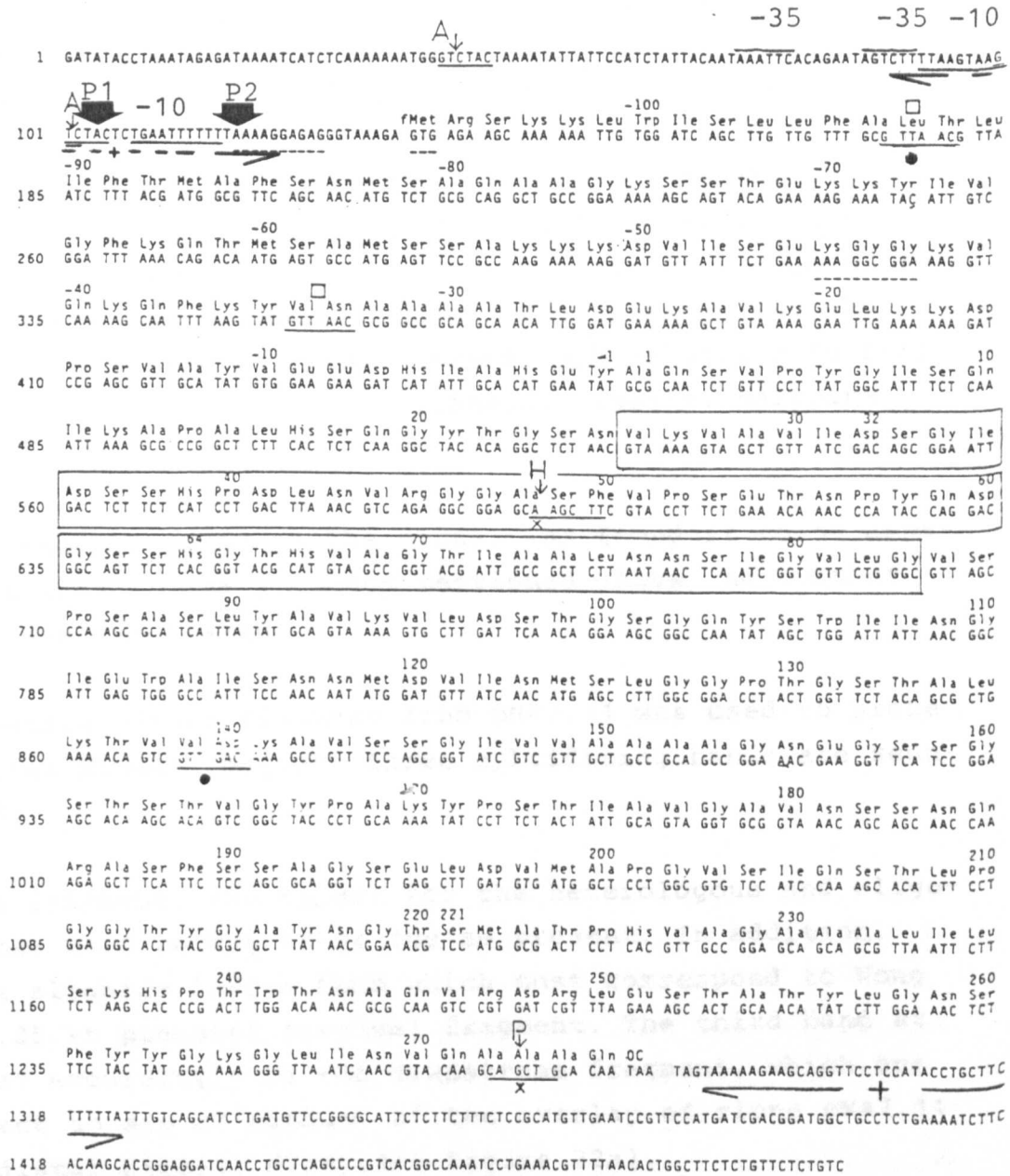
IV.4.2 Comparison of the Clone with Others from B.subtilis; Identification of Up- & Downstream HindIII Fragments

Shortly after these sequences were obtained, the sequence of the entire cloned gene of B. subtilis 168 was published (Stahl & Ferrari, 1984). Their sequence is shown in Figure 33. Those bases I had determined (from codon +26 to +83) (boxed in Fig. 33) proved to be identical to their sequence. As indicated in the figure , there is a PvuII site in codon +273, ie.673 bp from the proximal HindIII site. From chromosomal mapping and analysis of clone pN47.11 the longer HindIII-PvuII fragment had been estimated to be 700 bp, which agrees very well. In addition, the only AccI site(s) in the subtilisin moeity of pN47.11 was at a position consistent with the closely spaced AccI sites shown in Fig.33.

Just prior to publication of the above-mentioned paper, Wong et al (1984) reported the sequence of a cryptic B. subtilis HindIII 1.25 kb fragment which showed stationary phase promoter activity in combination with a promoterless CAT gene. This sequence

FIGURE 33: SEQUENCE OF THE SUBTILISIN GENE OF *B. subtilis* 168

From Stahl and Ferrari (1984)



Sigma-37 -10, -35 regions under/overlined (Wong et al 1984).

BOXED SEQUENCE determined from pN47.11

□, •, × denote the endpoints of deletions mentioned in the text.

□ aprA3	Kawamura & Doi (1984)	A	Acc I
• apr-684	Stahl & Ferrari (1984)	H	Hind III
× aprMFJ157	This thesis	P	Pvu II

In vivo & in vitro transcriptional start sites (S1 mapping, Wong & Doi, 1984).

← + → palindrome

subsequently proved to be the promoter-proximal half of the subtilisin gene, and was identical to the corresponding portion of Stahl and Ferrari's clone. By deduction, most of the Wong sequences should also be present on clone pN47.11 (see below).

Having pinpointed the proximal HindIII site by sequencing, it is clear that sequence data obtained from the distal HindIII-PvuII 400 bp fragment must lie well downstream of the gene (see Fig. 32a). There was no open reading frame covering the whole of the 127 bp obtained on either strand, and it showed no striking features such as AT- or GC-richness, repeats, hairpins etc.

Having oriented the gene on the clone, and found it to be apparently identical to the 2 clones mentioned above, no further sequencing was undertaken.

When the entire cloned fragment from pN47.11 was used to probe a chromosomal HindIII digest, three hybridizing bands are seen (Figure 38).

The 1.1 kb fragment, the target for the heterologous and oligomeric probes used before is of course present. In addition, there is a slightly larger band which must correspond to Wong et al's 1.25 kb promoter-proximal fragment. The third band at 2.3 kb must accordingly be the downstream fragment, which one would expect to see on account of the overlap of clone pN47.11 with the distal HindIII site (see Figure 32a).

There was the possibility that this represented incomplete HindIII digestion (i.e.1.1 kb + 1.2 kb = 2.3 kb). In later experiments this was shown not to be the case. The upstream HindIII site was later cloned together with approx.1.4 kb of DNA flanking the pN47.11 clone in chromosomal integration/excision experiments in B. subtilis and subsequent recloning in E. coli, and shown to lie approx. 150-200 bp from the EcoRI site, as expected (pMJ22), confirming that the c.1.2 kb HindIII

band lies upstream.

As further sequencing seemed unnecessary in view of the coincident publication of the sequences cited above and the apparent identity of these with the pN47.11-carried gene, expression experiments, and the generation of a subtilisin deletion mutant of B. subtilis were undertaken directly.

IV.5 Expression of the Cloned Gene in B. subtilis

IV.5.1 Recloning on a Gram-Positive Vector

The 2.6kb ClaI-EcoRI fragment carrying the gene and about 600bp upstream was recloned on p1050, a pBD64 derivative of B. Diderichson. It was inserted between adjacent EcoRI and ClaI sites in the vector polylinker. To facilitate the cloning, the vector was supplied in two individually isolated fragments (EcoRI-BglI; BglI-ClaI; see Fig.34a). The resulting plasmid was called pMJ6.

IV.5.2 B.subtilis pMJ6 Transformants Overproduce Subtilisin

B. subtilis Cm^R transformants overproduce subtilisin.

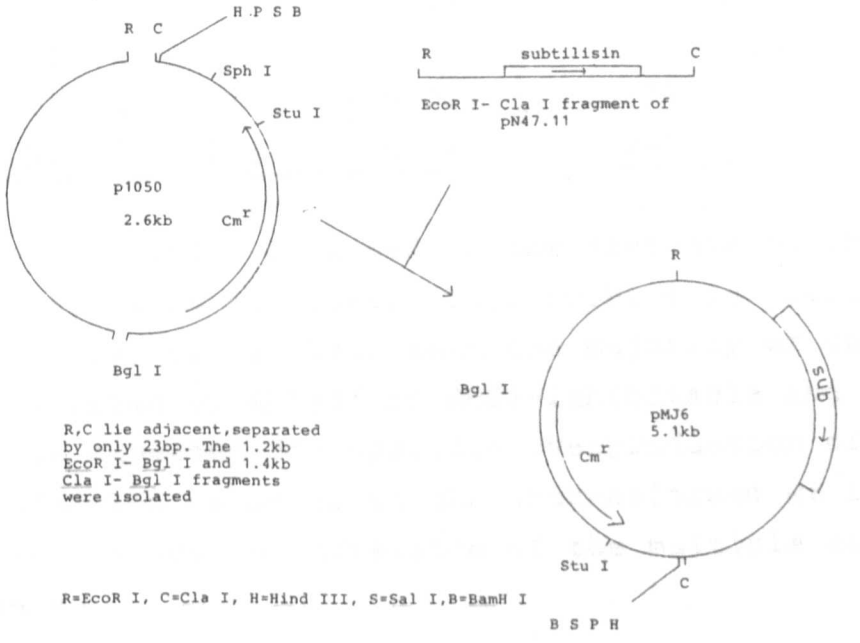
This was evident from the large halos about colonies on skim-milk plates (II.9.1.1; Fig.19: MFJ135 vs. MFJ93).

Extracellular protease was estimated with Azocoll as described in II.9.1.2, and production monitored during growth (Fig.34b). Medium and inoculation were as described in III.7.2.2. Wild-type MFJ93 and transformant MFJ135 were compared. Note the different scales on the ordinate axis. In the case of the overproducer MFJ135, a 3-fold smaller sample volume was used (viz. 0.15 ml), and the results are shown uncorrected for this difference.

Samples of supernatant were also tested for protease activity at 12h, and 20h of incubation, and results are shown below:

FIGURE 34

(A) TRANSFER OF CLONED B. subtilis SUBTILISIN GENE TO GRAM-POSITIVE VECTOR FOR EXPRESSION STUDIES



(B) OVERPRODUCTION OF SUBTILISIN FROM CLONED GENE

MFJ93 was transformed with a plasmid carrying the cloned subtilisin gene. The Cm^R pMJ6 transformant, MFJ135, overproduces serine protease. Culture medium was sampled,and exoproteolytic activity measured at different times during growth, as described in the text.

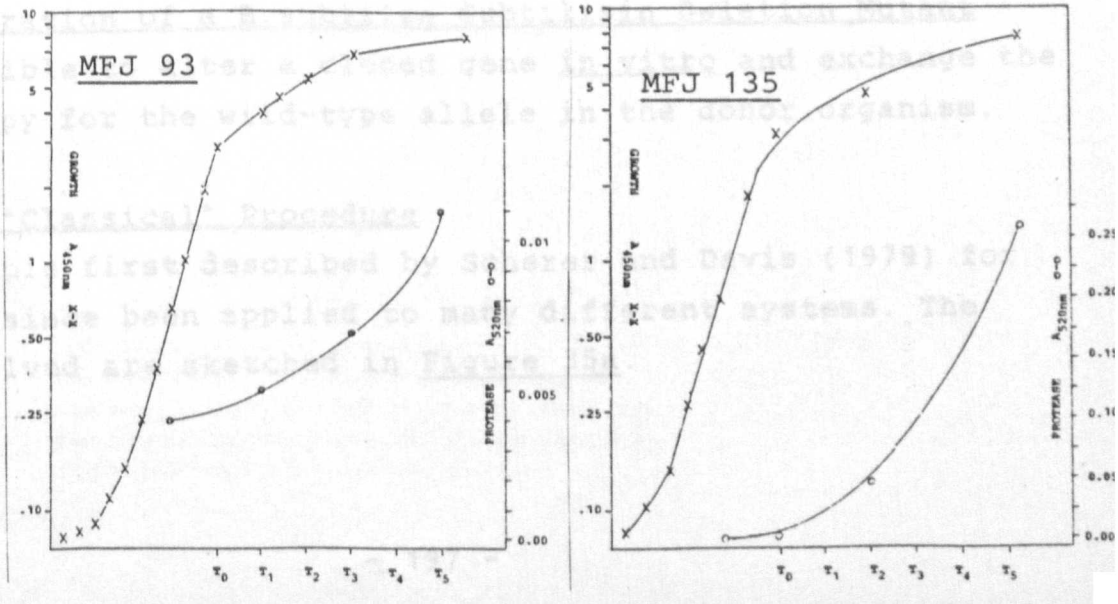


Table 9: Overproduction of Serine Protease by MFJ135

	MFJ93		MFJ135	
	A_{520}	Resid.Act.	A_{520}	Resid.Act.
12h (total)	0.040		0.353	
20h (total)	0.083		0.475	
+ PMSF	0.016	19%	0.002	<0.1%
+ EDTA	0.044	53%	0.446	94%

Note: The sum of activities obtained in the presence of the inhibitors does not match the total. This problem is discussed later. From the table, it is clear that the majority of the extra protease secreted by MFJ135 is PMSF-inhibitable and EDTA-resistant, just as expected for specific overproduction of subtilisin. Overproduction relative to the untransformed wt is about 30-fold, and is due to expression of the multiple copies of the cloned gene.

Even though exoprotease production by MFJ135 is so high, activity is detected only after stationary phase has been reached ie. after T_o. This restriction of subtilisin activity is due to transcriptional control by *inter alia* minor sigma factors, and probably catabolite repression, (see discussion in Chapter I), and parallels the results of others (Ferrari et al,1986).

The promoter region of the *B. subtilis* 168 subtilisin gene has been characterized as far as transcriptional initiation sites are concerned (Wong & Doi, 1984; I.3.5.4; Fig. 33).

IV.6 Generation of a B.subtilis Subtilisin Deletion Mutant

It is possible to alter a cloned gene *in vitro* and exchange the altered copy for the wild-type allele in the donor organism.

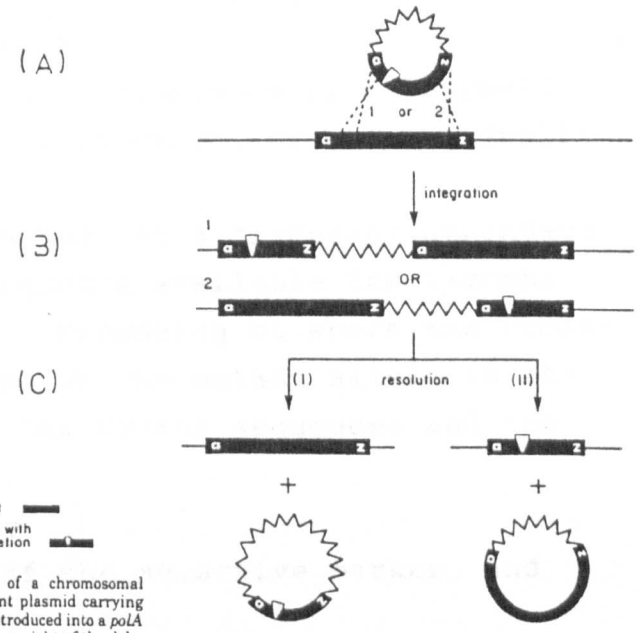
IV.6.1.1 "Classical" Procedure

The principle first described by Scherer and Davis (1979) for yeast has since been applied to many different systems. The steps involved are sketched in Figure 35a.

FIGURE 35 : GENERATION OF CHROMOSOMAL DELETIONS USING IN VITRO ALTERED CLONED SEQUENCES

(A) "CLASSICAL" ROUTE

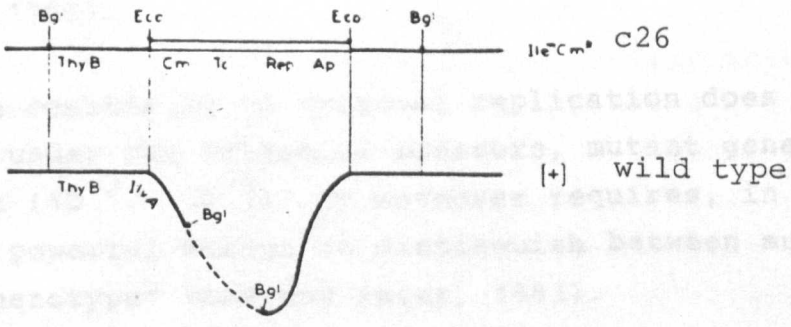
From Gutterson and Koshland (1983)



Expected pathway for replacement of a chromosomal segment with an altered segment. A recombinant plasmid carrying bacterial sequences with an internal deletion is introduced into a *polA* strain. Recombination can occur to either the left or right of the deletion, giving rise to one of two possible chromosomal DNA structures with integrated plasmid. The plasmid integrates are obtained by selection for tetracycline resistance. After removal of tetracycline from the medium, tetracycline-sensitive segregates are obtained when plasmid DNA resolves from the chromosome via homologous recombination. The segregates retain either the original sequence or the altered sequence in the chromosome. Replacement of the original sequence occurs when the plasmid resolves via recombination on the opposite side of the deletion from the integration event.

(B) BY SUBSTITUTION WITH LINEAR FRAGMENT

From Niaudet et al (1982)



Structure of the *thyB* region of the clone c26 ($Ile^- Cm^R$) and the wild-type (+) chromosome. Chromosomal DNA is represented by thick lines, pHV32 DNA by double lines. Sequences possibly existing in the wild-type, and absent in the c26 chromosome are shown as an interrupted line.

(A) First the altered gene, which must have homology to the host genome on either side of the alteration is introduced into the host by transformation with a plasmid which is preferably unable to replicate in the recipient, but which carries an expressible selective marker.
(B) The plasmid integrates into the chromosome by a Campbell-like mechanism under the selective pressure, creating a duplication of host sequences.
(C) The selective presure is removed. At a frequency dependent upon the length of duplicated sequence available for recombination, the plasmid will be lost. Depending on where the cross-over occurs, either the wild-type or the mutant allele is retained in the chromosome, while the vector sequences and the other copy of the gene are lost.

Colonies are screened for loss of the selective marker, and then the phenotype of interest.

IV.6.1.2 Application to Systems Other than Yeast

Gutterson and Koshland (1983) extended this idea to E. coli and S. typhimurium using ColE1 replicons and polA strains to prevent extra-chromosomal plasmid replication.

The polA(TS) system can also be used to test whether a cloned gene is essential: if expression of the selective marker can only be achieved at the permissive temperature, then the cloned gene one is attempting to alter/delete is most probably essential (Date, 1983).

Where such a constraint on episomal replication does not force integration under the selective pressure, mutant generation is a rare event (10^{-3} - 10^{-4}). It moreover requires, in the first instance, a powerful screen to distinguish between mutant and wild-type phenotypes (Lee and Saier, 1983).

IV.6.1.3 Interaction With The B. subtilis Chromosome

1) Non-Replicating Plasmids

Plasmids which carry a region of homology to the chromosome, can also integrate into the Bacillus subtilis chromosome by a Campbell-type mechanism in a rec-proficient host.

This was first shown by Duncan et al (1978). B. subtilis (Thy^-) was transformed with the nonhomologous bacteriophage β22 thymidylate synthetase gene on pMB9, which also carried a short 0.51 Md region of homology with the chromosome. The plasmid cannot replicate in B. subtilis, and was found integrated into the chromosome of Thy^+ transformants.

Indeed, the ability to select for integration at specific sites of nonreplicable plasmids has since been exploited to map cloned genes (Haldenwang et al, 1980; Ferrari F. et al, 1983).

2) Replicable Plasmids

In rec- proficient hosts, plasmids which are able to replicate and carry a region of homology to the chromosome, interact with the chromosomal locus (Iglesias et al, 1981). A symmetrical gene conversion can take place, and be detected where the plasmid carries an allelic variant of the host gene (Iglesias and Trautner, 1983).

IV.6.1.4 Subtilisin Deletion Mutants

As far as generating a subtilisin deletion mutant is concerned, Stahl and Ferrari (1984) utilized the Scherer and Davis logic pictured in Fig.35a, to construct a strain lacking a 684 bp HincII fragment in the gene. Kawamura and Doi (1984) on the other hand, introduced their altered subtilisin gene (with a 178 bp HpaI deletion) on pUB110 and screened Km^R transformants for gene conversion i.e. reduced exo-protease production. (See Fig.33, where the endpoints of the various deletions in the gene are indicated). Both groups started out with already muta-

genized strains having greatly reduced exoprotease activity compared with wild-type. This was a prerequisite in the latter case in order to be able to detect the rare mutants.

As a suitable transformable mutant was not available, and as the classical Scherer procedure had already been used for the purpose, it was decided to attempt the generation of such a subtilisin mutant by a slightly different approach, suggested by recent results described by Ehrlich (Niaudet et al, 1982) (see below).

IV.6.2 Gene Conversion by Linear Fragments

A) B. subtilis: In studies of intergenotes transformed with hybrids of B. globigii and B. subtilis DNA, Harris-Warwick and Lederberg (1978) showed that linear heterologous DNA fragments flanked by homologous DNA regions could be integrated into the B. subtilis genome en bloc with high efficiency.

In transformation experiments using a ColE1 replicon carrying the chloramphenicol resistance gene and B. subtilis DNA, Niaudet et al (1982) found that integration into the B. subtilis chromosome sometimes led to a mutant phenotype.With plasmid pHV438, where a B. subtilis DNA insert specifying thyB was used, one Cm^R transformant C26, was found to require isoleucine for growth ("insertional mutagenesis"). When chromosomal DNA of C26 was examined the strain was shown to have suffered a deletion of approximately 6 kb of ilvA sequences, which map very close to thyB. The data could be explained by assuming that instead of a Campbell-like integration of the incoming plasmid, a replacement of chromosomal DNA had taken place by a double-crossover event involving sequences homologous to the chromosome on either side of the vector (see Fig.35b). This hypothesis had been tested by repeating the transformation with the original pHV438 plasmid both in intact form, as well as after linearization at a site within the B. subtilis-derived sequences. When intact plasmid was used, approximately half of the Cm^R transformants were Ile^+ (Campbell insertion). With

linearized pHV438, all the Cm^R transformants were Ile^- (double-crossover event).

Thus one could expect to be able to generate defined deletions using linearized non-replicable plasmids very easily. In contrast to the Scherer scheme, vector sequences remain stably in the chromosome. Even with non-linearized plasmid, there was 50% deletion generation, presumably due to the linearization process that occurs randomly during plasmid-uptake into competent B. subtilis cells (de Vos et al, 1981). In this connection it is worth mentioning that the model proposed for the "activation" of plasmid monomers by homologous DNA inserts in the transformation of B. subtilis cells (monomers with no homology to DNA already within the recipient cell cannot transform), put forward by Canosi et al (1981), demands that a break is present in the part of the plasmid carrying the homologous sequence in order to allow single-stranded invasion of the chromosome at the corresponding locus. This synapsing leads to cicularization and stabilization of the single-stranded molecule and finally to DNA synthesis and replication of the plasmid. The "plasmid marker-rescue" transformation system of Dubnau (Contente and Dubnau, 1979; Gryczan et al, 1980) supports this model, and it can be concluded that linear plasmid DNA, partly homologous either to the chromosome or to a resident plasmid, takes part in a transformation process entirely analogous to classical chromosomal transformation.

Ehrlich has recently enlarged on the C26 data, and discussed the deletion of essential genes by the linear DNA systems (Niaudet et al, 1985).

B) Other Systems Besides B. subtilis: Gene replacements by transformation with linear fragments carrying either an in vitro altered gene or else a heterologous gene flanked by sequences homologous to, and in the same orientation as, two closely spaced regions on the host chromosome, have recently been des-

cribed for yeast (Rudolf et al, 1985) and E. coli (Jasin and Schimmel, 1984).

In the yeast example, the linearized fragments were introduced by cotransformation with a self-replicating plasmid carrying the selective marker. In the E. coli case, the selective marker was flanked by the homologous sequences. The method was extended to the deletion of essential genes by providing the essential function on a plasmid with a ts replicon.

IV.6.3 Construction of Plasmids for Subtilisin Deletion Mutant Generation

In summary of the above discussion:
The apparent ease with which an in vitro altered gene could be substituted for the wild-type allele in one step by introducing it into B. subtilis as a linear fragment, with a strong selection, led to the construction of plasmid derivatives of the subtilisin clone, pN47.11, where a portion of the coding sequence had been removed and replaced with a selectable marker; specifically, the chloramphenicol acetyl transferase (CAT) gene of pC194 (pMJ13, IV.6.4,-5).

The alteration would thus be labelled with a selectable marker (CAT), and should thereafter be easily transferable to closely related strains by transformation.

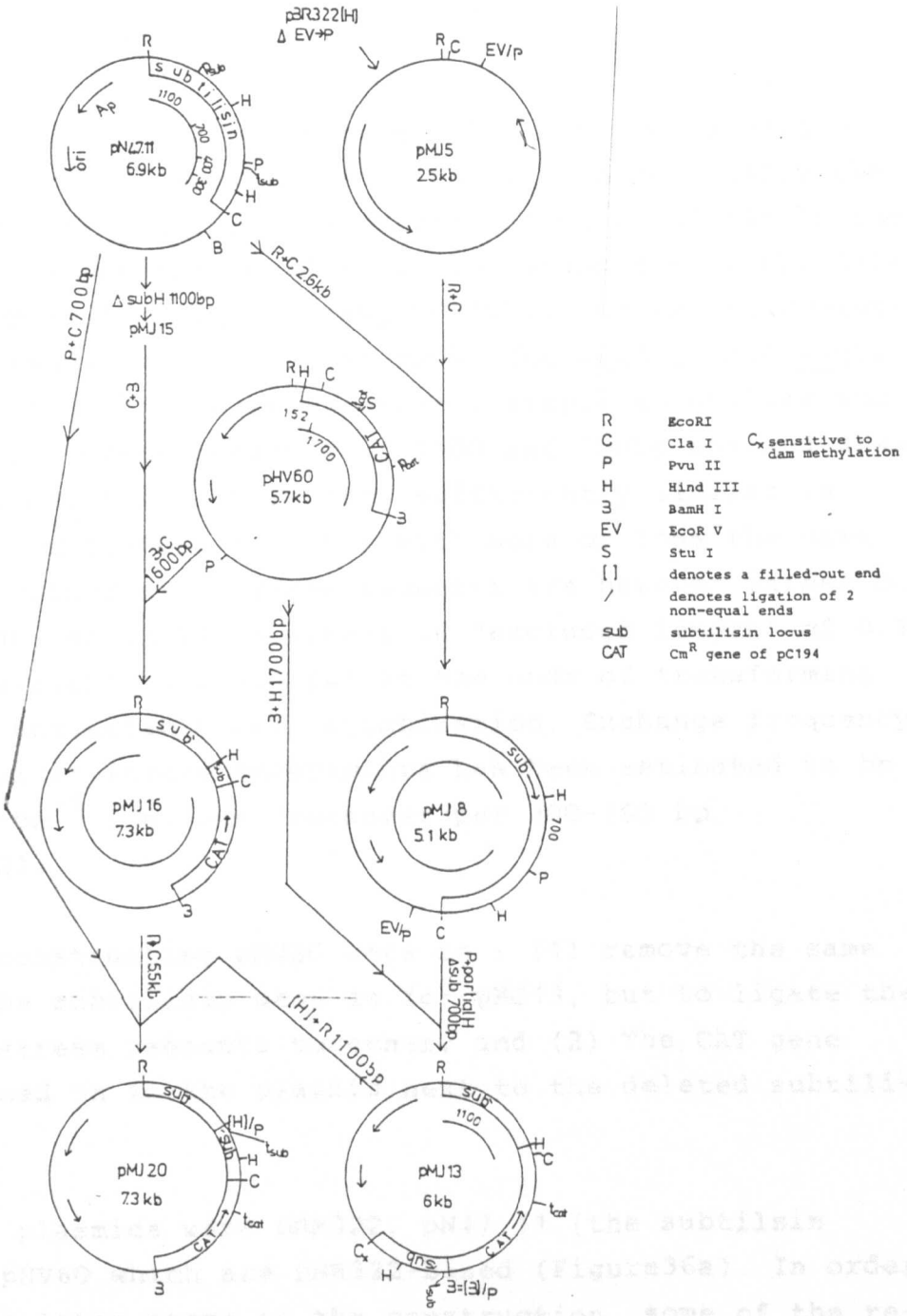
Precisely this method of interrupting/deleting part of a gene has since been adopted with the regulatory sacQ gene (Yang et al, 1986).

In addition to pMJ13, a derivative of pN47.11, which could be used to generate a subtilisin deletion mutant by the Scherer and Davis procedure of integration followed by excision was also prepared (pMJ20, IV.6.6).

The main steps involved in the construction of pMJ13 and pMJ20 are shown in Figure 36, and were carried out in MC1061 or the

FIGURE 36: CONSTRUCTION OF INTEGRATIVE PLASMIDS

(A) FOR GENERATION OF SUBTILISIN DELETION MUTANT



(B) FOR CLONING OF SEQUENCES FURTHER UPSTREAM VIA INTEGRATION-EXCISION

pMJ14 was a by-product of the scheme shown in (A).

dam⁻ derivative, MFJ92.

The aims in constructing pMJ13 were: (1) to remove a portion of the subtilisin coding sequence from pN47.11, specifically the c. 700 bp from the HindIII site corresponding to AA +48 to the PvuII site corresponding to AA +272. The endpoints of the deletion are shown in Figure 37 (Δapr MFJ157), and (2) to replace the deleted fragment with the CAT gene. The HindIII and PvuII sites were used as they were relatively simple to utilize and would leave homologous segments of 1100 and 700bp up- and downstream of the deletion. These are sufficiently similar in length for recombination to occur with more or less the same frequency on either side. These segments are however rather on the short side, as there is always an "excluded length" of 0.15 Mdalton (equivalent to c.230 bp) at the ends of transforming DNA, that is unavailable for recombination. Exchange frequency in plasmid marker-rescue experiments has been estimated to be 0.2-0.5 $Mdalton^{-1}$ i.e. one crossover per 300-760 bp (Dubnau, 1982).

The aims in constructing pMJ20 were to : (1) remove the same segment of the subtilisin gene as for pMJ13, but to ligate the up- and downstream remnants together, and (2) The CAT gene would be cloned on to the plasmid next to the deleted subtilisin gene.

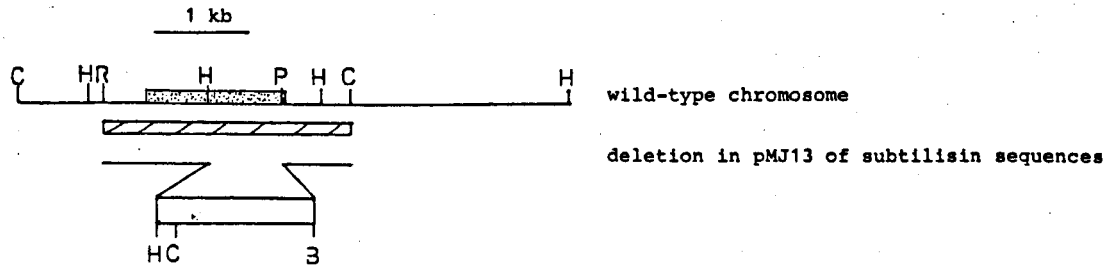
The starting plasmids were pBR322, pN47.11 (the subtilsin clone), and pHV60 which are pBR322-based (Figure36a). In order to facilitate later steps in the construction, some of the restriction sites present were eliminated viz. the HindIII site at base 29 in pBR322 was inactivated by "filling out " and religation.

Excision of DNA from the EcoRV site at 185 bp, to the PvuII site at 2066 bp eliminated the latter site as well as the BamHI site at 375, and eliminated the possible duplication of vector sequences when the CAT-carrying BamHI-HindIII fragment of pHV60

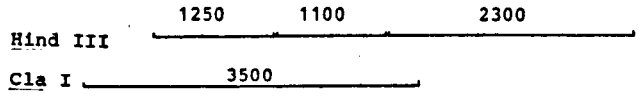
FIGURE 37: ANTICIPATED RESULTS OF SOUTHERN ANALYSIS OF MFJ93 :: pMJ13 INTEGRANTS

Chromosomal DNA is cleaved with HindIII and ClaI. Digests are probed with the subtilisin EcoRI-ClaI 2.6 kb fragment. The sizes of the expected probe-hybridizing digestion fragments are indicated below for the various integrant structures.

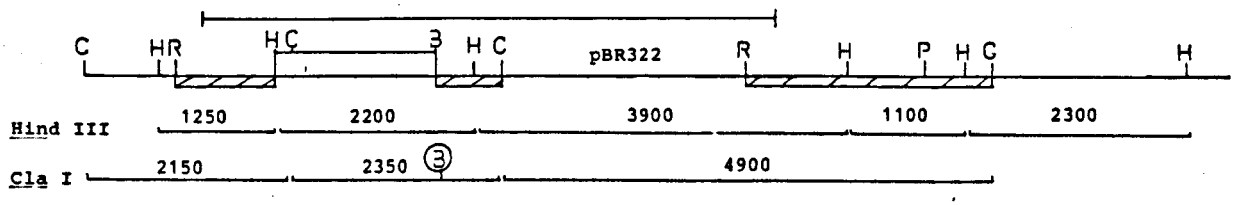
SUBTILISIN LOCUS: WT CHROMOSOME, pMJ13 DELETION



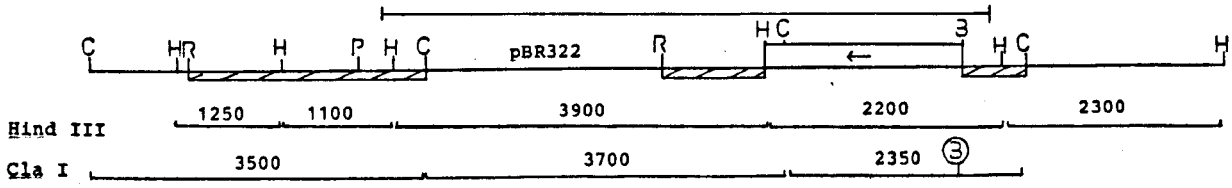
CASE 1: WILD-TYPE MFJ93 CHROMOSOME



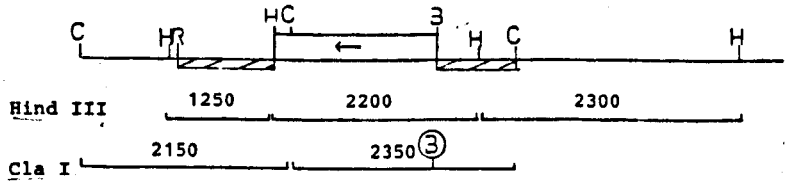
CASE 2: CAMPBELL INTEGRATION via HOMOLOGY 5´TO DELETION



CASE 3: CAMPBELL INTEGRATION via HOMOLOGY 3´TO DELETION



CASE 4: DOUBLE CROSSOVER, GENE REPLACEMENT



intergrated copy of pMJ13

restriction digest fragment (bp)

③ position at which Cla I fragment will be cleaved by BamH I

C Cla I
H Hind III
R EcoR I
B BamH I
P Pvu II

subtilisin gene

probe,and sequences homologous to probe= EcoR I-Cla I insert in pN47.11

CAT gene insert in pMJ13

pBR322= 322-derived sequences of vector

was added to the subtilisin plasmids.

Wherever possible, defined fragments were isolated on agarose gels. Every stage in the construction was rigorously checked by restriction analysis of intermediates. Note that ligation of a filled out BamHI end with a PvuII blunt-ended molecule regenerates the BamHI site.

As a by-product of the last step in the construction of pMJ13, which involved a partial HindIII digestion of PvuII-digested pMJ8, pMJ14 (Fig.36b) was obtained in which the alternative HindIII site had been cut. This plasmid was used in another experiment to clone sequences lying upstream of those in clone pN47.11. (IV.6.7).

IV.6.4.1 Identification & Characterization of apr Deletants

B. subtilis MFJ93 was transformed with pMJ13 that had been linearized with EcoRI and PstI, by congression with pUB110. No Cm^R colonies were obtained among the Km^R transformants.

As a control, intact pMJ13 had been used, and the equivalent amount of plasmid gave 200 Cm^R transformants.

The failure to obtain any transformants with linearized pMJ13 was attributed to the shortness (1100 bp) of the homologous DNA fragment after EcoRI digestion. During the transformation process, this would have been shortened further and evidently made a double crossover event impossible.

Varying Chloramphenicol Resistance; Amplification: Some of the Cm^R transformants from intact plasmid were restreaked several times on L agar containing 6 µg/ml Cm. It was striking that there were 2 types of colonies, viz. small and large. During the first restreaking procedure, some small colonies gave rise to large ones, suggesting amplification of integrated plasmid, such as had been seen by Gutterson and Koshland (1983). The purified colonies were tested at Cm concentrations of 6µg/ml

and 30 µg/ml. The small colonies could only grow at the lower concentration , while the large ones grew vigorously even at 30 µg/ml.

IV.6.4.2 Appearance on Skim-milk Plates

When a few purified small colonies were dotted on agar containing 1% (w/v) skim-milk, one of them (MFJ157) showed a very modestly reduced halo of proteolysis. The halo also displayed qualitative differences from the wild-type halo in that it was fuzzy. Supernatants of overnight L Broth cultures were tested on skim-milk agarose plates as described in II.9.1.1 and confirmed the above observation (Fig.19:MFJ157 vs.MFJ93). Chromosomal DNA was thus prepared from MFJ93, MFJ157, another small colony and a large one for analysis by Southern blot.

IV.6.4.3 Analysis of MFJ157 DNA by Southern Hybridization

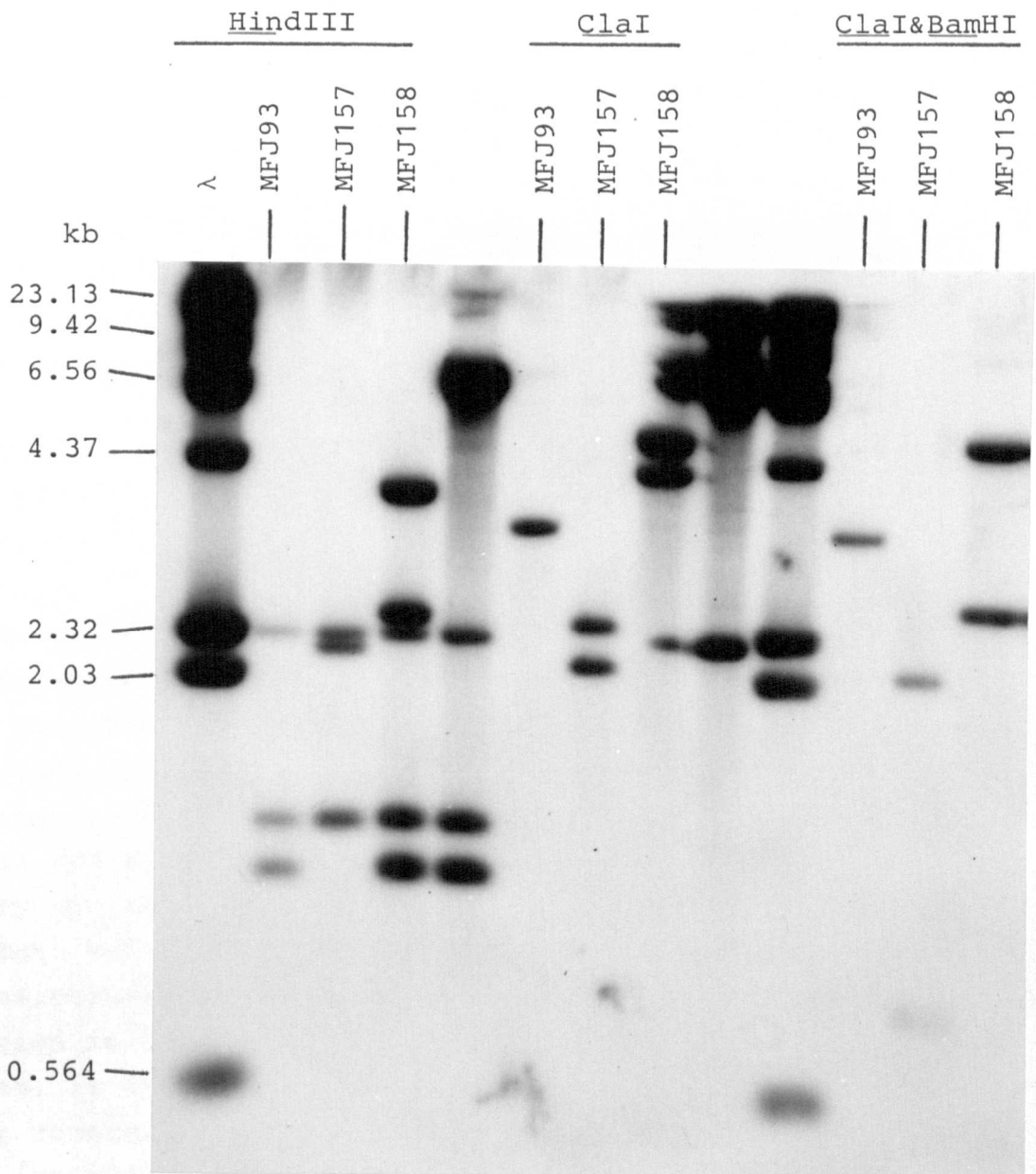
Figure 37 depicts the anticipated results of the Southern blotting experiment.The pre-integration step (cf. Fig. 35(A)a) is not shown. The diagram shows how one can distinguish between wild-type MFJ93 (CASE 1); Campbell integration of 1 copy of the plasmid (CASES 2, 3); and the desired deletant mutant (CASE 4) by probing chromosomal DNA digestions with the subtilisin insert of pN47.11, based on a prior knowledge of the chromosomal restriction map and the plasmid structure.

Figure 38, showing the results of the Southern blot for MFJ157 reveals that this strain gives the pattern of fragments expected to result from a double crossover event resulting in replacement of the wild-type subtilisin gene with the deleted gene interrupted by the CAT gene (carried originally on pMJ13) (CASE 4 in Figure 37). No other bands besides the ones expected are observed. Especially important is the ClaI-BamHI double digestion, which confirms the identity of the ClaI fragment including the CAT gene (subtilisin has no BamHI site), as being the slightly larger of the 2 ClaI bands. This disappears from the double digest, while a new smaller band is seen at about 700 bp, and corresponds to the remnant of the ClaI fragment

FIGURE 38: SOUTHERN ANALYSIS OF MFJ93 pMJ13 & pMJ14 INTEGRANTS

DNA was electrophoresed on 1% agarose. The filter was hybridized to nick-translated pMJ6 / EcoRI - ClaI 2.6 kb isolated subtilisin fragment.

MFJ157 = MFJ93::pMJ13; MFJ158 = MFJ93::pMJ14



(minus the CAT 1.7kb ClaI-BamHI segment) which now cannot be "seen" by the subtilisin probe. The blot was "stripped", and reprobed with pC194. As expected, in the case of the pMJ13 integrant only the upper ClaI fragment is seen and in the double digest a new c.1.7 kb band hybridized. Nothing at all is found in the corresponding MFJ93 tracks (data not shown).

Southern blot analysis of DNA from the other colonies showed in addition to the bands seen in MFJ157, also others, and were not examined further (data not shown).

IV.6.4.4 Measurement of Exoprotease Produced By MFJ157

To quantitate the difference in exoprotease activity between MFJ93 and MFJ157, the putative subtilisin deletion mutant, an Azocoll assay was used (described in II.9.1.2).

Schaeffer's medium was inoculated 1:100 from a fresh L Broth overnight culture. At 20 hours supernatants were taken for assay.

Absolute values obtained in the Azocoll assay varied slightly from one experiment to another. A consistent feature of these assays was that the sum of activities obtained in the presence of the inhibitors seldom approached 100%, and often was 70% of the activity seen in the absence of inhibitor. Why this should be so is not clear. It may be that some of the serine protease activity was compromised in the presence of EDTA. It has recently been established that Ca^{2+} ions form part of the subtilisin structure (see I.2.2.2), but there are no reports of such inhibition in the literature. ISP-I, the intracellular serine protease, is inactivated by chelators (I.2.3.2). Increasing inhibitor concentrations did not significantly increase the degree of inhibition seen. The organic solvent in which PMSF was dissolved had no inhibitory effect of its own. However, the pattern of inhibition observed was quite consistent within +/- 15%. Results from 2 typical experiments using different supernatants are shown below:

Table 10: Reduced Serine Protease Activity of MFJ157

	A_{520}corr			
	EXPT1: 1 ml SN		EXPT2: 0.3 ml SN	
	DN497	MFJ157	DN497	MFJ157
total	0.207	0.052	0.099	0.027
EDTA	0.147	0.016	0.052	0.004
PMSF	0.046	0.028	0.015	0.015
APR::NPR	3.2	0.57	3.5	0.27
% inhibition				
by EDTA	29	69	47	85
by PMSF	77	46	85	44

Less than half the activity in the NOVO reference strain MFJ93 was inhibited by EDTA, and very little was unaffected by PMSF ie, by far the majority of activity is due to serine protease(s). The ratio of APR to NPR produced by B. subtilis 168 is quoted variously as being approximately 1:1 (Prestidge et al, 1971; Higerd et al, 1972 ; Uehara et al, 1974); 4:1 (Millet, 1969); 3:1 (Ito & Spizizen,1972); or 1:4 (Hageman & Carlton, 1973). In some pleiotropic mutants more APR is secreted relative to the wt used to generate the mutant eg. YY88(pap-9), where the ratio ia APR/NPR = 3/1 (Mäntsälä & Zalkin, 1980). In B. subtilis nat-to this ratio is 1:13 (Uehara et al, 1974). Some of this variability is presumably due to different culture conditions and assay methods, and in the last case, to nprR2 regulation.

Mandelstam (Dancer & Mandelstam, 1975a) claimed that the metalloprotease is a substrate for the alkaline protease and progressively disappears from the medium in the course of sporulation, based on experiments where PMSF was added to culture supernants at intervals. These authors point out though that their results are different from those reported by others, where the ratio of the 2 types of protease secreted remain unchanged.

However, the most recent results, from Ferrari and coworkers with defined B. subtilis 168 deletion mutants of both apr and npr (Stahl & Ferrari, 1984; Yang et al, 1984) suggest that the

1:4 APR/NPR ratio is correct. Deletion of the neutral protease and alkaline protease genes respectively led to an 80% and 30% reduction in exoprotease activity, when measured by the Azocoll method. These estimates were consistent with halo sizes on casein plates.

The results obtained above with MFJ93 and 157 were therefore rather surprising, as the identical methodology had been used. As can clearly be seen from the table, the total activity in MFJ157 supernatants was only about 25% of that of the starting strain, and of this residual protease activity, the majority (69-85%) was inhibitable by EDTA ie. there had been a specific reduction in serine protease activity. One would still expect to find some serine protease activity in supernatants of a subtilisin deletant mutant due to cell lysis (which liberates intracellular protease), and due to bacillopeptidase F (see I.2). The results shown above (for freeze-thawed samples) were, however, not very different from those obtained with fresh, sterile filtered supernatants, so any extra contribution to activity by rupture of cells contaminating frozen samples was negligible.

The high contribution of APR to the total MFJ93 exoprotease activity evidently made it easy to detect a difference on skim-milk plates in the mutant MFJ 157. Other workers were not able to do so, since total activity of *B. subtilis* 168 had been reduced by only 20% upon deletion of the *apr* gene (Yang et al, 1984). Since the response on skim-milk plates is logarithmic (Montville, 1983), such a small decrease would barely be visible.

The MFJ157 deletion was therefore transferred to *B. subtilis* 168 (*trpC*) in order to compare results with those obtained by Ferrari & coworkers.

IV.6.5.1 Transfer of Deletion to 168 by Cm^R Selection

Having, so to speak, converted the subtilisin deletion into a selectable marker, it was very easy to transfer the mutation to closely related strains. This was demonstrated by transformation of MFJ93 competent cells with a few µg of chromosomal DNA from MFJ157. DNA was extracted from 6 of the hundreds of Cm^R transformants obtained. HindIII and ClaI digests were probed by Southern blotting with the subtilisin clone as described previously for MFJ157. All 6 isolates gave the same pattern as the mutant MFJ157 (cf. Fig.38). Moreover, when supernatants of overnight cultures of the 6 transformants were compared on skim-milk agarose plates with those of MFJ93 and MFJ157, all 6 showed the hazy reduced halos typical of the mutant (cf. Fig.19). Other, highly mutagenized derivatives of MFJ93 have also been transformed to the APR^-Cm^R phenotype (not shown).

In the same way B. subtilis 168 competent cells were transformed with MFJ157 chromosomal DNA. On skim-milk plates, supernatants of Cm^R transformants were not very different from the wild-type 168 (Fig.19: MFJ231 vs. 168), and Azocoll measurements confirmed a more modest reduction in total protease activity than for MFJ93 when the subtilisin gene is lost.

Under our conditions, a larger proportion of the extracellular protease activity of wt 168 is apparently due to NPR than is the case with MFJ93. While MFJ157 had at most 30% activity relative to the starting strain MFJ93, Cm^R derivatives of B. subtilis 168 have 50%. This is more in keeping with expectations, but still rather different from the results of Yang et al (1984). The mutants all had very low activity in the presence of EDTA, indicating a selective loss of serine protease activity.

The data suggest that strain MFJ93 (=DN497) has in fact an altered protease regulation relative to the archetypal wt, B. subtilis 168. It was isolated as a RUB200 derivative (Yoneda et al, 1979) that was apparently unchanged after transformation

with DNA from a mutagenized Spo^- $protease^{+/-}$ strain (DN243) (B. Diderichson, personal communication).

IV.6.5.2 Characterization of apr Deletants by Immune Diffusion,SDS-Gel Electrophoresis,Western Blotting

To supplement the exoprotease quantification data which indicated a lowered serine protease activity in the putative deletant mutants MFJ157 and 231, other approaches were used to confirm that subtilisin activity was absent.

Immune diffusion of supernatants of wt 168, apr deletant MFJ157, and overproducer MFJ135 against anti-subtilisin NOVO was performed (II.15). No cross-reaction could be detected with any L Broth cultures. An industrial medium, BP-X, was used to grow the strains to high cell density and thus improve protease yield. Figure 4C illustrates the results, which reveal no cross reaction between anti-NOVO and the MFJ157 sample (wells 14, 20 to 24). B. subtilis MFJ93 (well 16) gave a precipitin band of identity with authentic subtilisin NOVO (well 13). This was more intense in the case of the overproducer MFJ135 (well 18). When the 168 sample was diluted 4x as a control so that exoprotease levels in MFJ157 were matched, a precipitin band was still detectable in the 168 sample (not shown).

The results of SDS gel electrophoresis of L Broth samples from "wt" MFJ93 and its putative apr deletant MFJ157 are shown in Fig.21 (lanes 1 & 2). The apparent lack of any alteration in the extracellular protein pattern in the region corresponding to authentic subtilisin NOVO/BPN', and the intense band in the overproducer 135 (Fig.21, lane 3), occasioned doubt about the reliability of the Southern blotting data. In these same samples, a difference in exoprotease activity levels is readily demonstrable on skim -milk plates (Fig.19).

Differences between wt and MFJ157 are indeed seen elsewhere on the protein gel (Fig.21). There is, for example, a more intense

larger band in the sample from the mutant. It was possible that the mapping data might have been misinterpreted, and that a large fusion protein with subtilisin was secreted by MFJ157, accounting for both lowered exoprotease activity, and the larger polypeptide. However, hybrid protein synthesis should have been precluded. Well-characterized fragments had been used in the relevant plasmid constructions, and structures were checked rigorously.

It was only upon gross overloading of 4 day BP-X samples that a band, of mobility similar to the NOVO marker, was demonstrated in MFJ93, and in B. subtilis 168. This band was absent from supernatants of the corresponding Cm^R apr deletants MFJ157 and 231 resp.(Fig.39a). This is more readily apparent when dialysed, lyophilised samples were run (Fig.39b). Whether the band in question is subtilisin cannot be stated with certainty in the absence of confirmatory in situ immunological or activity data, or direct sequencing.The fact that the overproducer MFJ 135, containing multiple copies of the cloned homologous gene, displays an intense band at the same position, distinct from that of the authentic subtilisin NOVO marker, is most persuasive. The slight differences must be ascribed to strain variation. The commercial NOVO preparation originates probably from B. amyloliquefaciens, not B. subtilis. Similarly, subtilisin Carlsberg produced by MFJ29 differs slightly in mobility from that of the commercial "Alcalase".

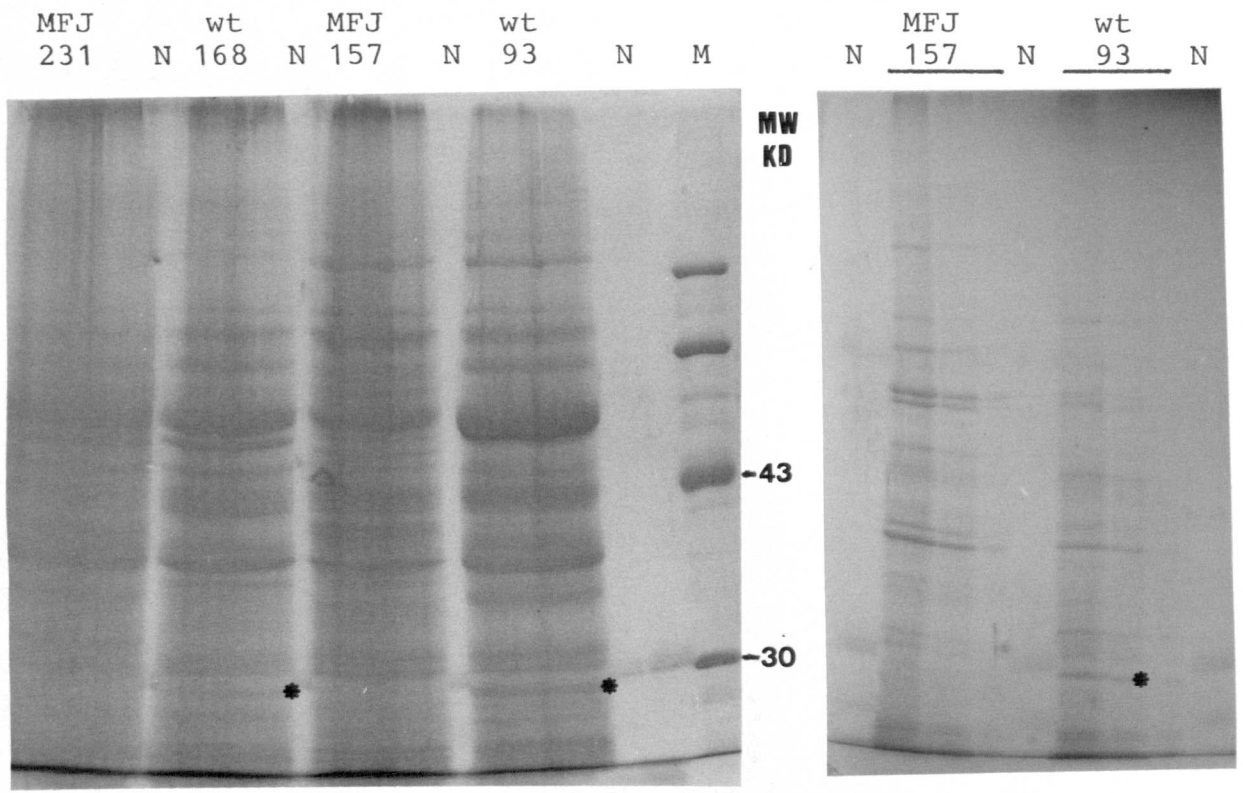
It seems likely that subtilisin levels in L Broth overnight samples are simply to low to be reflected as a discrete protein band on an SDS gel. The band common to wt and apr deletant, and which has a mobility similar to that of authentic subtilisin NOVO, is probably some other unrelated exoprotein (Fig.21).

An attempt was made to correlate the immunological and electrophoretic data on the deletants by the Western blotting technique (II.13). This was a complete failure, despite the adoption of several precautionary measures to reduce proteolysis of

FIGURE 39: apr DELETION MUTANTS FAIL TO SECRETE SUBTILISIN

(A) TCA precipitated samples

(B) Dialysed and lyophilysed



Four day supernatants of BP-X cultures were precipitated with TCA (A) or dialysed and lyophilysed (B), prior to running the exoproteins on SDS gels as described in II.12.3.

Authentic subtilisin NOVO (N) and molecular weight standards (M) were used as size markers.

The subtilisin secreted by B. subtilis wild-types MFJ93 and 168wt is indicated with ✱. The protease is not detectable in the culture medium of the respective Cm^R Δapr transformants MFJ157 and MFJ231.

the antibodies.

Samples were concentrated c.40x by TCA precipitation (II.12.4), and reacted with PMSF. EDTA was also added to inhibit metalloprotease. The effect of low temperature (5°C) incubation, addition of 4% BSA to minimize proteolysis of antibody, and the ommission of Tween, which both removes protein from the filter and activates some proteases, were assessed and found to be of no benefit.

Others have found that most manipulations, but especially electrophoresis, promote hydrolysis of the inactive phenylsulfonate ester of serine protease and thereby restore activity (Roitsch & Hageman, 1983). PMSF and EDTA were consequently added to buffers before incubation with antiserum.

BP-X samples of wt 168, MFJ135 and 213 were tested. The first two produce a NOVO type enzyme, the last, subtilisin Carlsberg. No bands at all were seen after ELISA with less than 0.25 ml of anti-NOVO in Washing Buffer (II.13). Using higher concentrations, non-specific hybridization to all Coomassie Blue-stainable material was achieved, excepting the subtilisin band of wt 168 and MFJ135. Significantly, the subtilisin produced by MFJ 213 was visualized by ELISA, due to cross-reaction with the antiserum (not shown).

Using the more discriminating anti-Carlsberg serum, non-specific cross-reaction was seen at all concentrations of serum high enough to give a signal in ELISA. Again the endogenous subtilisin secreted by B. subtilis was not visualized, but that may be due to the specificity of the immune serum (not shown).

When the filters themselves were stained with Coomassie Blue after the ELISA, subtilisin was found to be in situ, even where the ELISA had failed to show it up.

Successful Western blotting of subtilisin from B. amylolique-

faciens (ie. the authentic NOVO enzyme), synthesized from the cloned gene in a B. subtilis host, has recently been reported (Powers et al, 1986). The authors failed to detect the host enzyme. On the basis of my results described above, it seems likely that this has more to do with aggressive degradation of antibody by B. subtilis subtilisin than, as claimed, to reflect antibody specificity.

The fact that no cross-reaction can be detected on Ouchterlony plates between anti-Carlsberg and large amounts of B. subtilis subtilisin secreted by MFJ135 (Fig.4A, well 2), while cross-reactions between this serum and authentic NOVO enzyme can be demonstrated (not shown), tend to support this view.

This matter could be clarified by repeating the blotting using authentic subtilisin NOVO, instead of the 168 product, as a control.

These experiments were not pursued, as the method apparently could neither be used to detect subtilisin antigen in the deletion mutant, nor could it be applied as a sensitive means of determining low levels of Carlsberg synthesis.

IV.6.6 Transformation of B. subtilis with pMJ20.

MFJ93 was transformed with intact pMJ20 (Fig.36a). As was the case with pMJ13, both small and large transformants able to grow at Cm concentrations of 6 and 30 µg/ml respectively, were seen. Chromosomal DNA was prepared from 3 colonies that could not grow at the higher Cm concentration, and HindIII digests were analysed by Southern blotting, as described previously for pMJ13.

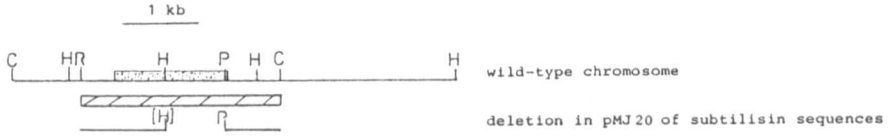
Figure 40a shows the predicted size of bands hybridizing to the probe in the case of integration both up- and downstream of the site of deletion in the subtilisin gene.

Figure 40b shows the pattern actually obtained. Two of the co-

FIGURE 40: STRUCTURE OF MFJ93:: pMJ20 INTEGRANTS

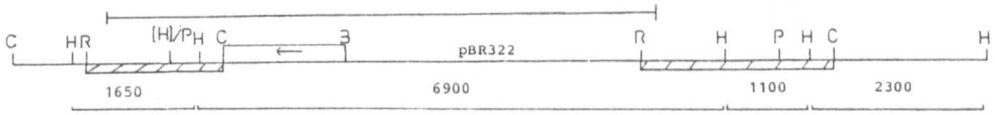
(A) PREDICTED PROBE-HYBRIDIZING HindIII DIGESTION FRAGMENTS

WT CHROMOSOME & pMJ20 SUBTILISIN DELETION

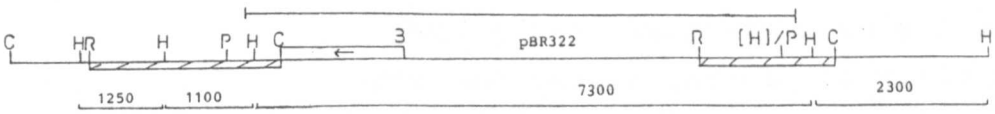


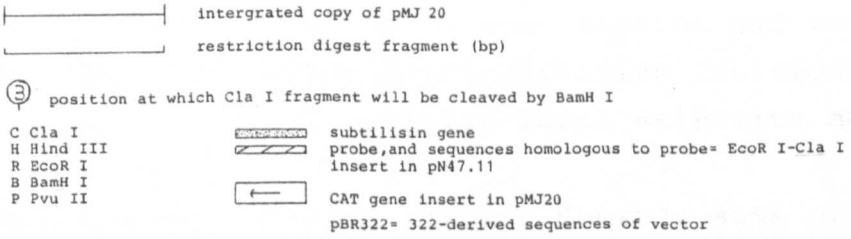
CASE 1: WT MFJ93 CHROMOSOME

CASE 2: CAMPBELL INTEGRATION 5´TO DELETION



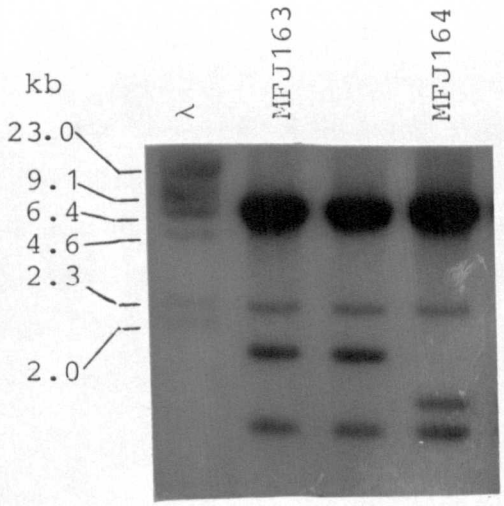
CASE 3: CAMPBELL INTEGRATION 3´TO DELETION





(B) SOUTHERN ANALYSIS OF MFJ93::pMJ20 INTEGRANTS

DNA was digested with HindIII and electrophoresed on 1% agarose. The probe was nick-translated pMJ6 EcoRI-ClaI 2.6kb subtilisin fragment.



lonies analysed, including MFJ163, conformed to CASE 2, and MFJ164 to CASE 3.

Either MFJ163 or 164 could be used to generate a Cm^S Δ*apr* strain by growing it in the absence of Cm and replica-plating on plates with and without Cm 6 μg/ml. One would need to screen about 10^4 colonies to have one Cm^S Δ*apr* (Stahl and Ferrari, 1984).

IV.7 Isolation of Upstream Sequences by Campbell Integration & Excision of Plasmid DNA

In general, once one has integrated a non-replicating plasmid carrying sequences homologous to the chromosome at a specific locus by recombination one can easily obtain an extension of the sequences originally cloned on the plasmid. This is achieved by digesting chromosomal DNA of the integrate with a restriction enzyme that will leave the vector intact, to yield a fragment carrying, in addition, sequences flanking the site of insertion of the plasmid. The DNA is then ligated and transformed into a host where the plasmid's replicative functions are active, with selection for the plasmid-borne selective marker.

Successful use of this technique usually demands some prior knowledge of the chromosomal restriction sites in the area, and has been used by Roeder and Fink (1980) in yeast, and by Ferrari F. et al (1983) and Ehrlich (Niaudet et al, 1982) in *B. subtilis*.

Plasmid pMJ14 (Fig.36b) was a by-product of the synthesis of pMJ13 (Fig.36a). This plasmid has no functional Gram-positive replicon, and can be assumed to integrate by a Campbell-type mechanism, preferentially via the homologous sequences that lie upstream of the subtilisin gene (where there is 1.8 kb available on pMJ14 for recombination), rather than downstream (only 0.3 kb available for recombination).

B. subtilis MFJ93 was transformed with pMJ14, and chromosomal DNA of one Cm^R transformant MFJ158 was examined by Southern blotting. The probe was the EcoRI-ClaI 2.5 kb subtilisin clone from pN47.11 (Fig.31c).

Figure 41a shows the expected chromosomal configuration of the integrant MFJ158 = MFJ93::pMJ14, and the size of the bands one would expect to hybridize to the probe after HindIII and ClaI digestions.

Sketching out the 2 alternative integrants viz. with integration via sequences respectively up- and downstream of the deleted subtilisin segment, reveals that HindIII subtilisin-specific fragments would be of the same size regardless of site of integration (only 1 case is shown in Fig.41a). The ClaI specific fragments however, differ in size in the 2 alternatives. Moreover, the ClaI & BamHI double digestion will also be diagnostic.

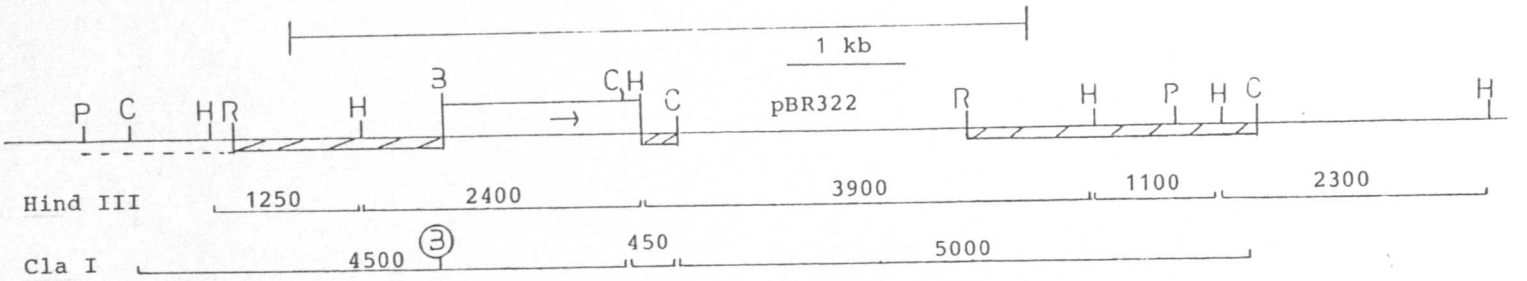
Figure 37 shows that the blotting results were in excellent agreement with the predicted integration drawn out in Figure 41.

pMJ14 integrant MFJ158 was also reprobed with pC194 as described previously for MFJ157, and only the band predicted to carry the CAT gene was found to hybridize, as one would expect.

Combination of the original chromosomal mapping data obtained with shorter probes (Fig.30c) and the structure of the inte-integrant confirmed by the blot show that digestion of chromosomal DNA with PvuII would leave the pBR322-derived replicon intact (recall: unique vector PvuII site inactivated in construction of pMJ14 precursor, pMJ5; Fig. 41a) and would give on one fragment all the material on pMJ14 at the start plus about 1.4 kb of new material lying further upstream of the subtilisin gene.

FIGURE 41: STRUCTURE OF MFJ93::pMJ14 INTEGRATE & EXCISION PRODUCT

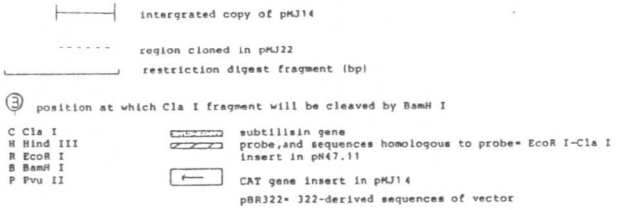
(A) STRUCTURE OF INTEGRATE RESULTING FROM CAMPBELL INTEGRATION VIA HOMOLOGY 5´TO THE CAT GENE AND PROBE-HYBRIDISING FRAGMENTS EXPECTED IN SOUTHERN ANALYSIS



Fragments expected from a single crossover downstream of the CAT gene

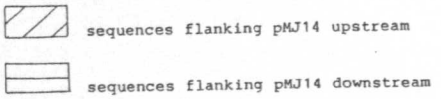
Hind III as above

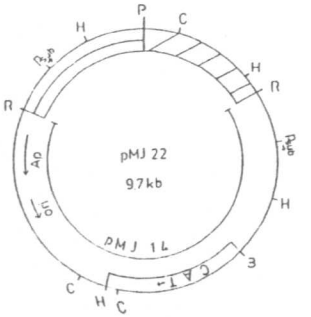
Cla I 5900 ③ 450 3500



(B) STRUCTURE OF EXCISION PLASMID pMJ22, WHICH CARRIES NEW CLONED SEQUENCES FROM THE SUBTILISIN PROMOTER-PROXIMAL REGION

With respect to MFJ93::pMJ14 chromosome:





The MFJ158 DNA was consequently digested with PvuII, ligated and transformed into E. coli. The plasmid found in the Ap^R transformants is pictured in Figure 41b (pMJ22).

Chapter V: CONCLUSIONS

Genes coding for two homologous exoproteases have been cloned from two different *Bacillus* spp.

The comparative primary structure of the Carlsberg (Fig.14) and BPN'-type (Fig.33) subtilisin genes showed a high degree of conservation not only in the region encoding the mature enzyme, but also about putative precursor processing sites (Fig.3 in Appendix 1). Details of the processing remain obscure, though it seems that maturation can be partly autocatalysed (Powers et al, 1986).

Secondary structure predictions revealed that folding over the entire precursor protein of both subtilisin types is potentially identical, despite the primary sequence divergence (Fig.4 in Appendix 1).

Inspection of the sequence upstream from the coding region of both enzymes revealed several features that have been associated with regulation and expression in other systems. These include promoter-like hexanucleotides, an AT-rich region, indirect repeats and a ribosome-binding site (*B. licheniformis* gene, III.6; *B. subtilis* gene, Wong et al, 1984; Stahl & Ferrari, 1984).

Analysis of the *in vivo* transcriptional initiation region of the BPN'-type enzyme revealed two closely-spaced start sites (*B. subtilis* gene, Wong & Doi,1984). When the author applied the same technique, viz. S1-mapping, to the *B.licheniformis*-derived Carlsberg gene, two *in vivo* sites were similarly seen, although in this case the localization was not precise enough to relate the sites to promoter-like sequences present in the region with any certainty (III.8.2). By cDNA sequencing however, only one start site was apparent (Fig.14 bases 297 &8; Appendix 3).

The cDNA data should be regarded as being the more reliable. While some discussion of the artefacts produced by both S1 and cDNA methods is found in the literature (Sharp et al, 1980; Szekely, 1977 resp.), it is now believed by workers utilizing these techniques that the following two factors play the greatest role where mRNA integrity is assumed. S1 nuclease can attack RNA/DNA hybrids at rU-dA regions due to the weak bonding involved (Martin & Tinoco, 1980). Premature termination of cDNA is attributed to contamination of reverse transcriptase preparations by RNAases specific for either single- or double-stranded RNA. In recent years such contamination has been minimized.

In the case of the B. subtilis gene, a run of T residues on the non-coding strand could certainly account for the shorter protected fragment (P2 in Fig.33), by increasing the S1 sensitivity of a longer mRNA-DNA hybrid. Although the Carlsberg gene is also pyrimidine-rich in the putative P1,P2 region, it lacks the extreme T-richness of its BPN' counterpart in the area (cf. Figs. 14 & 33), and thus makes artefactual generation of shorter "protected" fragments less likely. The discrepancy between the S1 and cDNA data remains unresolved.

The transcriptional initiation site indicated by cDNA sequencing is preceded only by a "perfect" -10 region for sigma-43, but no recognizable -35 homology is present at an appropriate distance. This highlights the inadequacies in our understanding of Gram-positive promoters (cf. Table 1, I.3.4,Fig.14).

Transcripts starting at this position, or further upstream, have rich potential for adoption of the various possible 2^{o} structures, as discussed in III.6. Both of the long palindromes identified (Fig.14) would be present on such a transcript, but whether these features of sequence have any regulatory role remains to be established. If indeed they do, their control could be exercised either at the DNA or RNA level.

By analogy to other systems, it is likely that the left-most palindrome functions as a chromosomal recognition site for a regulatory protein (Rosenberg & Court, 1979).

Because of the internal complementarities, subtilisin Carlsberg mRNA could adopt alternative, mutually exclusive secondary structures at the 5' end, involving either the one or the other of the two long palindromes situated just upstream from the start of the coding sequences (see Fig. 14). In the one conformation, the RBS would be sequestered in a stem, and might be inaccessible to ribosomes. There are well-documented reports of such sequestration having a negative influence on translation initiation ,eg. the lamB701, and B708 mutants of Schwartz (Hall et al, 1982). Other references are listed in III.6. Alternative mRNA secondary structures have been proposed to regulate the B. subtilis trp operon (Shimotsu et al, 1986).

In connection with possible post-transcriptional regulation, the considerable delay in subtilisin Carlsberg secretion from B. subtilis when expression was under heterologous promoter control, is intriguing. These heterologous promoters were placed upstream from the in vivo start site (III.7.2.2, III.7.3). Thus all the secondary structure possibilities for the authentic subtilisin Carlsberg mRNA, described above, would also be present for transcripts initiated at these heterologous promoters.

It needs to be clarified whether, in these experiments, transcription indeed took place from the promoters in exponential phase, as only indirect evidence of their activity in growth phase was presented. If, as anticipated, their vegetative phase activity is confirmed, it must be ascertained whether the observed delay is due to translational or post-translational factors (eg. secretion, maturation and release from the cell membrane, passage through the cell wall ; III.7.3) .

Should translational control be implicated, the putative role

of the palindromes must be directly established eg. by site-specific mutagenesis.

As has been mentioned, considerable difficulties were experienced by the author, and others, in cloning the Carlsberg gene in E. coli. Others have had problems in obtaining clones in B. subtilis hosts. It is tempting and reasonable to attribute such problems to strong promoter activity (see below and I.10.3.2, III.9), although strain MFJ29 does not secrete excessive amounts of subtilisin Carlsberg (III.8.1; Fig.19). The levels of overproduction of subtilisin Carlsberg by B. subtilis from the multicopy plasmid pMJ51 (III.7.5), were modest compared with those obtained with the endogenous subtilisin under similar conditions (with pMJ6, Fig.34). It has not yet been established whether the B. subtilis host utilizes the same start site(s) as does B. licheniformis.

In E. coli, the postulated "internal RBS" (III.6.3) may have led to the synthesis of an intracellular serine protease, and rendered cloning difficult. On the other hand, the ISP-I gene has been successfully cloned in E. coli, although there were instability problems in subcloning, and suggestions that these were due to removal of a protease inhibitor function (Koide et al,1986; Beppu, personal communication; see I.2.3.3.3).

The subtilisin gene was readily cloned from B. subtilis (Stahl & Ferrari, 1984; Chapter IV, this thesis) in multiple copies in E. coli. It is known that this gene is transcribed in E. coli (Wong & Doi, 1984), and as remarked on above, considerable overproduction of the gene product is obtained when the cloned gene is present on a multicopy plasmid in B. subtilis.

While expression of subtilisin Carlsberg by MFJ29 and in a B. subtilis host was rather feeble, this does not necessarily imply that the promoter per se is weak. Industrial strains of B. licheniformis, after extensive mutagenesis, secrete enormous amounts of exoprotease. This is probably due to changes at some

or all of the regulatory loci (see I.4 to I.6), and/or their target sites. The complex exoprotease control is exerted at various levels, and increased production could be due to many factors other than direct promoter up-mutations.

It cannot however be ruled out that a promoter down-mutation (i.e. a point mutation around the transcriptional start site) has been cloned. It has been shown that the sequences 5' from the Carlsberg gene were repeatedly deleted from clones, and that those lying more than c. 200 bp from the coding region failed to be cloned in this study (III.4). In contrast, no problems were experienced in the cloning of extensive regions upstream from the gene from B. subtilis. This was evidenced by the cloning of pMJ6, from which the gene was strongly expressed (IV.5), and the subsequent integration-excision experiment with pMJ14, in which additional sequences 5' from the gene were cloned (IV.7).

Possibly the sequences that could not be cloned here influence subtilisin Carlsberg promoter activity in B. licheniformis; sequences which would lead to excessively high levels of transcription in a foreign host, where specific modulating factors may be absent. There is evidence of negative regulation of the subtilisin gene of B. subtilis (see I.6.5.3).The instability of clones of the B. licheniformis and B. subtilis α-amylase genes in E. coli and B. subtilis has been reported (resp. for the different genes, Sibakov & Palva, 1984, inter alia; Yang et al, 1983). In the cases cited, removal of upstream sequences enhanced clonal stability, while (some) promoter activity remained. In another report, (Joyet et al, 1984), the B.licheniformis gene could not be cloned in E. coli on a certain restriction fragment which included a large segment upstream from the coding sequence, although a smaller fragment was stable in E.coli.

An alternative possibility to promoter strength that would account for cloning difficulties is that the sequences upstream

from the subtilisin Carlsberg gene may encode a product that is toxic to a cell if present in excess, or which titrate out an essential factor, but which are unrelated to subtilisin apart from adjacent chromosomal localization. A search of the cloned sequences (Fig.14) revealed no significant ORF on either strand upstream or overlapping the gene, which was preceded by a likely RBS.

Only isolation, probably by a single copy cloning method (III. 9), and characterization of these sequences, combined with a study of their expression in B. licheniformis, will resolve these questions.

Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*

Myra Jacobs[1], Margareta Eliasson[2], Mathias Uhlén[2,3,4*] and Jan-Ingmar Flock[4]

[1]Novo Research Institute, Novo Alle, DK-2880 Bagsvaerd, Denmark, [2]Department of Biochemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden, [3]European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG and [4]Department of Microbiology, The Biomedical Center, Box 581, S-751 23 Uppsala, Sweden

Received 11 October 1985; Revised 6 November 1985; Accepted 22 November 1985

ABSTRACT

The gene encoding subtilisin Carlsberg from Bacillus licheniformis has been isolated by molecular cloning using a mixture of synthetic oligonucleotides. The entire nucleotide sequence of the coding sequence as well as 5' and 3' flanking sequences have been determined. The deduced amino acid sequence reveals an N-terminal signal peptide consisting of 29 residues, a pro-peptide of 76 residues followed by the mature protein comprising 274 residues. The ATG initiator codon is preceeded by two putitative overlapping ribosomal binding sequences. A palindromic sequence typical for transcription termination is found downstream from the TAA stop codon. Structural comparisons between different known subtilisin genes reveal extensive homology, particularly in the parts coding for the pro-region and the mature protein. Expression studies in Bacillus subtilis show that the cloned fragment produces a functional enzyme when inserted after a B. subtilis promoter.

INTRODUCTION

The subtilisins are alkaline serine proteases produced by various species of Bacillus (1). There exists considerable confusion concerning the origin and relationship of the two most studied enzymes, subtilisin Carlsberg and subtilisin BPN' (1,2). Both proteases have been claimed to be produced by different strains of B. subtilis (3,4), but it has subsequently been shown that BPN' instead is produced by B. amyloliquifaciens (1) and that Carlsberg most likely is produced by a B. licheniformis or a B. pumilis strain (5). The original subtilisin, first described in 1954 (3) and later called subtilisin Carlsberg, has been thoroughly characterized and the complete amino acid sequence of the mature protein is known (6). Due to its thermostability, this enzyme has become industrially important and is in fact produced in greater quantities than any other

protein (7). The other enzyme, subtilisin BPN', has chemical and physical properties similar, but not identical, to subtilisin Carlsberg (1,2). The amino acid sequence analysis revealed that the two enzymes are approximately 70 % homologous (6).

Structural and functional comparisons of subtilisins from several strains and species of Bacillus suggest that all subtilisins belong to two groups, exemplified by Carlsberg and BPN' (1). This conclusion is based on amino acid composition, enzymatic activity, immunological properties (5) as well as complete amino acid sequence analysis (8). According to Keay and Moser (5) Bacillus pumilis and B. licheniformis produce a Carlsberg-like enzyme, while all other species in their study expressed a BPN'-like enzyme.

Recently, the cloning and characterization of the gene encoding subtilisin BPN' were described (9,10). This confirmed that the enzyme is produced by B. amyloliquifaciens. In addition, the subtilisin gene from B. subtilis 168 has been sequenced suggesting that this protein is closely related to BPN' (11,13) . In order to clarify the evolutionary relationships between the subtilisins, but also to enable structural and functional comparisons with genetic approaches, we decided to isolate the gene encoding subtilisin Carlsberg. Here we describe the cloning and sequencing of the entire coding sequence as well as the 5' and 3' flanking sequences. The sequence reveals an overall structure similar to the subtilisin BPN' coding region, including a 29 residue signal peptide and a 76 residue pro-region preceeding the C-terminal protein consisting of 274 residues. The Carlsberg and the BPN' pro-regions show extensive structural homology, in particular when they are compared by secondary structure predictions.

MATERIAL AND METHODS

Bacterial strains and vectors.

E. coli strains HB101 (12) and JM101 (12) and B. subtilis 168 (11) were used as bacterial hosts. Bacillus licheniformis NCIB 6816 was obtained from National Collection of Industrial Bacteria, Aberdeen, Scotland. The vectors used were plasmid pBR322 (13), pUC8 (13), pBD64 (14) and phages M13 mp8, M13 mp9

(13) and lambda 47.1 (13).

DNA preparations.

Chromosomal DNA was purified as described earlier (15). Plasmid DNA was prepared by a modified alkaline extraction method (16). Restriction endonucleases and other enzymes were used according to the suppliers recommendations. The transformation of E. coli and B. subtilis were made as described by Uhlén et al. (17). Southern analysis (18) and colony hybridizations (9) were performed as described earlier. The synthetic oligonucleotides were prepared by the phosphoamidite method (19) and purified by HPLC.

DNA sequencing.

Restriction fragments to be sequenced were cloned into appropriate restriction sites of M13 vectors and sequenced according to the dideoxy method (20). The samples were analyzed on 5 % denaturing polyacrylamide gels using wedge shaped gels as described by Olsson et al. (21). The sequences were analyzed using the UWGCG software (22).

Protein analysis.

Immune diffusion analysis was according to Ouchterlony (23). The presence of extracellular proteases was detected on milk agarose plates (24). SDS-polyacrylamide gels (25) were used to analyze the gene products.

RESULTS AND DISCUSSION

Identification of the subtilisin Carlsberg gene.

A number of Bacillus species were grown to late stationary phase and the culture medium was collected to find a bacteria that produces subtilisin Carlsberg. The presence of the enzyme was tested with a standard immune diffusion assay using specific antibodies raised against commercial subtilisin Carlsberg.

One strain was positive, namely a type strain of B. licheniformis, NCIB 6816 (data not shown). Chromosomal DNA from this strain was purified and subsequently used to characterize and clone the gene.

To identify the subtilisin Carlsberg gene, we used a synthetic hybridization probe that consisted of 16 14-mers. The sequences of the oligonucleotides were derived from the unique

pentapeptide Ala-Met-Lys-Gln-Ala which corresponds to amino acids 133 to 137 of the published protein sequence (6). The 16 oligonucleotides represent all possible combinations coding for this region omitting the wobble base of the C-terminal alanine residue. The oligonucleotides were mixed in approximately equimolar amounts, labeled with ^{32}P and subsequently used to characterize restriction enzyme digested chromosomal DNA from B. licheniformis NCIB 6816 using Southern blot hybridization (18). The analysis revealed several restriction fragments suitable for cloning (data not shown), most notable a 2.4 kb BclI-fragment and a 0.6 kb EcoRI/HaeIII-fragment.

Cloning of the subtilisin Carlsberg gene.

Initial experiments were aimed at cloning the 2.4 kb BclI fragment. Chromosomal DNA was digested with BclI, and the fragments were separated on preparative agarose gels. DNA fragments in the 2.2 to 2.6 kb range were isolated by the DEAE-paper procedure (12) and the fragments were ligated to BamHI digested pBR322 plasmid. The ligation mixture was used to transform competent E. coli HB101 and the ampicillin resistant colonies were screened for the presence of the subtilisin gene by hybridization with the labeled 14-mer oligonucleotide mixture. Although several recombinant plasmids were obtained which hybridized to the probe, no plasmid with the intact 2.4 kb fragment was found. The failure to clone the BclI-fragment may be due to technical difficulties, but this fragment may also be lethal to the E. coli host cell, due to the expression of the heterologous alkaline protease.

A similar approach was therefore used to clone the 0.6 kb EcoRI/HaeIII-fragment. Chromosomal DNA was digested with EcoRI and HaeIII and after separation on agarose gels, fragments between 0.6 and 0.8 kb were purified and ligated to pUC8 digested with EcoRI and SmaI. Several ampicillin resistant clones hybridized to the oligonucleotide mixture, and further analysis of one of the plasmids, pSUB1, showed the presence of a 0.6 kb insert. Partial DNA sequence analysis of the pSUB1 revealed that the insert encoded the C-terminal end of subtilisin , with the orientation of the gene from the EcoRI site towards the HaeIII site (Fig. 1). The EcoRI site

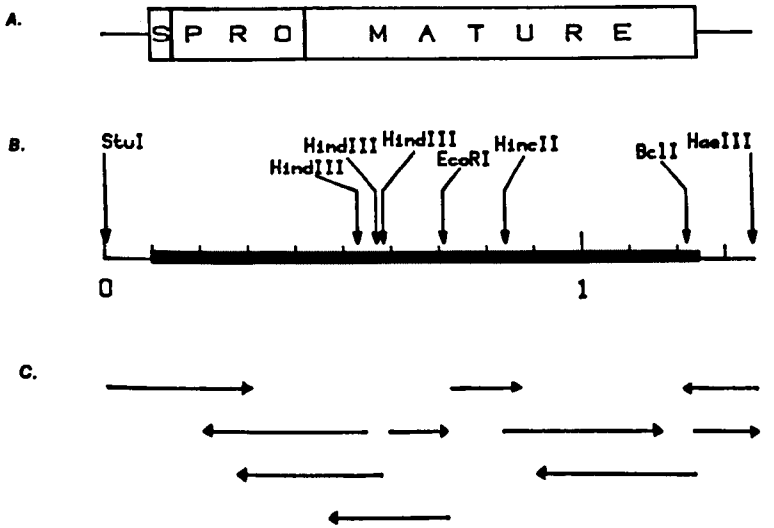


Fig. 1. Restriction map and sequencing strategy of the gene. A, schematic drawing of the gene coding for subtilisin Carlsberg with its different regions. S is a signal peptide, Pro is the highly charged pro-region and Mature is the mature enzyme as determined by protein sequence. B, partial restriction map of the corresponding DNA sequence. C, sequencing strategy of the 1.3 kb fragment.

corresponded to the amino acid residue 96 of the mature protein.

A lambda phage vector was used to clone the 5'-end of the gene. Southern blot analysis of B. licheniformis 6816 DNA has revealed a StuI site approximately 0.7 kb from the EcoRI site (data not shown). Chromosomal DNA was partially digested with Sau3A and separated by a sucrose gradient. Fragments in the 10 to 20 kb range were ligated into lambda 47.1 (13) digested with BamHI. The plaques were screened by hybridization with the 0.6 kb subtilisin gene fragment. 8 plaques were identified as positive and restriction analysis of one of the clones showed the expected 0.7 kb EcoRI/StuI-fragment. This fragment was subcloned into pUC8 giving the plasmid pSUB2. Partial DNA sequence analysis from the StuI site showed that pSUB2 contained the whole 5'-end of the structural gene.

Expression of the intact structural gene in B. subtilis.

In order to express a functional enzyme the gene was introduced into B. subtilis 168. The 1.3 kb StuI/HaeIII

fragment was assembled from the 0.7 kb StuI/EcoRI and the 0.6 kb EcoRI/HaeIII fragments and inserted into a derivative of the Bacillus vector pBD64 (14). After transformation into competent B. subtilis 168, restriction analysis of one plasmid, pSUB3, purified from a chloramphenicol resistant clone, revealed the expected 1.3 kb insert. However, when this clone was grown overnight and the culture medium was tested on milk agar plates, no enlarged clearing zone was observed around the well, suggesting that this plasmid cannot direct synthesis of extracellular protease. A B. subtilis promoter, was therefore inserted upstream of the structural gene of plasmid pSUB3 (data to be published elsewhere). B. subtilis host cells containing this plasmid, pSUB4, gave large clearing zones on milk agar plates when tested in the same way. Furthermore, the protein content of the culture medium of the B. subtilis host cells containing plasmids pSUB3 and pSUB4 were analyzed by SDS-polyacrylamide electrophoresis. This showed that cells containing pSUB4, but not pSUB3, produce an additional protein comigrating with subtilisin Carlsberg from B. licheniformis (data not shown). These results suggest that the 5'-region of the cloned gene does not contain functional signals for transcription in B. subtilis, but when a B. subtilis promoter is provided the enzymatically active protease is produced.

DNA sequence analysis.

Using the strategy outlined in Fig. 1C, the entire gene was sequenced according to the dideoxy method (20). The complete nucleotide sequence of the StuI/HaeIII-fragment is shown in Fig. 2. Starting from a double ATG initiator codon at nucleotide 109, there is an open reading frame of 1137 nucleotides terminating in a TAA stop codon at nucleotide 1246. The total protein consists of 379 amino acids with a molecular weight of 38,910. The ATG codons are preceeded by two overlapping sequences complimentary to the 3'-end of 16S rRNA of B. subtilis. In genes from Gram positive organisms, the distance between the A in the consensus sequence GGAGG of the ribosomal binding sequence and the initiation codon is around 10 nucleotides (15,26). Consequently, although we have not shown that the first ATG codon is used to initiate translation, the spacing to the

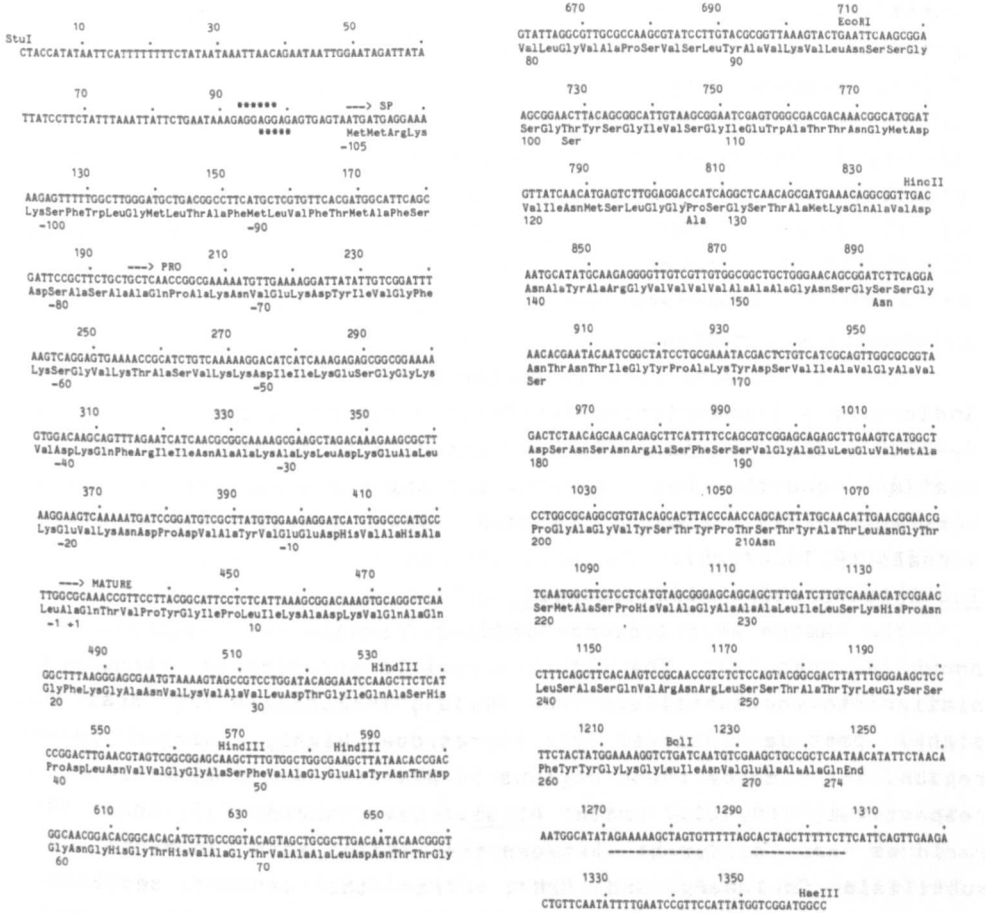


Fig. 2. Nucleotide and amino acid sequence of the subtilisin Carlsberg gene. Numbering of the amino acids starts at the N-terminal of the mature protein. Two possible ribosomal binding sequences (***) and a possible transcription termination sequence (---) are indicated. The amino acids that are different compared to the published sequence of the mature protein are shown. The putitative starting residues of the signal sequence (SP), the pro-region (PRO) and the mature protein (MATURE) are indicated.

two ribosomal binding sequences (12 and 9 nucleotides, respectively) suggests that this codon is used.

Although the expression data in B. subtilis suggests that a functional promoter does not exist in the 5'-flanking region of the gene, a number of sequences similar to promoters known to be

functional in B. subtilis can be found. An attractive sigma-43 promoter can be found at nucleotides 34 to 39 (-35) and 57 to 61 (-10), respectively. 5 out of 6 nucleotides are in both cases identical to the consensus sequence (TTGACA and TATAAT) and the spacing is the expected 17 nucleotides. In addition, there is a possible sigma-37 promoter at position 48 to 67 in which 8 out of 10 nucleotides are identical to the -10 consensus sequence GGAATTGTTT (11). However, no -35 region similar to the consensus of sigma-37 promoters can be detected upstream from the putitative -10 region.

The structural gene is followed by a palindromic sequence indicating a transcription termination signal as shown in Fig. 2 (---). The sequence is terminated in a stretch of T-rich residues and the free energy of the hairpin structure is -21.1 kcal/mole (22), similar to other Gram positive termination signals (9,10,27,28).

The deduced amino acid sequence.

The amino acid sequence deduced from the DNA sequence is shown in Fig. 2. The sequence reveals an overall structure similar to the subtilisin BPN' (9,10), including a 29 residue signal peptide followed by a 76 residue highly charged pro-region. The size of these regions in BPN' is 30 and 77 residues, respectively (9,10). Smith et al. have shown (6) that 84 residues are different between the amino acid sequence of subtilisin Carlsberg and BPN'. For the deduced sequence presented here, 81 out of the 84 residues are identical to the Carlsberg sequence. Furthermore, the proline residue reported to be missing at position 56 of subtilisin Carlsberg, as compared to BPN' (6), is lacking in the deduced sequence (Fig. 2). Thus, the predicted mature protein consists of the expected 274 residues, in contrast to the 275 residues of subtilisin BPN'. This confirms that the cloned gene is coding for subtilisin Carlsberg, as expected from the reactivity of the gene product to anti-subtilisin Carlsberg antisera.

The amino acid residues that differ from the earlier published protein sequence of the mature subtilisin Carlsberg (6) are also indicated in Fig. 2. There are only five differences observed, at residues 102, 128, 157, 160 and 211.

Most of them involve serine and aspargine residues, which, according to the authors (6), were hard to determine by the protein sequence method employed. However, since most of the differences can be explained by single point mutations, some of the divergence could be due to strain variation. Recently, the complete protein sequence of another Carlsberg-like subtilisin was reported (8). It is interesting to note that this enzyme, called subtilisin DY, contains a threonine, a proline and a serine at position 102, 128 and 157, respectively. Thus the residues at these postions are identical to the sequence reported here.

The signal sequence.

The putative signal sequence (Fig. 2) has similar properties to other signal sequences from Gram positive bacteria. This includes a basic N-terminal segment followed by a stretch of uncharged residues. The cleavage site most frequently follows the (-3,-1) rule, giving cleavage preferentially after the residues Ala-X-Ala (29). Based on this rule and other structural considerations (30), we predict that the signal peptidase processing occurs after the residues Ala-Ser-Ala as shown in Fig. 2. Abrahmsen et al. (31) have shown, that Gram positive signal peptides are unusual both in length and charge compared with other prokaryotic and eukaryotic organisms. The fact that the charge of the N-terminus is +3 and the size is 29, therefore demonstrates that the proposed subtilisin Carlsberg signal peptide is typical for Gram positive bacteria.

Homology plot analysis.

The complete amino acid sequence of four subtilisins have so far been determined, namely subtilisin Carlsberg, subtilisin BPN', subtilisin amylosacchariticus and subtilisin DY (8). Comparisons of their primary sequences have revealed that 62.5 % of all residues in these four subtilisins are identical. The structural studies also showed that the subtilisins may be divided into two groups according to their mutual similarities, i.e. subtilisin DY and Carlsberg into one, and subtilisin BPN' and amylosacchariticus into another (8). However, many amino acid similarities were also observed between, rather than within the groups, i.e. 14 of the substitutions between BPN' and

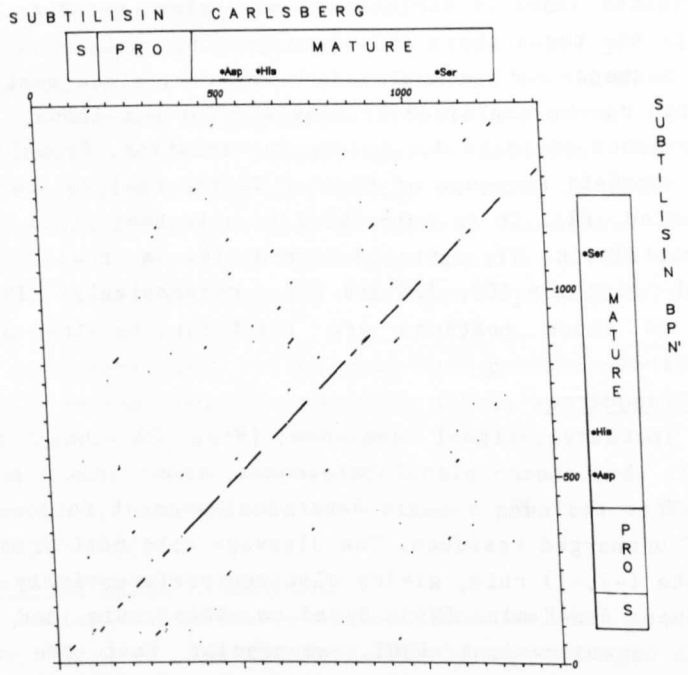


Fig. 3. Dot matrix comparison between the genes encoding subtilisin Carlsberg and BPN'. Each dot represents the center of a 21 nucleotide stretch where at least 14 nucleotides are identical (21). The numbers refer to the sequence of the subtilisin Carlsberg (Fig. 2) and subtilisin BPN' (9). A schematic drawing of the coding gene is shown for each axis with relevant features. The position of the amino acids (Asp, His and Ser) of the active site of the enzyme (1) is indicated.

amylosacchariticus are identical between BPN' and Carlsberg. The structural development of the different subtilisins must therefore have followed specific phylogenetic paths from a common ancestral precursor.

In order to clarify the structural homologies between the genes encoding BPN' and Carlsberg, homology plot analysis (27) was performed. As the genes probably have evolved from a common ancestral gene, a homologous region may imply evolutionary pressure to retain the amino acids. When the DNA sequences of subtilisin BPN' (9) and subtilisin Carlsberg (from Fig. 2) were analyzed, the plot shown in Fig. 3 was obtained. This reveals that the two sequences are highly conserved in the coding region for the mature protein as already suggested by the amino acid

sequences (6), but little homology is found in the 5' and 3' flanking regions of the gene. The 3'-end of the coding region can actually be distinguished simply by the sudden lack of homology between the sequences. In the 5'-end, the border between homologous regions and non-homologous regions is less defined. A conserved region can be observed in the C-terminal part of the pro-region, suggesting that there is a strong evolutionary pressure to retain the structure of this part of the protein. This region may be recognized by a proteolytic enzyme during maturation. In contrast, relatively little homology can be observed in the part coding for the N-terminus of the pro-region. The signal sequence has also diverged considerably between the two sequences confirming that signal peptides can vary considerably in primary structure (29).

It has been proposed from the analysis of the protein sequence (6) that subtilisins contain repetitions suggesting that the protein has evolved through amplification followed by point mutations. Detailed analysis of the DNA sequences that would be remnants of gene duplications, and especially searching for similar nucleotides in the wobble position of the codons, did not reveal any relationship (data not shown). This hypothesis can also be tested by comparing the DNA sequence of the gene with itself using a homology plot similar to the one shown in Fig. 3. Internal homology will show up as parallel lines outside the line of identity (29). Analysis of both the Carlsberg and the BPN' sequences did not reveal any related internal sequences (data not shown).

Secondary structure predictions.

Based on the primary amino acid sequence of the different subtilisins predictions of secondary stucture according to Chou and Fasman (32) and hydrophobicity according to Kyte and Dolittle (33) were performed. The result for the first 120 residues of subtilisin DY, BPN' and Carlsberg, are shown in Fig. 4. A remarkable similarity of α- and β-structures, as well as hydrophobicity, between the three proteases were encountered, in spite of the diverged primary sequence in particular between Carlsberg and the other two. The hydrophobicity prediction for the signal sequences show the expected pattern of a hydrophilic (basic) N-terminus followed by a hydrophobic stretch of

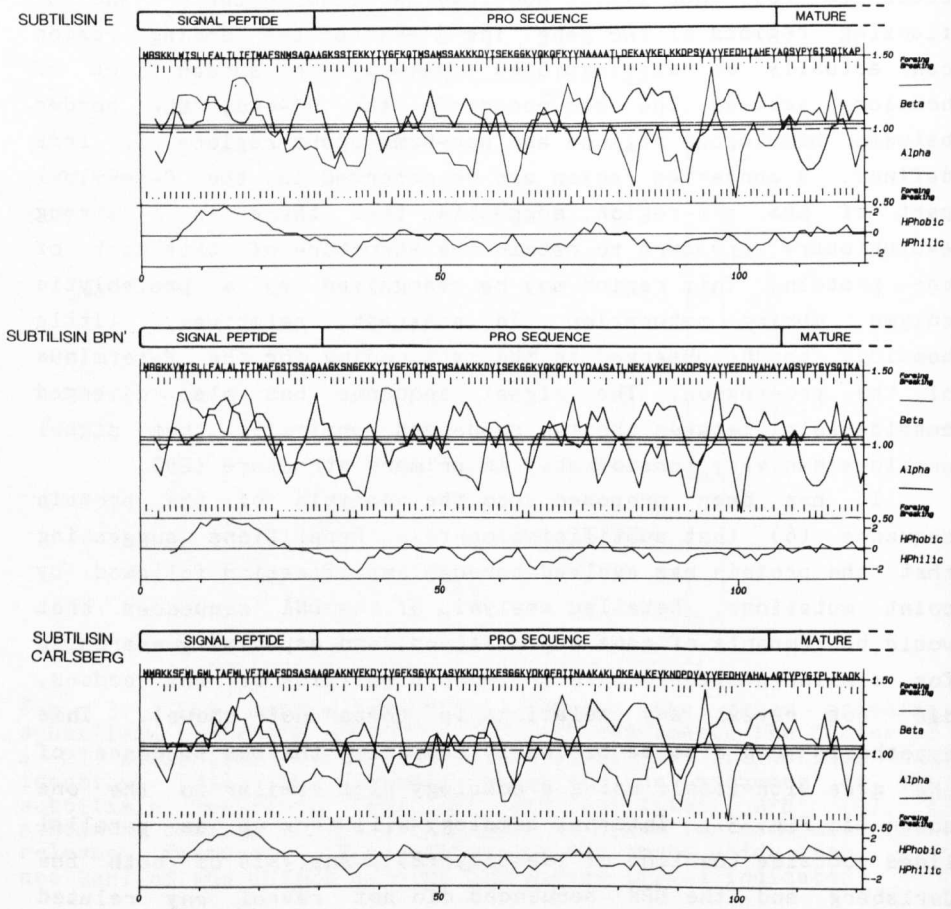


Fig. 4. Secondary predictions of the gene encoding the N-terminal part of subtilisin DY (11), BPN' (9) and Carlsberg (this paper). Each diagram show from top to bottom a schematic drawing of the coding regions,the amino acid sequence, the residues that break or form α-helix structures, the probability for α-structures and β-structures (30), respectively, the residues that break and form β-structures and the hydrophobicity of the sequence (31).

approximately 20 residues. The proposed peptidase cleavage site is for all three proteins located at the border between hydrophobic and hydrophilic sequences. All three sequences predict an α-helix structure around this site, but the possible functional significance of this is unclear.

The homology plot analysis revealed little homology in the

N-terminal end of the pro-region of subtilisin BPN' and Carlsberg. A comparison residue for residue yields around 60% homology. It is therefore interesting that the predictions of secondary structure show high similarities throughout the whole pro-region (Fig. 4). This indicates that the tertiary structure of this region is conserved. Furthermore, certain stretches of amino acids are identical in all three proteins, i.e. the sequence Tyr-Ile-Val-Gly-Phe-Lys which is 10 residues from the signal peptidase cleavage site. This sequence may be involved in protein interaction during maturation of the preprotein and thus under strong evolutionary pressure.

In this paper, we have characterized the gene encoding subtilisin Carlsberg. The gene coding for subtilisin BPN' has previously been characterized (9,10). Having both genes, it is now possible to perform refined biochemical studies with genetic approaches to reveal how serine proteases distinguish between different substrates or why the kinetic parameters for these two enzymes are significantly different for some substrates (1). Site-specific mutagenesis can be used to establish the biochemical basis for these differences. Furthermore, it might be possible to produce hybrid protein between the two enzymes with altered enzymatic activities. As the subtilisin can be used for stereo-specific hydrolysis of organic molecules, such studies might create new enzymes engineered to specifically catalyze these reactions.

ACKNOWLEDGEMENTS

This investigation was supported by grants from the Swedish Natural Research Council and the Swedish Natural Board for Technical Development. We are grateful to Drs Lennart Philipson, Niels Fiil and Børge Diderichsen for valuable advice and critical comments. We thank Heide Seifert for patient secretarial help.

*To whom correspondence should be addressed at: European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG

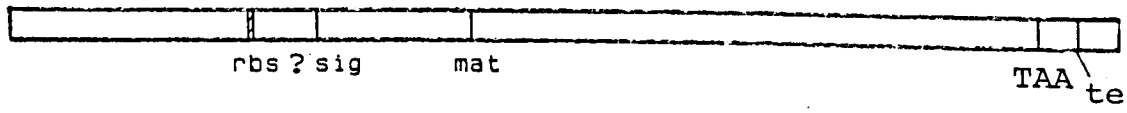
REFERENCES

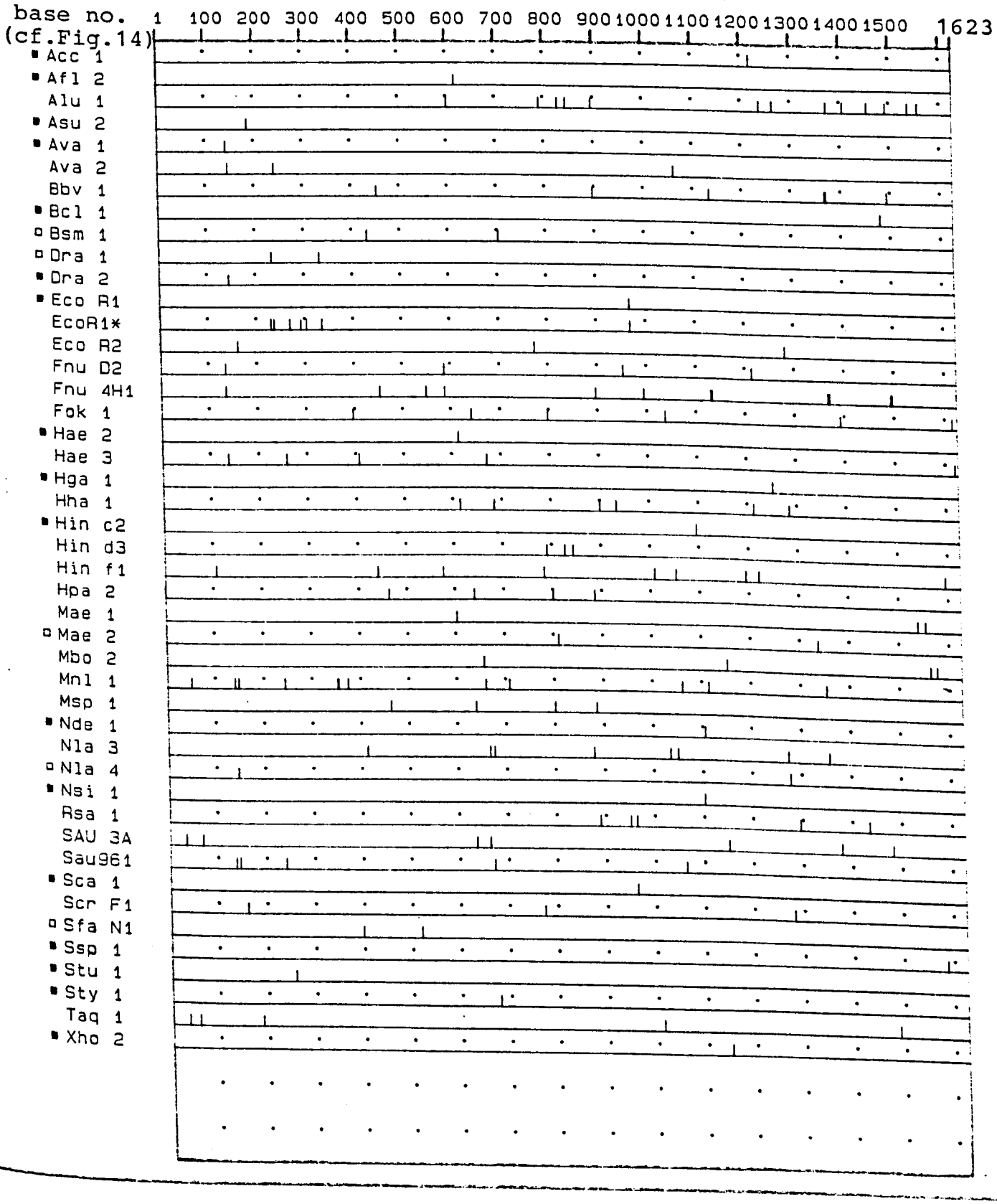
1. Ottesen, M. and Svendsen, I. (1971) in The Enzymes, ed. P.D. Boyer Vol. III, Acad. Press pp. 199-215.
2. Markland, F.S. and Smith, E.L. (1971) in The Enzymes, ed. P.D. Boyer Vol. III, Acad. Press pp. 561-608.
3. Guntelberg, A.V. and Ottesen, M. (1954), Compt Rend. Trav. Lab. Carlsberg 29, 36-41.

4. Hagihara, B. (1954) Ann. Rev. Sci. Works. Fac. Sci. Osaka Univ. 2, 35-38.
5. Keay, l. and Moser, P.W. (1969) Biochem. Biophys. Res. Commun. 34, 600-605.
6. Smith, E.S., DeLange, R.J., Evans, W.H., Landon, M. and Markland, F.S. (1968) J. Biol. Chem. 243, 2184-2191.
7. Aunstrup, K., Anderson, O., Falck, E.A. and Nielsen, T.K. (1979) in Microbial Technology, eds. Peppler, H.J. and Terlman, D. Vol. I, Acad. Press, pp. 281-309.
8. Nedkov, P., Oberthur, W. and Braunitzer, G. (1985) Biol. Chem. Hoppe-Seyer 366, 421-430.
9. Wells, J.A., Ferrari, E., Henner, D.J., Estell, D.A. and Chen, E.Y. (1983) Nucl. Acids Res. 11, 7911-7925.
10. Vasantha, N., Thompson, L.D., Rhodes, C., Banner, C., Nagel, J. and Filpula, D. (1984) J. of Bacteriol. 159, 811-819.
11. Wong, S., Price, C., Goldfarb, D.S. and Doi, R.H. (1984) Proc. Natl. Acad. Sci. USA 81, 1184-1188.
12. Maniatis, T., Fritsch, E.F. and Sambrock, J. (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, N.Y.
13. Stahl, M. and Ferrari, E. (1984) J. Bacteriol. 158, 411-418.
14. Gryczan, T., Shivakumar, A.G. and Dubnau, D. (1980) J. Bacteriol. 141, 246-253.
15. Löfdahl, S., Guss, B., Uhlén, M., Lindberg, M. and Philipson, L. (1983) Proc. Natl. Acad. Sci. USA 80, 697-701.
16. Kieser, T. (1984) Plasmid 12, 19-36.
17. Uhlén, M., Guss, B., Nilsson, B., Götz, F. and Lindberg, M. (1984) J. of Bacteriol. 159, 713-719.
18. Southern, E. (1975) J. Mol. Biol. 98, 503-517.
19. Nilsson, B., Holmgren, E., Josephson, S., Gatenbeck, S., Philipson, L. and Uhlén, M. (1985) Nucl. Acids Res. 13, 1151-1162.
20. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
21. Olsson, A., Moks, T., Uhlén, M. and Gaal, A. (1984) J. Biochem. and Biophys. Meth. 10, 83-90.
22. Devereaux, J., Häberli, P. and Smithies, O. (1984) Nucl. Acids Res. 12, 387-395.
23. Hudson, L. and Hay, F.C. (1976) Practical Immunology, Backwell Scientific.
24. Ryden, A.-C., Lindberg, M. and Philipson, L. (1973) J. Bacteriol. 116, 25-32.
25. Lämmli, U.K. (1970) Nature 227, 680-685.
26. McLaughin, J.R., Murray, C.L. and Rabinowitz, J.C. (1981) J. Biol. Chem. 256, 11283-11291.
27. Uhlén, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L. and Lindberg, M. (1984) J. Biol. Chem. 257, 1695-1702.
28. Takkinen, R., Pettersson, R.F., Kalkkinen, N., Palva, I., Söderlund, H. and Käriäinen, L. (1983) J. Biol. Chem. 258, 1007-1013.
29. von Heine, G. (1984) J. Mol. Biol. 173, 243-251.
30. von Heine, G. (1985) J. Mol. Biol. 184, 99-105.
31. Abrahmsen, L., Moks, T., Nilsson, B., Hellman, U. and Uhlén, M. (1985) EMBO J. submitted.
32. Chou, G. and Fasman, G.D. (1978) Adv. Enz. 47, 45-147.
33. Kyte, J. and Dolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.

APPENDIX TWO

Restriction Sites in Sequence : SUBTILISIN Carlsberg





APPENDIX 3

Determination of mRNA Start Site By Primed cDNA Synthesis

As discussed in III.8.2, the S1 mapping data indicated that there were 2 subtilisin Carlsberg transcriptional start sites, and their location could be approximated (Fig.14: P1,P2; Fig.26).

To ascertain more precisely where the postulated P1 and P2 were situated, cDNA was made in collaboration with Helle Wøldike, by the method described briefly below (see II.18 for details). A 5'-labelled 16-mer synthetic primer (indicated in Fig.14, complementary to bases 393 to 409), reverse transcriptase, and T12 RNA from MFJ29 were used.

The labelled cDNA was visualized on, and isolated from, a 20% denaturing gel (Fig.42a, 4 hour exposure). Note that there is only one band in the region which would correspond to a strand terminating in the region of P1 or P2. This strand has an estimated length of 105 b, judging from the labelled markers. Upon prolonged exposure, a second, discrete, shorter fragment was detected. It seemed to be about 50 b long. It may be accounted for either by premature termination, possibly due to template 2^{o} structure, or more probably to primed synthesis on the "wrong" template.

The upper cDNA band was excised after autoradiography and extracted from the gel (c.5cps recovered), and the purine sequencing reaction performed (II.14.1). The products were analysed on an 10% 40cm. denaturing gel. Chemical sequencing reactions of another, unrelated fragment were used as size markers (Fig. 42B).

Comparison with the known sequence (Fig.14, coding strand shown) reveals that the cDNA terminates at bases 297 and 298 (on the strand complementary to that shown in Fig.14). The last chemical reaction product that can be seen (Fig.42B), corres-

FIGURE 42: PRECISE TRANSCRIPTIONAL START SITE DETERMINATION BY cDNA SYNTHESIS AND SEQUENCING

(A) cDNA SYNTHESIS

cDNA was synthesized from primer 299 on MFJ29 T12 RNA. Products of the reverse transcriptase reaction were electrophoresed on a 20% denaturing gel and visualized by autoradiography.

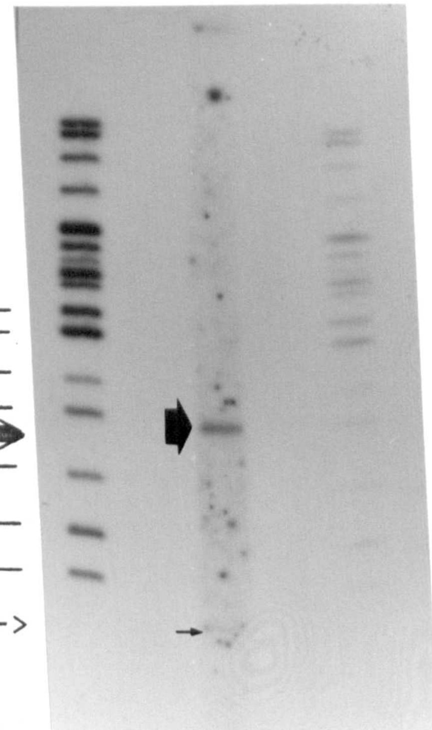
outer lanes: pBR322/ MspI size markers

central lane: cDNA

broad arrow: species whose sequence is shown below in (B).

fine arrow: position of less intense 2nd species

b
160
147
123
110
90
76
67
->



(B) cDNA SEQUENCING

The cDNA of approx. 105b was excised from gel shown in (A) and subjected to the A&G reaction (II.14.1).

Maxam and Gilbert sequencing reactions of Fragment X were used as markers.

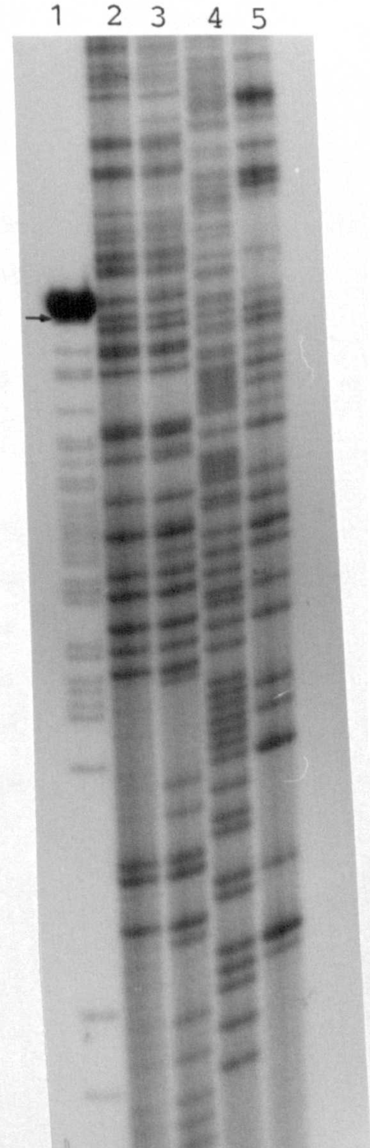
The results of the chemical sequencing were analysed on a 10% denaturing gel.

lane 1: cDNA A&G reaction
lanes 2 to 5: Fragment X, resp. G,A&G,T&C and C reactions.

The arrow shows the position of the last G that can be read off the cDNA. The purine pattern shown in lane 1 should be compared with the pattern of pyrimidines in the subtilisin Carlsberg sequence 5´to the coding region shown in Figure 14.

1 2 3 4 5

->



ponds to a labelled chain ending at the C (or properly, G in cDNA), at base 299. At a distance of 2 nt from this fragment (as estimated from the marker tracks), is an intense band (= unreacted full-length labelled cDNA). This resolves into 2 discrete bands on brief exposure (not shown). There appears thus to be a heterogeneity of 1 nt in the length of the cDNA synthesized. This may be an artefact of the synthesis, or else reflect heterogeneity in the template.

Base 297 must be regarded as representing the closest possible transcriptional initiation position in relation to the gene coding sequences, since it reflects the minimum length of the subtilisin-specific mRNA which served as template for cDNA synthesis.

Just before the putative transcriptional start site (bases 286 -291), (and also between 268-273), lies a typical -10 region for sigma-43 (cf.Table 1). However, at the permitted interval of 16-18bp upstream from these hexanucleotides (see I.3.4), no match with the canonical -35 region can be found.

These results illustrate clearly why no expression of subtilisin Carlsberg was obtained with clones interrupted at the StuI, for it lies at most 36 bp from the mRNA start site, and undoubtedly is involved in direct interaction with RNA polymerase (I.3.1).

It remains to be established why a discrepancy exists between the S1 mapping results (2 well-separated start sites), and the cDNA synthesis (quite clearly only 1 species - albeit with microheterogeneity at the 5' end - in the correct size range). To resolve the question, the S1 experiment ought perhaps to be repeated, possibly utilizing the NlaIII site, as discussed in III.8.2.

APPENDIX 4

Transformation of B. licheniformis by Plasmid DNA

It was desirable to have a plasmid transformation system for B. licheniformis both to facilitate cloning (see III.9) and for studying the expression of fragments cloned from the organism in the species of origin.

A protoplast transformation procedure applicable to B. licheniformis has been described (Akamutsu & Sekiguchi, 1982), where selection for Cm^R transformants was on a regeneration medium containing up to 25 microgram/ml of antibiotic.

When transformation of MFJ29 by pC194 was attempted by a modification of this method (II.4.5), regenerants acquired a Cm-resistant phenotype regardless whether they had been treated with plasmid or not. This was not the case with B. subtilis protoplasts treated in parallel as a control. Some Bacillus spp. have chromosomally-determined Cm resistance, eg. B.pumilus (Shaw,1975; Harwood et al,1983). It is not known whether B. licheniformis is such a species. Cells spread directly on to Cm-containing plates never display resistance.

Transformation with chromosomal DNA has a longer tradition in the species (Dubnau & Pollack, 1965). It appears that classes of spontaneously arising morphological variants differ widely in their ability to be transformed (Leonard & Mattheis, 1965).

A variety of B. licheniformis strains that have been used in transformation studies were acquired from BGSC. They included 5A4,5,13,15,16,20,21,23,24. These were screened on Penassay plates containing a variety of antibiotics. The results of this study are found in Table 4. There was widespread resistance to low levels of Tet, Emy, Cm, and Km.

Different classes of transformable mutants were shown to require different media for the development of competence (Thorne & Stull, 1966). The spp. listed above were grown to competence in both BLSG and NBSG-X as described in II.4.4.1.

They were transformed with oligomerized, Gram-positive plasmids carrying Km^R, or Cm^R, and that had either no homologous chromosomal sequences, or else DNA derived from MFJ29. In no case was a transformant obtained. It was also the experience of Imanaka et al (1981), that this procedure could not be applied to plasmid transformation.

A strain AL20222 was obtained from Apothekenes Laboratorium. It had been mutagenized and reportedly could take up plasmid DNA by the method described in II.4.4.2, using special media. In 3 attempts to transform the strain with pUB110, there were consistently 10x more apparently Km^R colonies in test samples than in controls (not treated with plasmid). These colonies were however not viable upon repeated repurification.

From the antibiotic screen described above, it would appear that high concentrations of Cm, Km and Tet are suitable for selection of transformants.

The most promising approach seems to lie in protoplast transformation, with Km rather than Cm selection (Imanaka et al, 1981; P-E Pederson, personal communication).

Abrahmsen, L., Moks, T., Nilsson, B., Hellman, U. and Uhlen, M. (1985). EMBO J. 4, 3901-3906: Analysis of signals for secretion in the staphyloccoccal protein A gene.

Achberger, E.C. and Whiteley, H.R. (1980).J. Biol. Chem. 255, 11964-11957: The interaction of E. coli core RNA polymerase with specificity-determining subunits derived from unmodified and SP82-modified B. subtilis RNA polymerase.

Adler, L-A., and Arvidson, S. (1984). FEMS Microbiol. Letts. 23, 17-20: Immunological crossreaction between proteins supposed to be involved in protein secretion in S. aureus and B. subtilis.

Akamatsu, T. and Sekiguchi, J. (1982). Agric. Biol. Chem. 46, 1617-1621: Transformation of Bacillus protoplasts by plasmid pTP4 DNA.

Akamatsu, T. and Sekiguchi, J. (1984). Agric. Biol. Chem. 48 , 651-655: An improved method of protoplast regeneration for Bacillus species and its application to protoplast fusion and transformation.

Akiyama, Y. and Ito, K. (1985). EMBO J. 4, 3351-3356: The SecY membrane component of the bacterial protein export machinery: analysis by new electrophoretic mehods for integral membrane proteins.

Albertini, A.M., Hofer, M., Calos, M.P. and Miller, J.H. (1982). Cell 29, 319-328: On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions.

Ambulos, N.P. Jr., Mongkolsuk, S., Kaufman, J.D. and Lovett, P.S. (1985). J. Bact. 164, 696-703: Chloramphenicol-induced translation of cat-86 mRNA requires two cis-acting regulatory regions.

Ambulos, N.P.Jr., Duvall, E.J. and Lovett, P.S. (1986). J. Bact. 167, 842-849: Analysis of the regulatory sequences needed for induction of chloramphenicol acetyltransferase gene cat-86 by chloramphenicol and amicetin.

Aoyama, T. and Takanami, M. (1985). Nucl. Ac. Res. 13, 4085-4096: Essential structure of E. coli promoter II. Effect of the sequences around the RNA start point on promoter function.

Aronson, A.I., Angelo, N. and Holt, S.C. (1971). J. Bact. 106, 1016-1025: Regulation of extracellular protease production in B. cereus T. Characterization of mutants producing altered amounts of protease.

Artman, M. and Westhamer, S. (1974). J. Bact. 120, 980-983: Effect of cyclic guanosine 3',5'-monophosphate on the synthesis of enzymes sensitive to catabolite repression in intact cells of E. coli.

Aubert, E., Klier, A. and Rapoport, G. (1985). J. Bact. 161, 1182-1187: Cloning and expression in E. coli of the regulatory sacU gene from B. subtilis.

Aunstrup, K., Anderson, O., Falck, E.A. and Nielsen, T.K. (1979). Production of Microbial Enzymes. In Microbial Technology, Vol. I, pp. 281-309 (Peppler, H.J. & Perlman, D. eds.) (Academic Press).

Aymerich, S., Gonzy-Treboul, G. and Steinmetz, M. (1986). J. Bact. 166, 993-998: 5'-noncoding region sacR is the target of all identified regulation affecting the levansucrase gene in B. subtilis.

Ayusawa, D., Yoneda, Y., Yamane, K. and Maruo, B. (1975). J. Bact. 124, 459-469: Pleiotropic phenomena in autolytic enzyme(s) content, flagellation, and simultaneous hyper-production of extracellular alpha-amylase and protease in a B. subtilis mutant.

Bachmann, B.J. and Low, K.B. (1980). Microbiol. Rev. 44, 1-56: Linkage map of E. coli K-12, edition 6.

Bale, A., d'Alarco, M. and Marinus, M.G. (1979). Mutation Research, 59, 157-165: Characterization of DNA adenine methylation mutants of E. coli K-12.

Ballassa, G. (1969). Mol. Gen. Genet. 104, 73-103: Biochemical genetics of Bacterial Sporulation I. Unidirectional pleio-tropic interactions among genes controlling sporulation in B. subtilis.

Banner, C.D.B., Moran, C.P. Jr. and Losick, R. (1983). J. Mol. Biol. 168, 351-365: Deletion analysis of a complex promoter for a developmentally regulated gene from B. subtilis.

Benson, S.A., Hall, M.N. and Silhavy, T.J. (1985). Ann. Rev. Biochem. 54, 101-134: Genetic analysis of protein export in E. coli K-12.

Berk, A.J. and Sharp, P.A. (1977). Cell 12, 721-732: Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1-endonuclease-digested hybrids.

Bernlohr, R.W. (1964). J. Biol. Chem. 239, 538-543: Post log phase metabolism of sporulating micro-organisms I. Protease of B. licheniformis.

Bernlohr, R.W., Maddox, M.K. and Goldberg, N.D. (1974). J. Biol. Chem. 249, 4329-4331: Cyclic guanosine 3',5'-mono-phosphate in E. coli and B. licheniformis.

Binnie, C., Lampe, M. and Losick, R. (1986). Proc. Nat. Acad. Sci. USA 83, 5943-5947: Gene encoding the sigma-37 species of RNA polymerase sigma factor from B. subtilis.

Blobel, G. and Dobberstein, B. (1975). J. Cell. Biol. 67, 835-851: Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immuno-globulin light chains on membrane-bound ribosomes of murine myeloma.

Blow, D.M., Birktoft, J.J. and Hartley, B.S. (1969). Nature 221, 337-340: Role of a buried acid group in the mechanism of action of chymotrypsin.

Bochner, B.R., Huang, H., Schieven, G.L. and Ames, B.N. (1980). J. Bact. 143, 926-933: Positive selection for loss of tetracycline resistance.

Bode, W., Papaoloos, E., Musil, D., Seemueller, U. and Fritz, H. (1986) EMBO J. 5, 813-818: Refined 1.2 A crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin C. Molecular structure of eglin and its detailed interaction with subtilisin.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Berlach, M.C., Heyneker, H.L. and Boyer, H.W. (1977). Gene 2, 95-113: Construction and characterization of new cloning vehicles. II.A Multipurpose cloning system.

Both, G.W., McInnes, J.L., Hanlon, J.E., May, B.K. and Elliot, W.H. (1972). J. Mol. Biol. 67, 199-207: Evidence for an accumulation of messenger RNA specific for extracellular protease and its relevance to the mechanism of enzyme secretion in bacteria.

Bott, K.F. and Davidoff-Abelson, R. (1966). J. Bact. 92, 229-240: Altered sporulation and respiratory patterns in mutants of B. subtilis induced by acridine orange.

Bouvier, J., Stragier, P., Bonamy, C. and Szulmajster, J. (1984). Proc. Nat. Acad. Sci. USA 81, 7012-7016: Nucleotide sequence of the spoOB gene of B. subtilis and regulation of its expression.

Boyer, H.W. and Carlton, B.C. (1968). Arch. Biochem. Biophys. 128, 442-455: Production of two proteolytic enzymes by a transformable strain of B. subtilis.

Boylan, R.J., Mendelson, N.H., Brooks, D. and Young, F.E. (1972). J. Bact. 110, 281-290: Regulation of the bacterial cell wall: Analysis of a mutant of B. subtilis defective in biosynthesis of teichoic acid.

Brehm, S.P., Staal, S.P. and Hoch, J.A. (1973). J. Bact. 115, 1063-1070: Phenotypes of pleiotropic-negative sporulation mutants of B. subtilis.

Briat, J-F., Gilman, M.Z. and Chamberlin, M.J. (1985). J. Biol. Chem. 260, 2038-2041: B. subtilis sigma-28 and E. coli sigma-32 (HptR) are minor sigma factors that display an overlapping promoter specificity.

Brosius, J., Cate, R.L. and Perlmutter, A.P. (1982). J. Biol. Chem. 257, 9205-9210: Precise location of two promoters for the beta-lactamase gene of pBR322.

Brosius, J. (1984). Gene 27, 161-172: Toxicity of an overproduced foreign gene product in E. coli and its use in plasmid vectors for the selection of transcriptional terminators.

Buck, M., Miller, S., Drummond, M. and Dixon, R. (1986). Nature, 320, 374-378: Upstream activator sequences are present in the promoters of nitrogen fixation genes.

Burnett, T.J., Shankweiler, G.W. and Hageman, J.H. (1986). J. Bact. 165, 139-145: Activation of Intracellular Serine Proteinase in B. subtilis cells during sporulation.

Burns, D.M. and Beacham, I.R. (1985). Febs Letts. 189, 318-324: Rare codons in E. coli and S. typhimurium signal sequences.

Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S. and Gross, C.A. (1981). Nucl. Acids Res. 9, 2889-2903: The nucleotide sequence of the cloned rpoD gene for the RNA polymerase sigma subunit from E. coli K-12

Burton, Z.F., Gross, C.A., Wanatabe, K.K. and Burgess, R.R. (1983). Cell 32, 335-349: The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in E. coli K-12.

Calos, M.P. and Miller, J.H. (1980). Cell 20, 579-595:Transposable Elements.

Canosi, U., Morelli, G. and Trautner, T.A. (1978). Mol. Gen. Genet. 166, 259-267: The relationship between molecular structure and transformation efficiency of some S. aureus plasmids isolated from B. subtilis.

Canosi, U., Iglesias, A. and Trautner, T.A. (1981). Mol. Gen. Genet. 181, 434-440: Plasmid transformation in B. subtilis: Effects of insertion of B. subtilis DNA into plasmid pC194.

Casabadan, M.J. and Cohen, S.N. (1980). J. Mol. Biol. 138, 179-207: Analysis of gene control signals by DNA fusion and cloning in E. coli.

Cashel, M. and Gallant, J. (1969). Nature 221, 838-841: Two compounds implicated in the function of the RC gene of E. coli.

Caulfield, M.P., Horiuchi, S., Tai, P.C. and Davis, B.D. (1984). Proc. Nat. Acad. Sci. USA 81, 7772-7776: The 64-kilodalton membrane protein of B. subtilis is also present as a multi-protein complex on membrane-free ribosomes.

Caulfield, M.P., Furlong, D., Tai, P.C. and Davis, B.D. (1985). Proc. Nat. Acad. Sci. USA 82, 4031-4035: Secretory S complex of B. subtilis forms a large, organized structure when released from ribosomes.

Chang, S. and Cohen, S.N. (1979). Mol. Gen. Genet. 168, 111-115: High frequency transformation of B. subtilis protoplasts by plasmid DNA.

Chelm, B.K., Duffy, J.J. and Geiduschek, E.P. (1982). J. Biol. Chem. 257, 6501-6508:Interaction of B. subtilis RNA polymerase core with two specificity-determining subunits.

Chenchick, A., Beabealashvilli, R. and Mirzabekov, A. (1981). FEBS Letts. 128, 46-50: Topography of interaction of E. coli RNA polymerase subunits with lac UV5 promoter.

Cheng, Y.S.E. and Aronson, A.I. (1977). Proc. Nat. Acad. Sci. USA 74, 1254-1258: Alteration of spore coat processing and protein turnover in a B. cereus mutant with a defective post-exponential intracellular protease.

Chi, N-Y.W., Ehrlich, S.D. and Lederberg, J. (1978). J. Bact. 133, 816-821: Functional expression of two B. subtilis chromosomal genes in E. coli.

Chirgwin, J.W., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979). Biochemistry, 18, 5294-5299: Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease.

Chou, P.Y. and Fasman, G.D. (1978). Adv. Enz. 47, 45-147: Prediction of the secondary structure of proteins from their amino acid sequence.

Coleman, G. (1967). J. Gen. Microbiol. 49, 421-431: Studies on the regulation of extracellular enzyme formation by B. subtilis.

Connors, M.J., Mason, J.M. and Setlow, P. (1986). J. Bact. 166, 417-425: Cloning and nucleotide sequencing of genes for three small, acid-soluble proteins from B. subtilis spores.

Constanzo, M. and Pero, J. (1983). Proc. Nat. Acad. Sci. USA 8o, 1236-1240: Structure of a B. subtilis bacteriophage SPo1 gene encoding a RNA polymerase sigma factor.

Constanzo, M., Brzyustowicz, L., Hannett, N. and Pero, J. (1984). J.Mol. Biol. 180, 533-547: Bacteriophage SP01 genes 33 and 34: Location and primary structure of genes encoding regulatory subunits of B. subtilis RNA polymerase.

Contente, S. and Dubnau, D. (1979). Plasmid 2, 551-571: Marker rescue transformation by linear plasmid DNA in B. subtilis.

Corfield, V.A., Reid, S.J., Bodmer, J. and Thomson, J.A. (1984). Gene, 30, 17-22: A modified protoplast-regeneration protocol facilitating the detection of cloned exoenzyme genes in B. subtilis.

Dagert, M., Jones, I., Goze, A., Romac, S., Niaudet, B. and

Ehrlich, S.D. (1984). EMBO J. 3, 81-86: Replication functions of pC194 are necessary for efficient plasmid transduction by M13 phage.

Dancer, B.N. and Mandelstam, J. (1975a). J. Bact. 121, 406-410: Production and possible function of serine protease during sporulation of B. subtilis.

Dancer, B.N. and Mandelstam, J. (1975b). J. Bact. 121, 411-415: Criteria for categorizing early biochemical events occuring during sporulation of B. subtilis.

Date, T. (1983). J. Bact. 154, 76-83: Demonstration by a novel genetic technique that leader peptidase is an essential enzyme of E. coli.

Davison, J. (1984). Gene 28, 1-15: Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids- A review.

Debabov, V.G. (1982). The industrial use of bacilli. In The Molecular Biology of the Bacilli, Vol. I. Dubnau, D.A. ed. (Academic Press), pp. 331-37o.

De Boer, H.A., Hui, A., Comstock, L. J., Wong, E. and Vasser, M. (1983). DNA 2, 231-235: Portable Shine-Dalgarno regions: A systematic study of defined alterations of nucleotide sequences within E. coli ribosome-binding sites.

DelMar, E.G., Largman, C., Brodrick, J.W. and Geokas, M.C. (1979). Anal. Biochem. 99, 316-320: A sensitive new substrate for chymotrypsin.

Devereaux, J., Haberli, P. and Smithies, O. (1984). Nucl. Acids Res. 12, 387-395: A comprehensive set of sequence analysis programs for the VAX.

DiCioccio, R.A. and Strauss, N. (1973). J. Mol. Biol. 77, 325-336: Patterns of transcription in B. subtilis during sporulation.

Dod, B. and Balassa, G. (1978). Mol. Gen. Genet. 163, 57-63: Spore control (Sco) mutations in B. subtilis.I.

Dod, B., Balassa, G., Raulet, E. and Jeannoda, V. (1978). Mol. Gen. Genet. 163, 45-56: Spore control (Sco) mutations in B. subtilis.II.

Doi, R.H. and Igarashi, R.T. (1964). Proc. Nat. Acad. Sci. USA 52, 755-762: Genetic transcription during morphogenesis.

Doi, R.H. (1982): RNA polymerase of B. subtilis. In the Molecular Biology of the Bacilli, Vol. I. Dubnau, D.A. ed. (Academic Press), pp. 71-110.

Drenth, J., Hol, W.G.J., Janosius, J.N. and Koekoek, R.

(1972). Eur. J. Biochem. 26, 177-181: Subtilisin Novo: The 3-D structure and its comparison with subtilisin BPN'.

Dubnau, D. (1982): Genetic transformatin in B. subtilis. In The Molecular Biology of the Bacilli, Vol. I. Dubnau, D.A. ed. (Academic Press), pp. 147-175

Duffy, J.J. and Geiduschek, E.P. (1977). Nature, 270, 28-32: Purification of a positive regulatory subunit from phage SP01-modified RNA polymerase.

Duncan, C.H., Wilson, G.A. and Young, F.E. (1978). Proc. Nat. Acad. Sci. USA. 75, 3664-3668:Mechanism of integrating foreign DNA during transformation of B. subtilis.

Ehrlich, S.D. (1978). Proc. Nat. Acad. Sci. USA 75, 1433-1436: DNA cloning in B. subtilis.

Ehrlich, S.D., Niaudet, B. and Michel, B. (1981): Use of plasmids from S. aureus for cloning of DNA in B. subtilis. In Current Topics in Microbiol. Immunol. 96, 19-29.

Emr, S.D., Hedgpeth, J., Clement, J-M, Silhavy, T.J. and Hofnung, M. (1980). Nature, 285, 82-85: Sequence analysis of mutations that prevent export of lambda receptor, an E. coli outer membrane protein.

Endo, S. (1962). J. Fermentation Technol. 40, 346-353: Studies on protease produced by thermophilic bacteria.

Errington, J., Fort, P. and Mandelstam, J. (1985). FEBS. Lett. 188, 184-188: Duplicated sporulation genes in bacteria: Implications for simple developmental systems.

Espinosa, M., Lopez, P. and Lacks, S.A. (1984). Gene 28, 301-310: Transfer and expression of recombinant plasmids carrying pneumococcal mal genes in B. subtilis.

Estell, D.A., Graycar, T.P. and Wells, J.A. (1985). J. Biol. Chem. 260, 6518-6521: Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation.

Fahnestock, S.R. and Fisher, K.E. (1986). J. Bact. 165, 796-804: Expression of the staphylococcal protein A gene in B. subtilis by gene fusions utilizing the promoter from a B. amyloliquefaciens alpha-amylase gene.

Fahnestock, S.R., Saunders, C.W., Guyer, M.S., Lofdahl, S., Guss, B., Uhlen, M. and Lindberg, M. (1986). J. Bact. 165, 1011-1014: Expression of the staphylococcal protein A gene in B. subtilis by integration of the intact gene into the B. subtilis chromosome.

Ferrari, E. and Hoch, J.A. (1983). Mol. Gen. Genet. 189, 321-325: A single copy, transducible system for complemen tation and dominance analyses in B. subtilis.

Ferrari, E., Henner, D.J. and Yang, M.Y. (1985). Biotechnology 3, 1003-1007: Isolation of an alanine racemase gene from B. subtilis and its use for plasmid maintenance in B. subtilis.

Ferrari, E., Howard, S.M.H. and Hoch, J.A. (1986). J. Bact. 166, 173-179: Effect of stage 0 sporulation mutations on subtilisin expression.

Ferrari, F.A., Nguyen, A., Lang, D. and Hoch, J.A. (1983). J. Bact. 154, 1513-1515: Construction and properties of an integrable plasmid for B. subtilis.

Ferrari, F.A., Trach, K., LeCoq, D., Spence, J., Ferrari, E. and Hoch, J.A. (1985). Proc. Nat. Acad. Sci. USA 82, 2647-2651: Characterization of the spo0A locus and its deduced product.

Fersht, A.S. (1985). In Enzyme Structure and Mechanism, 2nd Edition. (W.H. Freeman & Co., New York).

Fiil, N.P., Willumsen, B.M., Friesen, J.D. and von Meyenburg, K. (1977). Mol. Gen. Genet. 150, 87-101: Interaction of alleles of the relA, relC, and spoT genes in E. coli: Analysis of the interconversion of GTP, ppGpp and pppGpp.

Fisher, S.H. and Magasanik, B. (1984). J. Bact. 157, 942-944: Isolation of B. subtilis mutants pleiotropically insensitive §o glucose catabolite repression.

Fisher, S.H. and Sonenshein, A.L. (1984). J. Bact. 157, 612-621: B. subtilis glutamine synthetase mutants pleiotropically altered in glucose catabolite repession.

Fishman, S.E. and Hershberger, C. (1983). J. Bact. 155, 459-466: Amplified DNA in Stretomyces fradiae.

Fliss, E.R. and Setlow, P. (1985). Gene, 151-157: Genes for B. megaterium small, acid-soluble spore proteins: nucleotide sequence of two genes and their expression during sporulation.

Fliss, E.R., Loshon, C.A. and Setlow, P. (1986). J. Bact. 165, 467-473: Genes for B. megaterium small, acid-soluble spore protein: Cloning and nucleotide sequence of three additional genes from this multigene family.

Fogarty, W.M., Griffin, P.J. and Joyce, A.M. (1974a). Process Biochem. 9, (6), 11-18, 24: Enzymes of Bacillus species - Part I.

Fogarty, W.M., Griffin, P.J. and Joyce, A.M. (1974b). Process Biochem. 9, (7), 27-34: Enzymes of Bacillus species - Part II.

Fox, J.L. (1985). ASM News 51, 566-568: Protein Engineering: Biotechnology's new wave.

Frankena, J., van Versefeld, H.W. and Stouthamer, A.H. (1985).

Appl. Microbiol. Biotechnol. 22, 169-176: A continuous culture study of the bioenergetic aspects of growth and production of exocellular protease in B. licheniformis.

Fujii, M., Takagami, M., Imanaka, T. and Aiba, S. (1983). J. Bact. 154, 831-837: Molecular cloning of a thermostable neutral protease gene from B. stearothermophilus in a vector plasmid and its expression in B. stearothermophilus and B. subtilis.

Fukumoto, J. and Negoro, H. (1951). Proc. Japan Acad. 27, 441-444: Crystallization of bacterial proteinase.

Gay, P., LeCoq, D., Steinmetz, M., Ferrari, E., and Hoch, J. (1983). J. Bact. 153, 1424-1431: Cloning structural gene sacB, which codes for exoenzyme levansucrase of B. subtilis: Expression of the gene in E. coli.

Gentz, R., Langner, A., Chang, A.C.Y., Cohen, S.N. and Bujard, H. (1981). Proc. Nat. Acad. Sci. USA 78, 4936-4940: Cloning and analysis of strong promoters is made possible by the downstream placement of a RNA termination signal.

Gerdes, K., Bech, F.W., Jørgensen, S.T., Løbner-Olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S. and von Meyenburg, K. (1986). EMBO J. 5, 2023-2029: Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon.

Ghosh, P.K., Reddy, V.B., Swinscoe, J., Choudary, P.V., Lebowitz, P. and Weissman, S.M. (1978). J. Biol. Chem. 253, 3643-3647: The 5′ terminal leader sequence of late 16S mRNA from cells infected with simian virus 40.

Gilkes, N.R., Kilburn, D.G., Miller, R.C. and Warren, R.A.J. (1984). Biotechnology 2, 259-263: A mutant of E. coli that leaks cellulase activity encoded by cloned cellulase genes from Cellulomonas fimi.

Gilman, M.Z., Wiggs, J.L. and Chamberlin, M.J. (1981). Nucl. Acids Res. 9, 5991-6000: Nucleotide sequences of two B. subtilis promoters used by B. subtilis sigma-28 RNA polymerase.

Gilman, M.Z. and Chamberlin, M.J. (1983). Cell 35, 285-293: Development and genetic regulation of B. subtilis genes transcribed by sigma-28 RNA polymerase.

Gilman, M.Z., Glenn, J.S., Singer, V.L. and Chamberlin, M.J. (1984). Gene, 32, 11-20: Isolation of sigma-28-specific promoters from B. subtilis DNA.

Gitt, M.A., Wang, L-F. and Doi, R.H. (1985). J. Biol. Chem. 260, 7178-7185: A strong sequence homology exists between the major RNA polymerase sigma factors of B. subtilis and E. coli.

Glass, R.E., Jones, S.T. and Ishihama, A. (1986). Mol. Gen. Genet. 203, 265-268: Genetic studies on the beta subunit of E. coli RNA polymerase.

Godefroy-Colburn, T., Wolfe, A.D., Dondon, J., Grunberg-Manago, M., Dessen, P. and Pantaloni, D. (1975). J. Mol. Biol. 94, 461-478: Light-scattering studies showing the effect of Initiation Factors on the reversible dissociation of E. coli ribosomes.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. and Stormo, G. (1981). Annu. Rev. Microbiol. 35, 365-403: Translational initiation in prokaryotes.

Goldfarb, D.S., Doi, R.H. and Rodriguez, R.L. (1981). Nature 293, 309-311: Expression of Tn-9-derived chloramphenicol resistance in B. subtilis.

Goldfarb, D.S., Wong, S.L., Kudo, T. and Doi, R.H. (1983). Mol. Gen. Genet. 191, 319-325: A temporally regulated promoter from B. subtilis is transcribed only by an RNA polymerase with a 37,000 dalton sigma factor.

Gould, A.R., May, B.K. and Elliot, W.H. (1975). J. Bact. 122, 34-40: Release of extracellular enzymes from B. amyloliquefaciens.

Gray, G.L., Mainzer, S.E., Rey, M.W., Lamsa, M.H., Kindle, K.L., Carmona, C. and Requadt, C. (1986). J. Bact. 166, 635-643: Structural genes encoding the thermophilic alpha-amylases of B. stearothermophilus and B. licheniformis.

Gray, O. and Chang, S. (1981). J. Bact. 145, 422-428: Molecular cloning and expression of B. licheniformis beta-lactamase gene in E. coli and B. subtilis.

Gribskov, M. and Burgess, R.R. (1986). Nucl. Acids Res. 14, 6745-6763: Sigma factors from E. coli, B. subtilis, ´phage SP01, and ´phage T4 are homologous proteins.

Grosjean, H.. and Fiers, W. (1982). Gene, 18, 199-209: Review: Preferential codon usage in prokaryotic genes: The optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes.

Grossman, A.D., Erickson, J.W. and Gross, C.A. (1984). Cell, 38, 383-390: The htpR gene product of E. coli is a sigma factor for heat-shock promoters.

Gryczan, T.J. and Dubnau, D. (1978). Proc. Nat. Acad. Sci. USA 75, 1428-1432: Construction and properties of chimeric plasmids in B. subtilis.

Gryczan, T.J., Contente, S. and Dubnau, D.A.(1978a). J. Bact. 134, 318-329: Characterization of S. aureus plasmids introduced by transformation into B. subtilis.

Gryczan, T., Contente, S. and Dubnau, D. (1980). Mol. Gen. Genet. 177, 459-467: Molecular cloning of heterologous chromosomal DNA by recombination between a plasmid vector and a homologous resident plasmid in B. subtilis.

Gryczan, T., Shivakumar, A.G. and Dubnau, D.A. (1980a). J. Bact. 141, 246-253: Characterization of chimeric plasmid cloning vehicles in B. subtilis.

Guntelberg, A.V. (1956). Comptes Rendus des travaux du Laboratoire Carlsberg 29, 27-35: A method for the production of the plakalbumin-forming proteinase from B. subtilis.

Guntelberg, A.V. and Ottesen, M. (1956). Compt. Rend. des Trav. du Lab. Carlsberg 29, 36-48: Purification of the proteolytic enzyme from B. subtilis.

Gutterson, N.I. and Koshland, D.E. (1983). Proc. Nat. Acad. Sci. USA 80, 4894-4898: Replacement and amplification of bacterial genes with sequences altered in vitro.

Hageman, J.H. and Carlton, B.C. (1973). J. Bact. 114, 612-617: Effects of mutational loss of specific intracellular proteases on the sporulation of B. subtilis.

Hager, P.N. and Rabinowitz, J.C. (1985a): Translational specificity in B. subtilis. In The Molecular Biology of the Bacilli, Vol. II. Dubnau, D.A., ed.,pp. 1-29.

Hager, P.W. and Rabinowitz, J.C. (1985b). J. Biol. Chem. 260, 15163-15167: Inefficient translation of T7 late mRNA by B. subtilis ribosomes: Implications for species-specific translation.

Hahn, J. and Dubnau, D.A. (1985). J. Bact. 162, 1014-1023: Analysis of plasmid deletional instability in B. subtilis.

Haider, M.Z., Knowles, B.H. and Ellar, D.J. (1986). Eur. J. Biochem. 156, 531-540: Specificity of B. thuringiensis var. colmeri insecticidal delta-endotoxin is determined by differential proteolytic processing of the protoxin by larval gut proteases.

Haldenwang, W.G. and Losick, R. (1979). Nature, 282, 256-260: A modified RNA polymerase transcribes a cloned gene under sporulation control in B. subtilis.

Haldenwang, W.G. and Losick, R.(1980). Proc. Nat. Acad. Sci. USA 77, 7000-7004: Novel RNA polymerase sigma factor from B. subtilis.

Haldenwang, W.G., Banner, C.D.B., Ollington, J.F., Losick, R., Hoch, J.A., O'Connor, M.B. and Sonenshein, A.L. (1980). J. Bact. 142, 90-98: Mapping a cloned gene under sporulation control by insertion of a drug resistance marker into the B. subtilis chromosome.

Haldenwang, W.G., Lang, N. and Losick, R. (1981). Cell 23, 615-624: A sporulation-induced sigma-like regulatory protein from B. subtilis.

Haldenwang, W.G. and Truitt, C.L. (1982). J. Bact. 151, 1624-1626: Peptide maps of regulatory subunits of B. subtilis RNA polymerase.

Hall, F.F., Kunkel, H.O. and Prescott, J.M. (1966). Arch. Biochem. Biophys. 114, 145-153: Multiple proteolytic enzymes of B. licheniformis.

Hall, M.N., Gabay, J., Debarbouille, M. and Schwartz, M. (1982). Nature 295, 616-618: A role for mRNA secondary structure in the control of translation initiation.

Hall, M.N., Gabay, J. and Schwartz, M. (1983). EMBO J. 2, 15-19: Evidence for a coupling of synthesis and export of an outer membrane protein in E. coli.

Hanson, R.S. (1975): Role of small molecules in regulation of gene expression and sporogenesis in bacilli. In Spores VI, Gerhardt, P., Costilow, R.N. and Sadoff, H.L. (eds.) (Am. Soc. for Microbiol., Washington D.C.), pp. 318-326.

Harding, N.E., Cleary, J.M., Smith, D.W., Michon, J.J., Brusilow, W.S.A. and Zyskind, J.W. (1982). J. Bact. 152, 983-993: Chromosomal replication origins (oriC) of Enterobacter aerogenes and Klebsiella pneumoniae are functional in E. coli.

Harris- Warrick, R.M. and Lederberg, J. (1978). J. Bact. 133, 1246-1253: Interspecies transformation in Bacillus: Mechanism of heterologous intergenote transformation.

Harwood, C.R., Williams, D.M. and Lovett, P.S. (1983). Gene, 24, 163-169: Nucleotide sequence of a B. pumilus gene specifying chloramphenicol acetyltransferase.

Hashiguchi, K., Tanimoto, A., Nomura, S., Yamane, K., Yoda, K., Harada, S., Mori, M., Furusato, T., Takatsuki, A., Yamasaki, M. and Tamura, G.(1985). Agric. Biol. Chem. 49, 545-550: Gene amplification of the amyE-tmrB region in B. subtilis.

Hashiguchi, K., Tanimoto, A., Nomura, S., Yamane, K., Yoda, K., Harada, S., Mori, M., Furusato, T., Takatsuki, A., Yamasaki, M. and Tamura, G.(1986). Mol. Gen. Genet. 204, 36-43: Amplification of the amyE-tmrB region on the chromosome in tunicamycin-resistant cells of B. subtilis.

Hayashi, S. and Wu, H.C. (1983). J. Bact. 156, 773-777: Biosynthesis of B. licheniformis penicillinase in E. coli and B. subtilis.

Hayashi, S., Chang, S-Y., Chang, S. and Wu, H.C. (1986). J.

Bact. 165, 678-681: Processing of B. licheniformis penicillinase lacking a lipoprotein modification site in E. coli.

Henner, D.J. and Hoch, J.A. (1980). Microbiol. Rev. 44, 57-82: The B. subtilis chromosome.

Hentschel, C., Irminger, J-C., Bucher, P. and Birnstiel, M.L. (1980). Nature 285, 147-151: Sea urchin histone mRNA termini are located in gene regions downstream from putative regulatory sequences.

Herrnstadt, C., Soares, G.C., Wilcox, E.R. and Edwards, D.L. (1986). Biotechnology 4, 305-308: A new strain of B. thuringiensis with activity against coleopteran insects.

Higerd, T.B., Hoch, J.A. and Spizizen, J. (1972). J. Bact. 112, 1026-1028: Hyperprotease-producing mutants of B. subtilis.

Higerd, T.B. and Spizizen, J. (1973). J. Bact. 114, 1184-1192: Isolation of two acetyl esterases from extracts of B. subtilis.

Hirata, H., Fukazawa, T., Negoro, S. and Okada, H. (1986). J. Bact. 166, 722-727: Structure of a beta-galactosidase gene of B. stearothermophilis.

Hiroishi, S. and Kadota, H. (1976). Agric. Biol. Chem. 40, 1047-1049: Intracellular proteases of B. subtilis.

Hitotsuyanagi, K., Yamane, Y. and Maruo, B. (1978). Agric. Biol. Chem. 42, 585-592: Independent control of the production and localization of alkaline phosphatase from the production of alpha-amylase and protease in B. subtilis.

Hoch, J.A. (1976). Genetics of Bacterial Sporulation. In Advances in Genetics, Vol. 18 (Caspar, E.W. ed.) (Academic Press), pp. 69-98.

Hoch, J.A. and Mathews, J. (1973). Genetics 73, 215-228: Chromosomal location of pleiotropic negative sporulation mutations in B. subtilis.

Hoch, J.A., Shiflett, M.A., Trowsdale, J. and Chen, S.M.H. (1978). In Spores VII, Chambliss, G. and Vary, J.C. (eds.) (Am. Soc. for Microbiol, Washington D.C.), pp. 127-130.

Hoch, J.A., Trach, K., Kawamura, F. and Saito, H. (1985). J. Bact. 161, 552-555: Identification of the transcriptional suppressor sof-1 as an alteration inthe spoOA protein.

Hogan, A. and Faust, E.A. (1984). Molecular and Cellular Biol. 4, 2239-2242: Short direct repeats mediate spontaneous high frequency deletions in DNA of minute virus of mice.

Holmes, W.M., Platt, T. and Rosenberg, M. (1983). Cell 32, 1029-1032: Termination of transcription in E. coli.

Honjo, M., Akaoka, A., Nakayama, A., Shimoda, H. and Furutani, Y. (1985). J. Biotechnol. 3, 73-84: Construction of a secretion vector containing the prepro- structure coding region of B. amyloliquefaciens neutral protease gene and secretion of B. subtilis alpha-amylase and human inteferon-beta in B. subtilis.

Honjo, M., Akaoka, A., Nakayama, A. and Furatani, Y. (1986). J. Biotchnol. 4, 63-71: Secretion of human growth hormone in B. subtilis usng propropeptide coding region of B. amyloliquefaciens neutral protease gene.

Hopp, T.P. and Woods, K.R. (1981). Proc. Nat. Acad. Sci. USA 78, 3824-3828: Prediction of protein antigenic determinants from amino acid sequences.

Horinouchi, S. and Weisblum, B. (1982). J. Bact. 150, 815-825: Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance.

Horiuchi, S., Tai, P.C. and Davis, B.D. (1983a). Proc. Nat. Acad. Sci. USA 80, 3287-3291: A 64-kd membrane protein from B. subtilis covered by secreting ribosomes.

Horiuchi, S., Marty-Mazar, D., Tai, P.C. and Davis, B.D. (1983b). J. Bact. 154, 1215-1221: Localization and quantification of proteins characteristic of the complexed membrane of B. subtilis.

Holm. L. (1986). Nucl. Acids Res. 14, 3075-3087: Codon usage and gene expression.

Hu, J.C. and Gross, C.A. (985). Mol. Gen. Genet. 199, 7-13: Mutations in the sigma subunit of E. coli RNA polymerase which affect positive control of transcription.

Hu, M. C-T. and Davidson, N. (1986). Gene 42, 21-29: Mapping transcriptional start points on cloned genomic DNA with T4 DNA polymerase: A precise and convenient technique.

Hui, A., Hayflick, J., Dinkelspiel, K. and De Boer, H.A. (1984). EMBO J. 3, 623-629: Mutagenesis of the three bases preceding the start codon of the beta-galactosidase mRNA and its effect on translation in E. coli.

Hulett, F.M., Wang, P-Z., Sussman, M. and Lee, J-W. K. (1985). Proc. Nat. Acad. Sci USA 82, 1035-1039: Two alkaline phosphatase genes positioned in tandem in B. licheniformis MC14 require different RNA polymerase holoenzymes for transcription.

Ide, M. (1971). Arch. Biochem. Biophys. 144, 262-268: Adenyl cyclase of bacteria.

Iglesias, A., Bensi, G., Canosi, U. and Trautner, T.A. (1981). Mol. Gen. Genet. 184, 405-409: Plasmid transformation in B. subtilis:Alterations introduced into the recipient - homologous DNA of hybrid plasmids can be corrected in transformation.

Iglesias, A. and Trautner, T.A. (1983). Mol. Gen. Genet. 189, 73-76: Plasmid transformation in B. subtilis: Symmetry of gene conversion in transformation with a hybrid plasmid containing chromosomal DNA.

Ikemura, T. (1981). J. Mol. Biol. 146, 1-21: Correlation between the abundance of E. coli transfer RNAs and the occurrence of the respective codons in its protein genes.

Ikeuchi, T., Kudoh, J. and Tsunasawa, S. (1986). Mol. Gen. Genet. 203, 371-376: Amino-terminal structure of spoOA protein and sequence homology with spoOF and spoOB proteins.

Imanaka, T., Tanaka, T., Tsunekawa, H. and Aiba, S. (1981). J. Bact. 147, 776-786: Cloning of the genes for penicillinase, penP and penI, of B. licheniformis in some vector plasmids and their expression in E. coli, B. subtilis, and B. licheniformis.

Inouye, M. and Halegoua, S. (1980). CRC Crit. Rev. in Biochem. 7, 339-371: Secretion and membrane localization of proteins in E. coli.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K. and Inouye, M. (1982). Proc. Nat. Acad. Sci. USA §9, 3438-3441: Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane.

Iserentant, D. and Fiers, W. (1980). Gene 9, 1-12: Secondary structure of mRNA and efficiency of translation initiation.

Isono, S. and Isono, K. (1975). Eur. J. Biochem. 56, 15-22: Role of ribosomal protein S1 in protein synthesis: Effects of its addition to B. stearothermophilus cell-free system.

Ito, J. and Spizizen, J. (1972): Early-blocked asporogenous mutants of B. subtilis 168. In Spores V, Halvorson, H.O., Hanson, R. and Campbell, L.L. (Eds.) (American Soc. for Microbiol.), pp. 107-112.

Ito, J. and Spizizen, J.(1973). Colloq. Int. CNRS 227, 81-82: Genetic studies of catabolite repression insensitive sporulation mutants of B. subtilis.

Ito, K., Bassford, P.J. Jr. and Beckwith, J. (1981). Cell 24, 707-717: Protein localization in E. coli: Is there a common step in the secretion of periplasmic and outer-membrane proteins?

Jacobs, M., Eliasson, M., Uhlen, M. and Flock, J-I. (1985). Nucl. Acids Res. 13, 8913-8926: Cloning, sequencing and expression of subtilisin Carlsberg from B. licheniformis.

Jaehning, J.A., Wiggs, J.L. and Chamberlain, M.J. (1979). Proc. Nat. Acad. Sci. USA 76, 5470-5474: Altered promoter selection by a novel form of B. subtilis RNA polymerase.

Janniere, L., Niaudet, B., Pierre, E. and Ehrlich, S.D. (1985).Gene 40, 47-55: Stable gene amplification in the chromosome of B. subtilis.

Jasin, M. and Schimmel, P. (1984). J. Bact. 159, 783-786: Deletion of an essential gene in E. coli by site-specific recombination with linear DNA fragments.

Johnson, W.C., Moran, C.P. Jr. and Losick, R. (1983). Nature 302, 800-804: Two RNA polymerase sigma factors from B. subtilis discriminate between overlapping promoters for a developmentally regulated gene.

Johnsrud L.(1979). Mol. Gen. Genet. 169,213-218: DNA sequence of the transposable element IS1.

Jolliffe, L.K., Doyle, R.J. and Streips, M.N.(1980). J. Bact. 141, 1199-1208: Extracellular proteases modify cell wall turnover in B. subtilis.

Joyet, P., Guerineau, M. and Heslot, H. (1984). FEMS Microbiol. Lett. 21, 353-358: Cloning of a thermostable alpha-amylase gene from B. licheniformis and its expression in B. subtilis.

Kaiser, E.T., Lawrence, D.S. and Rokita, S.E. (1985). Ann. Rev. Biochem. 54, 565-595: The chemical modification of enzymatic specificity.

Kallio, P. (1986). Eur. J. Biochem. 158, 491-495: The effect of the inverted repeat structure on the production of the cloned B. amyloliquefaciens alpha-amylase.

Kalnins, A., Otto, K., Ruther, U. and Muller-Hill, B. (1983). EMBO J. 2, 593-597: Sequence of the lacZ gene of E. coli.

Kastelein, R.A., Berkhout, B., Overbeek, G.P. and Van Duin, J. (1983). Gene 23, 245-254: Effect of the sequences upstream from the RBS on the yield of protein from the cloned gene for phage MS2 coat protein.

Kato, C., Kudo, T., Wanatabe, K. and Horikoshi, K. (1983). Eur. J. Appl. Microbiol. Biotechnol. 18, 339-343: Extracellular production of Bacillus penicillinase by E. coli carrying pEAP2.

Kato, C., Kudo, T. and Horikoshi, K. (1984). Agric. Biol. Chem. 48, 397-401: Gene expression and production of Bacillus no. 170 penicillinase in E. coli and B. subtilis.

Kawamura, F. and Doi, R.H. (1984). J. Bact. 160, 442-444: Construction of a B. subtilis double mutant deficient in extracellular alkaline and neutral proteases.

Kawamura, F. and Saito, H. (1983). Mol. Gen. Genet. 192, 330-334: Isolation and mapping of a new suppressor mutation of an early sporulation gene spo0F mutation in B. subtilis.

Kawamura, F., Wang, L-F. and Doi, R.H.(1985). Proc. Nat. Acad. Sci. USA 82, 8124-8128: Catabolite-resistant sporulation (crsA) mutations in the B. subtilis RNA polymerase sigma-43 gene (rpoD) can suppress and be suppressed by mutations in spo0 genes.

Keay, L. (1969). Biochem. Biophys. Res. Commun. 36, 257-265: Neutral proteases of the genus Bacillus.

Keay, L. and Moser, P.W. (1969). Biochem. Biophys. Res. Commun. 34, 600-604: Differentiation of alkaline proteases from Bacillus species.

Keay, L. and Wildii, B.S. (1970). Biotechnol. Bioeng. 12, 179-212: Proteases of the genus Bacillus. I. Neutral proteases.

Kendall, D.A., Bock, S.C. and Kaiser, E.T. (1986). Nature 321, 706-708: Idealization of the hydrophobic segment of the alkaline phosphatase signal peptide.

Kennedy, M.B. and Lennarz, W.J. (1979). J. Biol. Chem. 254, 1080-1089: Characterization of the extracellular lipase of B. subtilis and its relationship to membrane-bound lipase found in a mutant strain.

Kerjan, P., Keryer, E. and Szulmajster, J. (1979). Eur. J. Biochem. 98, 353-362: Characterization of a thermosensitive sporulation mutant of B. subtilis affected in the structural gene of an intracellular protease.

Kieser, T. (1984). Plasmid 12, 19-36: Factors affecting the isolation of cccDNA from Strep. lividans and E. coli.

Kleckner, N (1981). Ann. Rev. Genet. 15, 341-404: Transposable elements in prokaryotes.

Kobayashi, T., Kato, C., Kudo, T. and Horikoshi, K. (1986). J. Bact. 166, 728-732: Excretion of the penicillinase of an alkalophilic Bacillus sp. through the E. coli outer membrane is caused by insertional activation of the kil gene in plasmid pMB9.

Koide, Y., Nakamura, A., Uozumi, T. and Beppu, T. (1986). J. Bact. 167, 110-116: Cloning and sequencing of the major Intracellular Serine Protease gene of B. subtilis.

Kreft, J., Bernhard, K. and Goebel, W. (1978). Mol. Gen.

Genet. 162, 59-67: Recombinant plasmids capable of replication in B. subtilis and E. coli.

Kreft, J. and Hughes, C. (1981). Curr. Topics in Microbiol. Immunol. 96, 1-17: Cloning vectors derived from plasmids and phage of Bacillus.

Kudo, T., Kato, C. and Horikoshi, K. (1983) 156, 949-951: Excretion of the penicillinase of an alkalophilic Bacillus sp. through the E. coli outer membrane.

Kudo, T., Yoshitake, J., Kato, C., Usami, R. and Horikoshi, K. (1985). J. Bact. 161, 158-163: Cloning of a developmentally regulated element from alkalophilic B. subtilis DNA.

Kumamoto, C.A., Oliver, D.B. and Beckwith, J. (1984). Nature 308, 863-864: Signal sequence mutations disrupt feedback between secretion of an exported protein and its synthesis in E. coli.

Kumamoto, C. and Beckwith, J. (1985). J. Bact. 163, 267-274: Evidence for specificity at an early step in protein export in E. coli.

Kunst, F., Pascal, M., Lepesant-Kejzlarova, J., Billault, A and Dedonder, R. (1974). Biochimie 56, 1481-1489: Pleiotropic mutations affecting sporulation conditions and the synthesis of extracellular enzymes in B. subtilis 168.

Kurihara, M., Markland, F.S. and Smith, E.L. (1972). J. Biol. Chem. 247, 5619-563: Subtilisin Amylosacchariticus III. Isolation and sequence of the chymotryptic peptides and the complete amino acid sequence.

Kurotsu, T., Marahiel, M.A., Muller, K-D. and Kleinkauf, H. (1982). J. Bact. 151, 1466-1472: Characterization of an intracellular serine protease from sporulating cells of B. subtilis.

Kurzchalia, T.V., Wiedmann, M., Girschovich, A.S., Bochkareva, E.S., Bielka, H. and Rapoport, T.A. (1986). Nature 320, 634-636: The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle.

Kyte, J. and Doolittle, R.F. (1982). J. Mol. Biol. 157, 105-132: A simple method for displaying the hydropathic character of a protein.

Lamond, A.I. and Travers, A.A. (1985). Cell 40, 319-326: Genetically separable functional elements mediate the optimal expression and stringent regulation of a bacterial tRNA gene.

Lampe, M.F. and Bott, K.F. (1984). Nucl. Acids Res. 12, 6307-6323: Cloning the gyrA gene of B. subtilis.

Landick, R., Vaughn, V., Lau, E.T., VanBogelen, R.A.,

Erickson, J.W. and Neidhardt, F.C. (1984). Cell 38, 175-182: Nucleotide sequence of the heat shock regulatory gene of E. coli suggests its protein product may be a transcription factor.

Lauffer, L., Garcia, P.D., Harkins, R.N., Coussens, L., Ullrich, A. and Walker, P. (1985). Nature 318, 334-338: Topology of signal recognition particle receptor in endoplasmic reticulum membrane.

Lecadet, M-M., Lescourret, M. and Klier, A. (1977). Eur. J. Biochem. 79, 329-338: Characterization of an intracellular protease isolated from B. thuringiensis sporulating cells and able to modify homologous RNA polymerase.

Lee, C.A. and Beckwith, J. (1986). J. Bact. 166, 878-883: Suppression of growth and protein secretion defects in E. coli secA mutants by decreasing protein synthesis.

Lee, C.A. and Saier, M.H. Jr. (1983). J. Bact. 153, 685-692: Use of cloned mtl genes of E. coli to introduce mtl deletion mutations into the chromosome.

Lee, G. and Pero, J. (1981). J. Mol. Biol. 152, 247-265: Conserved nucleotide sequences in temporally controlled bacteriophage promoters.

Lee, G., Talkington, C. and Pero, J. (1980). Mol. Gen. Genet. 180, 57-65: Nucleotide sequence of a promoter recognized by B. subtilis RNA polymerase.

LeGrice, S.F.J. and Sonenshein, A.L. (1982). J. Mol. Biol. 162, 551-564: Interaction of B. subtilis RNA polymerase with a chromosomal promoter.

Lehrach, H., Diamond, D., Wozney, J.M. and Boedtker, H. (1977). Biochemistry 16, 4743-4751: RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical re-examination.

Lehtovaara, P., Ulmanen, I. and Palva, I. (1984). Gene 30, 11-16: In vivo transcription initiation and termination sites of an alpha-amylase gene from B. amyloliquefaciens cloned in B. subtilis.

Leighton, T.J. and Doi, R.H. (1971). J. Biol. Chem. 246, 3189-3195: The stability of messenger ribonucleic acid during sporulation in B. subtilis.

Leighton, T.J., Doi, R.H., Warren, R.A.J. and Kelln, R.A. (1973). J. Mol. Biol. 76, 103-122: The relationship of serine protease activity to RNA polymerase modification and sporulation in B. subtilis.

Leonard, C.G. and Mattheis, M.J. (1965). J. Bact. 90, 558-559: Different transforming characteristics of colonial variants from auxotrophic mutants of B. licheniformis.

Leung, A., Rubinstein, S., Yang, C., Jing-Wei, L. and Leighton, T. (1985). Mol. Gen. Genet. 201, 96-98: Suppression of defective sporulation phenotypes by mutation in the major sigma factor gene (rpoD) of B. subtilis.

Levinsohn, S. and Aronson, A. (1967). J. Bact. 93, 1023-1030: Regulation of extracellular protease production in B. cereus.

Levy, P.C., Pangburn, M.K., Burnstein, Y., Ericson, L.H., Neurath, H. and Walsh, K.A. (1975). Proc. Nat. Acad. Sci. USA 72, 4341-4345: Evidence of homologous relationship between thermolysin and neutral protease A of B. subtilis.

Lewin, B. (1974). In Gene expression, Vol. I: Bacterial genomes (John Wiley & Sons), p.80.

Li, Y. and Yousten, A.A. (1976). Appl. Microbiol. 30, 354-361: Metalloprotease from B. thuringiensis.

Lilley, D. (1986). Nature 320, 14-15: Bacterial chromatin: A new twist to an old story.

Lin, G-K., Goldfarb, D.S., Doi, R.H. and Rodrigues, R.L. (1985). Proc. Nat. Acad. Sci. USA 82, 173-177: Mutations that affect the translation efficiency of Tn9- derived cat gene in B. subtilis.

Linn, T.G., Greenleaf, A.L., Shorenstein, R.G. and Losick, R. (1973). Proc. Nat. Acad. Sci. USA 70, 1865-1869: Loss of the sigma activity of RNA polymerase of B. subtilis during sporulation.

Linn, T.G. and Losick, R. (1976). Cell 8, 103-104: The program of protein synthesis during sporulation in B. subtilis.

Little, J.W. (1979). Mol. Gen. Genet. 177, 13-22: Construction and characterization of a plasmid coding for a fragment of the E. coli recA protein.

Loenen, W.A.M. and Brammar, W.J. (1980).
Gene 20, 249-259: A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes.

Lofdahl, S., Guss, B., Uhlen, M., Philipson, L. and Lindberg, M. (1983). Proc. Nat. Acad. Sci. USA 80, 697-670: Gene for staphylococcal protein A.

Lopez, P., Espinosa, M., Greenberg, B. and Sacks, S.A. (1984). Proc. Nat. Acad. Sci. USA 81, 5189-5193: Generation of deletions in pneumococcal mal genes cloned in B. subtilis.

Loshon, C.A., Swerdlow, B.M. and Setlow, P. (1982). J. Biol. Chem. 257, 10838-10845: B. megaterium Spore Protease: synthesis and processing of precursor forms during sporulation and germination.

Losick, R. (1982): Sporulation genes and their regulation. In The Molecular Biology of the Bacilli, Vol. I. B. subtilis (Dubnau, D.A. ed.)(Academic Press) , pp. 179-201.

Losick, R. and Pero, J. (1981). Cell 25, 582-584: Cascades of sigma factors.

Losick, R., Shorenstein, R.G. and Sonenshein, A.L. (1970). Nature 227, 910-913: Structural alteration of RNA polymerase during sporulation.

Losick, R. and Sonenshein, A.L. (1969). Nature 224, 35-37: Change in the template specificity of RNA polymerase during sporulation of B. subtilis.

Maas, R. (1983). Plasmid 10, 296-298: An improved colony hybridization method with significantly increased sensitivity for detection of single genes.

Mahillon, J., Seurinck, J., Van Rompuy, L., Delcour, J. and Zabeau, M. (1985). EMBO J. 4, 3895-3899: Nucleotide sequence and structural organization of an insertion sequence element (IS231) from B. thuringiensis strain berlinger 1715.

Maizel, J.V. Jr. and Lenk, R.P. (1981). Proc. Nat. Acad. Sci. USA 78, 7665-7669: Enhanced graphic matrix analysis of nucleic acid sequences.

Maloy, S.R. and Nunn, W.D. (1981). J. Bact. 145, 1110-1112: Selection for loss of tetracycline resistance by E. coli.

Mamas, M. and Millet, J. (1975). Biochimie 57, 9-16: Purification et proprietes d'une esterase excretee pendant la sporulation de B. subtilis.

Mandelstam, J. and Waites, W.M. (1968). Biochem. J. 109, 793-801: Sporulation in B. subtilis: the role of exoprotease.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982): A Laboratory Manual (Cold Spring Harbor Laboratory).

Mantsala, P. and Zalkin, H. (1980) J. Bact. 141, 493-501: Extracellular and membrane-bound proteases from B. subtilis.

Markaryan, A.N., Ostoslavskaya, V.I., Svyadas, K., Yakusheva, D., Lyublinskaya, L.A., Stongin, A. Ya. and Stepanov, V.M. (1981). Int. J. Biochem. 13, 201-206: The specificity of the B. amyloliquefaciens intracellular serine protease: a comparison with the specificity of secretory subtilisins.

Markland, F.S., Shaw, E. and Smith, E.L. (1968). Proc. nat. Acad. Sci. USA 61, 1440-1447: Identification of Histidine 64 in the active site of subtilisin.

Markland, F.S. and Smith, E.L. (1967). J. Biol. Chem. 242, 5198-5211: Subtilisin BPN' VII: Isolation of cyanogen bromide peptides and the complete amino acid sequence.

Markland, F.S. and Smith, E.L. (1971). In The Enzymes, Vol. III (Boyer, P.D. ed.) (Academic Press),pp. 561-608.

Marmur, J., Seaman, E. and Levine, J. (1963). J. Bact. 85, 461-467: Interspecific transformation in Bacillus.

Matsubara, H. (1970): Purification and assay of thermolysin. In Meth. in Enz. Vol. XIX (Perlman, G.E. and Lorand, L. eds.) (Academic Press), pp. 642-651.

Matsubara, H. and Feder, J. (1971). Other bacterial, mold and yeast proteases. In The Enzymes, Vol. III. (Boyer, P.D., ed.) (Academic Press), pp. 721-795.

Martin, F. and Tinoco, I. (1980). Nucl. Acids Res. 8, 2295-2300: DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription.

Maxam, A.M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. USA 74, 560-564: A new method for sequencing DNA.

McConn, J.D., Tsuru, D. and Yasunobu, K.T. (1964). J. Biol. Chem. 239, 3706-3715: B. subtilis neutral proteinase I. A zinc enzyme of high specific activity.

McKenzie, T., Hoshino, T.,Tanaka, T. and Sueoka, N. (1986). Plasmid 15, 93-103: The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation.

McLauglin, J.R., Murray, C.L. and Rabinowitz, J.C. (1981a). J. Biol. Chem. 256, 11282-11291: Unique features in the RBS sequences of the Gram-positive S. aureus beta-lactamase gene.

McLaughlin, J.R., Murray, C.L. and Rabinowitz, J.C. (1981b). Proc. Nat. Acad. Sci. USA 78, 4912-4916: Initiation-factor independent translation of mRNAs from Gram-positive bacteria.

McLaughlin, J.R., Chang, S-Y. and Chang, S. (1982). Nucl. Acids Res. 10, 3905-3919: Transcriptional analyses of the B. licheniformis penP gene.

McPhalen, C.A., Schnebli, H.P. and James, M.N.G. (1985a). FEBS lett. 188, 55-58: Crystal and molecular structure of the inhibitor eglin from leeches in complex with subtilisin Carlsberg.

McPhalen, C.A., Svendsen, I., Jonassen, I. and James, M.N.G. (1985b). Proc. Nat. Acad, Sci. USA 82, 7242-7246: Crystal and molecular structure of chymotrypsin inhibitor 2 from barley seeds in complex with subtilisin Novo.

Meloun, B., Baudys, M., Kostka, V., Hausdorf, G., Frommel, C. and Hohne, W.E. (1985). FEBS lett. 183, 195-200: Complete primary structure of thermitase from Thermoactinomyces

vulgaris and its structural features related to the subtilisin-type proteinases.

Messing, J., Crea, R. and Seeburg, P.H. (1981). Nucl. Acids Res. 9, 309-321: A system for shotgun DNA sequencing.

Meyer, D.I., Krause, E. and Dobberstein, B. (1982). Nature 297, 647-650: Secretory protein translocation across membranes - the role of the "docking protein".

Michaelis, S., Hunt, J.F. and Beckwith, J. (1986). J. Bact. 167, 160-167: Effects of signal sequence mutations on the kinetics of alkaline phosphatase export to the periplasm in E. coli.

Michel, B., Niaudet, B. and Ehrlich, S.D. (1982). EMBO J. 1, 1565-1571: Intramolecular recombination during plasmid transformation of B. subtilis competent cells.

Michel, J.F. and Cami, B. (1969). Ann. Inst. Pasteur 116, 3-18: Selection de mutants de B. subtilis bloques au debut de la sporulation- nature des mutations selectionees.

Michel, J.F. and Millet, J. (1970). J. Appl. Bacteriol. 33, 220-227: Physiological studies on early-blocked sporulation mutants of B. subtilis.

Milhaud, P., Balassa, G. and Zucca, J. (1978). Mol. Gen. Genet. 163, 35-44: Spore control (sco) mutations in B. subtilis.

Miller, J.H. (1972): Expriments in Molecular Genetics (Cold Spring Harbor Laboratory).

Millet, J. (1969). Bull. Soc. Chim. Biol. 51, 457-469: Characterization de deux endopeptidases excretees par B. subtilis Marburg au cours de la sporulation.

Millet, J. (1970). J. Appl. Bacteriol. 23, 207-219: Characterisation of proteinases excreted by B. subtilis Marburg strain during sporulation.

Millet, J. (1971). C.R. Acad. Sci. Paris 272, 1806-1809: Characterisation d'une endopeptidase cytoplasmique chez B. megaterium en voie de sporulation.

Millet, J. (1977). FEBS lett. 74, 59-61: Characterization of protein inhibitor of intracellular protease from B. subtilis.

Millet, J., Archer, R. and Aubert, J-P. (1969). Biotechnol. Bioeng. 11, 1233-1246: Biochemical and physilogical properties of an extracellular protease produced by B. megaterium.

Millet, J. and Aubert, J.P. (1969). Ann. Inst. Pasteur 117, 461-473: Etude de la megateriopeptidase, protease exocellulaire de B. megaterium III: Biosynthese et role physiologique.

Millet, J., Kerjan, P., Aubert, J-P. and Szulmajster, J. (1972). FEBS lett. 23, 47-50: Proteolytic conversion in vitro of B. subtilis vegetative RNA polymerase into the homologous spore enzyme.

Millet, J., Larribe, M. and Aubert, J-P. (1976). Biochimie 58, 109-117: Mutant thermosensible de B. subtilis affecte dans la sporulation et la serylprotease extracellulaire.

Minkley, E.G. and Pribnow, D. (1973). J. Mol. Biol. 77, 255-277: Transcription of the early region of bacteriophage T7: selective initiation with dinucleotides.

Mitsui, Y., Satow, Y., Sakamaki, T. and Itaka, Y. (1977). J. Biochem. 82, 295-298: Crystal structure of a protein proteinase inhibitor, Streptomyces subtilisin inhibitor, at 2.3Å resolution.

Mongkolsuk, S., Chiang, Y-W., Reynolds, R.B. and Lovett, P.S. (1983). J. Bact. 155, 1399-1406: Restriction fragments that promote activity during post-exponential growth of B. subtilis.

Mongkolsuk, S. and Lovett, P.S. (1984). Proc. Nat. Acad. Sci. USA 81, 3457-3460: Selective expression of a plasmid cat gene at a late stage of B. subtilis sporulation.

Monroe, R.E. (1961). Biochem. J. 81, 225-232: Protein turnover and the formation of protein inclusions during sporulation of B. thuringiensis.

Montville, T.J. (1983). Appl. and Environ. Microbiol. 45, 200-204: Dual substrate plate diffusion assay for proteases.

Moran, C.P. Jr., Lang, N. and Losick, R. (1981a). Nucl. Acids Res. 9, 5979-5990: Nucleotide sequence of a B. subtilis promoter recognized by a B. subtilis RNA polymerase containing sigma-37.

Moran, C.P. Jr., Lang, N., Banner, C.D.B., Haldenwang, W.G. and Losick, R.(1981b). Cell 25, 783-791: Promoter for a developmentally regulated gene in B. subtilis.

Moran. C.P. Jr., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J. and Losick, R. (1982a). Mol. Gen. Genet. 186, 339-346: Nucleotide sequences that signal the initiation of transcription and translation in B. subtilis.

Moran, C.P. Jr., Johnson, W.C. and Losick, R. (1982b). J. Mol. Biol. 162, 709-713: Close contacts between sigma-37 RNA polymerase and a B. subtilis chromosomal promoter.

Mori, M., Tanimoto, A., Yoda, K., Harada, S., Koyama, N., Hashiguchi, K., Obinata, M., Yamasaki, M. and Tamura, G. (1986). J. Bact. 166, 787-794: Essential structure in the

cloned transforming DNA that induces gene amplification of the B. subtilis amyE-tmrB region.

Morihara, K. (1974). Adv. in Enzymology. 41, 178-243: Comparative specificity of microbial proteases.

Mosbach, K., Birnbaum, S., Hardy, K., Davies, J. and Bulow, L. (1983). Nature 302, 543-545: Formation of proinsulin by immobilized B. subtilis.

Mottes, M., Grandi, E., Sgaramella, V., Canosi, U., Morelli, G. and Trautner, T.A. (1979). Mol. Gen. Genet. 174, 281-286: Different specific activities of the monomeric and oligomeric forms of plasmid DNA in transformation of B. subtilis and E. coli.

Muller, M., Ibrahimi, I., Chang, C.N., Walter, P. and Blobel, G. (1982). J. Biol. Chem. 257, 11860-11863: A bacterial secretory protein requires signal recognition particle for translocation across mammalian endoplasmic reticulum.

Murray, C.L. and Rabinowitz, J.C. (1982). J. Biol. Chem. 257, 1053-1062: Nucleotide sequences of transcription and translation initiation regions in Bacillus phage Ø29 early genes.

Musso, R.E., Dilauro, R.and Adhya, S. (1977). Cell 12, 847-854: Dual control for transcription of the galactose operon by cyclic AMP and its receptor protein at two interspersed promoters.

Nagahari, K., Kanaya, S., Munakata, K., Aoyagi, Y and Mitushima, S. (1985). EMBO J. 4, 3589-3592: Secretion into the culture medium of a foreign gene product from E. coli: use of the ompF gene for secretion of human beta-endorphin.

Nagai, K. and Thøgersen, H.C. (1984). Nature 309, 810-812: Generation of beta-globin by sequence-specific proteolysis of a hybrid protein produced in E. coli.

Nagami, Y. and Tanaka, T. (1986). J. Bact. 166, 20-28: Molecular cloning and nucleotide sequence of a DNA fragment from B. natto that enhances production of extracellular proteases and levansucrase in B. subtilis.

Nagata, Y., Yamaguchi, K. and Maruo, B. (1974). J. Bact. 119, 425-430: Genetic and biochemical studies on cell-bound alpha-amylase in B. subtilis Marburg.

Nakahama, K., Miyazaki, T. and Kikuchi, M. (1985). Gene 36, 179-182: A promoter for B. subtilis expression vector.

Nakamura, K., Furusato, T., Shiroza, T. and Yamane, K. (1985). Biochem. Biophys. Res. Commun. 128, 601-606: Stable hyper-production of E. coli beta-lactamase by B. subtilis grown on a 0.5M succinate medium using a B. subtilis alpha-amylase secretion vector.

Nedkov, P., Obethur, W. and Braunitzer, G. (1985) Biol. Chem. Hoppe-Seyler 366, 421-430: Determination of the complete amino acid sequence of subtilisin DY and its comparison with the primary structures of the subtilisins BPN', Carlsberg and Amylosaccharitícus.

Neidhardt, F.C., VanBogelen, R.H. and Vaughn, V. (1984). Annu. Rev. Genet. 18, 295-329: The genetics and regulation of heat shock proteins.

Neugebauer, K., Sprengel, R. and Schaller, H. (1981). Nucl. Acids Res. 9, 2577-2588: Penicillinase from B. licheniformis: nucleotide sequence of the gene and implications for the biosynthesis of a secretory protein in a Gram-positive bacterium.

Neurath, H. (1975): Limited proteolysis and zymogen ac tivation.
In Proteases and Biological control (Reich, E., Rifkin, D.B. and Shaw, E. eds.)(Cold Spring Harbor Laboratory),pp.51-64.

Neurath, H. (1984). Science 224, 350-357: Evolution of Proteolytic Enzymes.

Niaudet, B., Janniere, L. and Ehrlich, S.D. (1985). J. Bact. 163, 111-120: Integration of linear, heterologous DNA molecules into the Bacillus chromosome: mechanism and use in induction of predictable rearrangements.

Niaudet, B., Goze, A and Ehrlich, S.D. (1982). Gene 19, 277-284: Insertional mutagenesis in B. subtilis: mechanism and use in gene cloning.

Niaudet, B., Janniere, L. and Ehrlich, S.D. (1984). Mol. Gen. Genet. 197, 46-54: Recombination between repeated DNA sequences occurs more often in plasmids than in the chromosome of B. subtilis.

Nicaud, J-M., Mackman, N. and Holland, I.B. (1986). J. Biotechnol. 3, 255-270: Minireview: Current status of secretion of foreign proteins by microorganisms.

Nicholson, W.L. and Chambliss, G.H. (1985). J. Bact. 161, 875-881: Isolation and characterization of a cis-acting mutation conferring catabolite repression resistance to alpha-amylase synthesis in B. subtilis.

Nicholson, W.L., Chambliss, G.H., Buckbinder, L., Ambulos, N.P. Jr. and Lovett, P.S. (1985). Gene 35, 113-120: Isolation and expression of a constitutive variant of the chloramphenicol-inducible plasmid gene cat-86 under control of the B. subtilis 168 amylase promoter.

Nicholson, W.L. and Chambliss, G.H. (1986). J. Bact. 165, 663-670: Molecular cloning of cis-acting regulatory alleles of the B. subtilis amyR region by using gene conversion transformation.

Nilsson, B., Abrahmsen, L. and Uhlen, M. (1985). EMBO J. 4, 1075-1080: Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors.

Nomura, S., Yamane, K., Sasaki, T., Yamasaki, M., Tamura, G. and Maruo, B. (1978). J. Bact. 136, 818-821: Tunicamycin-resistant mutants and chromosomal locations of mutational sites in B. subtilis.

Ochi,K., Kandala, J.C. and Freese, E. (1981). J. Biol. Chem. 256, 6866-6875: Initiation of B. subtilis sporulation by the stringent response to partial amino acid deprivaion.

Ochi, K., Kandala, J. and Freese, E. (1982). J. Bact. 151, 1062-1065: Evidence that B. subtilis sporulation induced by the stringent response is caused by the decrease in GTP or GDP.

Ogasawara, N. (1985). Gene 40, 145-150: Markedly unbiased codon usage in B. subtilis.

Ohmura, K., Shiroza, T., Nakamura, K., Nakayama, A., Yamane, K., Yoda, K., Yamasaki, M. and Tamura, G. (1984a). J. Biochem. 95, 87-93: A B. subtilis secretion vector system derived from the B. subtilis alpha-amylase promoter and signal sequence region, and secretion of E. coli beta-lactamase by the vector system.

Ohmura, K., Nakamura,K., Yamazaki, H., Shiroza, T., Yamane, K., Jigami, Y., Tanaka, H.,Yoda, K., Yamasaki, M. and Tamura, G. (1984b). Nucl. Acids Res.12, 5307-5319: Length and structural effect of signal peptides derived from B. subtilis alpha-amylase on secretion of E. coli beta-lactamase in B. subtilis cells.

Okada, J., Shimogaki, H., Murata, K. and Kimura, A. (1984). Appl. Microbiol. Biotechnol. 20, 406-412: Cloning of the gene responsible for the extracellular proteolytic activities of B. licheniformis.

Oliver, D. (1985a). Ann. Rev. Microbiol. 39, 615-648: Protein secretion in E. coli.

Oliver, D.B. (1985b). J. Bact. 161, 285-291: Identification of five new essential genes involved in the synthesis of a secreted protein in E. coli.

Ollington, J.F., Haldenwang, W.G., Huynh, T.V. and Losick, R. (1981). J. Bact. 147, 432-442: Developmentally regulated transcription in a cloned segment of the B. subtilis chromosome.

Oostra, B.A., Kok, K., Van Vliet, A.J., Ab, G. and Gruber, M. (1981). Mol. Gen. Genet. 183, 54-58: A mutation in the RNA polymerase beta' subunit causing depressed ribosomal RNA synthesis in E. coli.

Ottesen, M. and Svendsen, I. (1970). The Subtilisins. In Meth. in Enzymol., Vol. XIX (Perlman, G.E. and Lorand, L. eds.) (Academic Press), pp. 199-215.

Paddon, C.J. and Hartley,R.W. (1986). Gene 40, 231-239: Cloning, sequencing and transcription of an inactivated copy of B. amyloliquefaciens extracellular ribonuclease (barnase).

Palva, I. (1982). Gene 19, 81-87: Molecular cloning of alpha-amylase gene from B. amyloliquefaciens and its expression in B. subtilis.

Palva, I., Petterson, R.F., Kalkkinen, N., Lehtovaara, P., Sarvas, M., Soderlund, H., Takkinen, K. and Kaariainen, L. (1981). Gene 15, 43-51: Nucleotide sequence of the promoter and N-terminal signal peptide region of the alpha-amylase gene from B. amyloliquefaciens.

Palva, I., Sarvas, M., Lehtovaara, P., Sibakov, M. and Kaariainen, L. (1982). Proc. Nat. Acad. Sci. USA 79, 5582-5586:
Secretion of E. coli beta-lactamase from B. subtilis by the aid of alpha-amylase signal sequence.

Palva, I., Lehtovaara, P., Kaariainen, L., Sibakov, M., Cantell, K., Schein, C.H., Kashiwagi, K. and Weissman, C. (1983). Gene 22, 229-235: Secretion of interferon by B. subtilis.

Peden, K.W.C. (1983). Gene 22, 277-280: Revised sequence of the tetracycline-resistance gene of pBR322.

Perlman, D. and Halvorson, H.O. (1983). J. Mol. Biol. 167, 391-409: A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides.

Pero, J., Nelson, J. and Fox, T.D. (1975). Proc. Nat. Acad. Sci. USA 72, 1589-1593: Highly asymmetric transcription by RNA polymerase containing phage SP01-induced polypeptides and a new host protein.

Peschke, U., Beuck, V., Bujard, H., Gentz, R. and LeGrice, S. (1985). J. Mol. Biol. 186, 547-555: Efficient utilization of E. coli transcriptional signals in B. subtilis.

Piggott, P.J. and Coote, J.G. (1976). Bacteriol. Rev. 40, 908-962: Genetic aspects of bacterial endospore formation.

Piggott, P.J. and Hoch, J.A. (1985). Microbiol. Rev. 49, 158-179: Revised geneic linkage map of B. subtilis.

Postemky, C.J., Dignam, S.S. and Setlow, P. (1978). J. Bact. 135, 841-850: Isolation and characterization of B. megaterium mutants containing decreased levels of Spore Protease.

Power, S.D.,Adams, R.M. and Wells, J.A. (1986). Proc. Nat.

Acad. Sci. USA 83, 3096-3100: Secretion and autoproteolytic maturation of subtilisin.

Prestidge, L., Gage, V. and Spizizen, J. (1971). J. Bact. 107, 815-823: Protease activities during the course of sporulation in B. subtilis.

Price, C.W. and Doi, R.H. (1985). Mol. Gen. Genet. 201, 88-95: Genetic mapping of rpoD implicates the major sigma factor of B. subtilis RNA polymerase in sporulation initiation.

Priest, F.C. (1977). Bacteriol. Rev. 41, 711-753: Extracellular enzyme synthesis in the genus Bacillus.

Prozorov, A.A., Bashkirov, V.I., Khasanov, F.K., Glumova, F. and Irich, V.Y. (1985). Gene 34, 39-46: Insertion of eukaryotic DNA into the B. subtilis genome by means of a temperature sensitive plasmid vector.

Pugsley, A.P. and Schwartz, M. (1985). FEMS Microbiol. Rev. 32, 3-38: Export and secretion of proteins by bacteria.

Queen, C. and Rosenberg, M. (1981). Cell 25, 241-249: Differential translational efficiency explains discoordinate expression of the galactose operon.

Randall, L.L. and Hardy, S.J.S. (1977). Eur. J. Biochem. 75, 43-53: Synthesis of exported proteins by membrane-bound polysomes from E. coli.

Rappaport, H.P., Riggsby, W.S. and Holden, D.A. (1965). J. Biol. Chem. 240, 78-86: A B. subtilis proteinase.

Rapoport, G., Klier, A., Billault, A., Fargette, F. and Dedonder, R. (1979). Mol. Gen. Genet. 176, 239-245: Construction of a colony bank of E. coli containing hybrid plasmids representative of the B. subtilis 168 genome.

Ray, G. L. and Haldenwang, W.G. (1986). J. Bact. 166, 472-478: Isolation of B. subtilis genes transcribed in vitro and in vivo by a major sporulation-induced , DNA-dependent RNA polymerase.

Redman, C. M., Siekevitz, P. and Palade, G.E. (1966) J. Biol. Chem. 241, 1150-1158: Synthesis and transfer of amylase in pigeon pancreatic microsomes.

Reysset, G. and Millet, J. (1972). Biochem. Biophys. Res. Commun. 49, 328-334: Characterization of an intracellular protease in B. subtilis during sporulation.

Rhaese, H-J., Grade, R. and Dichtelmuller, H. (1976) European J. Biochem. 64, 205-213: Studies on the control of development. Correlation of initiation of differentiation with synthesis of highly phosphorylated nucleotides in B. subtilis.

Rickenberg, H.V. (1974). Annu. Rev. Microbiol. 28, 353-369: Cyclic AMP in prokaryotes.

Te Riele, H., Michel, B. and Ehrlich, S.D. (1986). Proc. Nat. Acad. Sci. USA. 83, 2541-2545: Single-stranded plasmid DNA in B. subtilis and S. aureus.

Roberts, D., Hooper, B.C., McClure, W.R. and Kleckner, N. (1985) Cell, 43, 117-130: IS10 transposition is regulated by DNA adenine methylation.

Roeder, G.S. and Fink, G.R. (1980) Cell 21, 239-249: DNA rearrangement associated with a transposable element in yeast.

Roitsch, C.A. and Hageman, J.H. (1983) J. Bact. 155, 145-152: Bacillopeptidase F: Two forms of a glycoprotein Serine protease from B. subtilis 168.

Rong, S., Rosenkrantz, M.S. and Sonenshein, A.L. (1986) Bact. 165, 771-779: Transcriptional control of the B. subtilis spoIID gene.

Rood, J.I., Sneddon, M.L.C. and Morrison, J.F. (1980). J. Bact. 144, 552-559: Instability in tyrR strains of plasmids carrying the tyrosine operon: Isolation and characterization of plasmid derivatives with insertions or deletions.

Rose, J.K., and Shafferman, A. (1981) Proc. Natl. Acad. Sci. USA 78, 6670-6674: Conditional expression of the vesicular stomatitis virus glycoprotein gene in E. coli.

Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319-353: Regulatory sequences involved in the promotion and termination of RNA transcription.

Rubin, E.M., Wilson, G.A. and Young, F.E. (1980) Gene 10, 227-235: Expression of Thymidilate synthetase activity in B. subtilis upon integration of a cloned gene from E. coli.

Rudolf, H., Koenig-Rauseo, I. and Hinnen, A. (1985) Gene 36, 87-95: One-step gene replacement in yeast by cotransformation.

Ruther, U., Koenen, M., Otto, K. and Muller-Hill, B. (1981) Nucl. Acids. Res. 9, 4087-4098: pUR222, a vector for cloning and rapid chemical sequencing of RNA.

Ryter, A. (1965) Ann. Inst. Pasteur 108, 40-60: Etude morphologique de la sporulation de B. subtilis.

Ryter, A., Schaeffer, P. and Ionesco, H. (1966) Ann. Inst. Pasteur 107, 384-400: Classification cytologique par leur stade de blocage des mutantes de sporulation de B. subtilis Marburg.

Sadaie, Y. and Kada, T. (1985) J. Bact. 163, 648-653: B. subtilis gene involved in cell division, sporulation and exoenzyme secretion.

Sadoff, H.L., Celikkol, E. and Englebrecht, H.L. (1970). Proc. Nat. Acad. Sci. USA 66, 844-849: Conversion of bacterial aldolase from vegetative to spore form by a sporulation-specific protease.

Sanders, R. and McGeoch, D. (1973) Proc. Nat. Acad. Sci. USA 70, 1017-1021: A mutant transcription factor that is activated by 3′:5′-cyclic guanosine monophosphate.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5463-5467: DNA sequencing with chain-terminating inhibitors.

Sanger, F. and Shaw, D.C. (1960) Nature 187, 872-873: Amino acid sequence about the reactive serine of a proteolytic enzyme from B. subtilis.

Saunders, C.W., Schmidt, B.J., Mirot, M.S. Thompson, L.D. and Guyer, M.S. (1984a) J. Bact. 157, 718-726: Use of chromosomal integration in tne establishment and expression of blaz, a S. aureus beta-lactamase gene in B. subtilis.

Saunders, C.W., Banner, C.D.B., Fahnestock, S.R., Lindberg, M. Mirot, M.S., Rhodes, C.S., Rudolph, C.F., Schmidt, B.J., Thompson, L.D., Uhlen, M. and Guyer, M.S. (1984b): Use of chromosomal integration for the establishment of heterologous genes in B. subtilis. In Protein Transport and Secretion (D.L. Oxendes, ed.)., Publ. by: Alan R. Liss, New York.

Schaeffer, P. (1969) Bact. Rev. 33, 48-71: Spores and the production of antibiotics, exoenzymes and exotoxins.

Schaeffer, P., Millet, J. and Aubert, J.P. (1965) Proc. Nat. Acad. Sci. USA 54, 704-711: Catabolic repression of bacterial sporulation.

Schatz, G. (1986) Nature 321, 108-109: Protein translocation: A common mechanism for different membrane systems?

Schechter, Y., Rafaeli-Eshkol, D. and Hershko, A. (1973) Biochem. Biophys. Res. Commun. 54, 1518-1524: Influence of protease inhibitors and energy in intracellular protein breakdown in starving E. coli.

Scherer, S. and Davis, R.W. (1979) Proc. Nat. Acad. Sci. USA 76, 4951-4955: Replacement of chromosome segments with altered DNA sequences constructed in vitro.

Schreier, H.J. and Sonenshein, A.L. (1986) J. Bact. 167, 35-43: Altered regulation of the glnA gene in glutamine synthetase mutants of B. subtilis.

Seki, T., Yoshikawa, H., Kawadmura, F., Kazami, J. and Saito, H. (1986) J. Gen. Appl. Microbiol. 32, 67-72: The construction of single copy promoter-probe vectors for B. subtilis.

Sekiguchi, J., Takada, N., Okada, H. (1975) J. Bact. 121, 688-694: Genes affecting the productivity of alpha-amylase in B. subtilis Marburg.

Selker, E. and Yanofsky, C. (1979) J. Mol. Biol. 130, 135-143: Nucleotide sequlence of the trpC-trpB intercistronic region from Salmonella typhimurium.

Setlow, P. (1973) Biochem. Biophys. Res. Commun. 52, 365-372: Inability to detect cyclic AMP in vegetative or sporulating cells or dormant spores of B. megaterium.

Sharp, P.A. Berk, A.J. Berget, S.M. (1980). Meth. in Enz. 65, 750-768: Transcription maps of adenovirus.

Sharrock, W.J., Gold, B.M. and Rabinowitz J.C. (1979) J. Mol. Biol. 135, 627-638: Protein synthesis in B. subtilis II. Selective translation of natural mRNAs and its possible relation to the species-specific inhibition of protein synthesis by lincomycin and erythromycin.

Sharrock, W.J. and Rabinowitz, J.C. (1979) J. Mol. Biol. 135, 611-626: Protein synthesis in B. subtilis I. Hydrodynamics and in vitro functional properties of ribosomes from B. subtilis W168.

Sharrock, R.A. and Leighton, T. (1982) Mol. Gen. Genet 186, 432-438: Suppression of defective-sporulation phenotypes by the B. subtilis mutation rev4.

Sharrock, R.A., Rubinstein, S. Chan, M., Leighton, T. (1984) Mol. Gen. Genet. 194, 260-264: Intergenic suppression of spo0 phenotypes by the B. subtilis mutation rvtA.

Shaw, W.V. (1975) Meth. in Enz. 43, 737-755: Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria.

Sherratt, D.J. and Collins, J.F. (1973) J. Gen. Microbiol. 76, 217-230: Analysis by transformation of the penicillinase system in B. licheniformis.

Shiba, K., Ito, K. and Yura, T. (1986) J. Bact. 166, 849-856: Suppressors of the secY24 mutation: Identification and characterization of additional ssy genes in E. coli.

Shimatake, H. and Rosenberg, M. (1981) Nature 292 128-132: Purified lambda regulatory protein cII positively activates promoters for lysogenic development.

Shimotsu, H. and Henner, D.J. (1986) Gene 43, 85-94: Construction of a single-copy integration vector and its use in analysis of regulation of the trp operon of B. subtilis.

Shimotsu, H., Kawamura, F., Kobayashi, Y. and Saito, H. (1983) Proc. Natl. Acad. Sci. USA 80, 658-662: Early sporulation gene spo0F: Nucleotide sequence and analysis of gene product.

Shimotsu, H., Kuroda, M.I., Yanofsky, C. and Henner, D.J. (1986) J. Bact. 166, 461-471: Novel form of transcription attenuation regulates expression of the B. subtilis tryptophan operon.

Shine, J. and Dalgarno, L. (1974) Proc. Nat. Acad. Sci. USA 71, 1342-1346: The 3'-terminal sequence of E. coli 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites.

Shinke R., Aoki, K., Nishira, H. and Yuki, S. (1979) J. Ferment. Technol. 57, 53-55: Isolation of rifampin-resistant, asporogenous mutant from B. cereus and its high beta-amylase productivity.

Shoer, R. and Rappaport, H.P. (1972) J. Bact. 109, 575-583: Analysis of a B. subtilis proteinase mutant.

Shorenstein, R.G. and Losick, R. (1973a) J. Bact. 248, 6163-6169: Purification and properties of the sigma subunit of ribonucleic acid polymerase from vegetative B. subtilis.

Shorenstein, R.G. and Losick, R. (1973b) J. Bact. 248, 6170-6173: Comparative size and properties of the sigma subunits of ribonucleic acid polymerase from B. subtilis and E. coli.

Sibakov, (1986a) Ph.D. thesis, University of Helsinki. Application of a Bacillis secretion vector in protein production.

Sibakov, M. (1986b) Eur. J. Biochem. 155, 577-581: High expression of B. licheniformis alpha-amylase with a Bacillus secretion vector.

Sibakov, M., Sarvas, M. and Palva, I. (1983) FEMS Microbiol. Letts. 17, 81-85: Increased secretion of alpha-amylase from B. subtilis caused by multiple copies of alpha-amylase gene from B. amyloliquefaciens is not further increased by genes enhancing the basic level of secretion.

Sibakov, M. and Palva, I. (1984) Eur. J. Biochem. 145, 567-572: Isolation and the 5'-end nucleotide sequence of B. licheniformis alpha-amylase gene.

Siebenlist, U., Simpson, R.B. and Gilbert, W. (1980) Cell 20, 269-281: E. coli RNA polymerase interacts homologously with two different promoters.

Silhavy, T.J., Benson, S.A. and Emr, S.D. (1983) Microbiol. Rev. 47, 313-344: Mechanisms of protein localization.

Simpson, R. (1979) Cell 18, 277-285: The molecular topography of RNA polymerase-promoter interaction.

Smith, E.L., DeLange, R., Evans, W.H., Landon, M. and Markland, F.S. (1968) J. Biol. Chem. 243, 2184-91: Subtilisin Carlsberg.V. The complete sequence; comparison with subtilisin BPN'; evolutionary relationships.

Smith, I., Paress, P. and Pestka, S. (1978). Proc. Nat. Acad. Sci. USA 75, 5993-5997: Thiostrepton-resistant mutants exhibit relaxed synthesis of RNA.

Smith, M.T., Aynsley, C. and Wake, R.G. (1985) Gene 38, 9-17: Cloning and localization of the B. subtilis chromosome replication terminus terC

Sonenshein, A.L. and Losick, R. (1970) Nature 227, 906-909: RNA polymerase mutants blocked in sporulation.

Southern, E.M. (1975) J. Mol. Biol. 98, 503-517: Detection of specific sequences among DNA fragments separated by gel electrophoresis.

Srivastava, O.P. and Aronson, A.I. (1981). Arch. Microbiol. 129, 227-232: Isolation and characterization of a unique protease from sporulating cells of B. subtilis.

Stahl, M.L. and Ferrari, E. (1984) J. Bact. 158, 411-418: Replacement of the B. subtilis subtilisin structural gene with an in vitro-derived deletion mutation.

Stallcup, M.R., Sharrock, W.J. and Rabinowitz, J.C. (1976) J. Biol. Chem. 251, 2499-2510: Specificity of bacterial ribosomes and messenger ribonucleic acids in protein synthesis reactions in vitro.

Stanssens, P., Remaut, E. and Fiers, W. (1985) Gene 36, 211-223: Alterations upstream from the Shine-Dalgarno regions and their effect on bacterial gene expression.

Stassi, D.L., Dunn, J.J. and Lacks, S.A. (1982) Gene 20, 359-366: Nucleotide sequence of DNA controlling expression of genes for maltosaccharide utilization in S. pneumoniae.

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

Stauffer, C.E. and Etson, D. (1969) J. Biol. Chem. 244, 5333-5338: The effect on subtilisin activity of oxidizing a methionine residue.

Steinmetz, M., Kurst, F. and Dedander, R. (1976) Mol. Gen. Genet. 148, 281-285: Mapping of mutations affecting synthesis of exocellular enzymes in B. subtilis. Identity of the sacUh, amyB, and pap mutations.

Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Treboul, E. and Gay, P. (1985) Mol. Gen. Genet. 200, 220-228: The DNA sequence of the gene for the secreted B. subtilis enzyme levansucrase and its genetic control sites.

Steitz, J.A., Wahba, A.J., Laughrea, M., Moore, P.B. 1977) Nuc. Acids. Res. 4, 1-15: Differential requirements for polypeptide chain initiation complex formation at the three bacteriophage R17 intiator regions.

Stepanov, V.M., Strongin, A. Ya., Izotova, L.S., Abramov, Z.T., Lyublinskaya, L.A., Ermakova, L.M. Baratova, L.A. and Belyanova, L.P. (1977) Biochem Biophys. Res. Commun. 77, 298-: Intracellular serine protease from B. subtilis. Structural comparison with extracellular serine proteases - subtilisins.

Stephens, M.A., Lang, M., Sandman, K., and Losick, R. (1984). J. Mol. Biol. 176, 333-348: A promoter whose utilization is temporally regulated during sporulation in B. subtilis.

Stephens, M.A., Ortlepp, S.A., Ollington, J.F., and McConnell, D.J. (1984). J. Bact. 158, 369-372: Nucleotide sequence of the 5′ region of the B. licheniformis alpha-amylase gene: Comparison with the B. amyloliquefaciens gene.

Stewart, G.C. and Bott, K. F. (1983). Nucl. Acids Res. 11, 6289-6300: DNA sequence of the tandem ribosomal promoter for B. subtilis operon rrnB.

Stormo, G.D., Schneider, T.D. and Gold, L.M. (1982). Nucl. Acids Res. 10, 2971-2996: Characterization of translational initiation sites in E. coli.

Stragier, P. (1986). FEBS lett. 195, 9-11: Comment on duplicated sporulation genes in bacteria by Errington et al (1985).

Stragier, P., Bouvier, J., Bonamy, C. and Szulmajster, J. (1984). Nature 312, 376-378: A developmental gene product of B. subtilis homologous to the sigma factor of E. coli.

Strauch, K.L., Kumamoto, C.A., and Beckwith, J. (1986). J. Bact. 166, 505-512: Does SecA mediate coupling between secretion and translation in E. coli.

Strongin, A. Ya., Izotova, L.S., Abramov, Z.T., Gorodetsky, D.I., Ermakova, L.M., Baratova, L.A., Belyanova, L.P. and Stepanov, V.M. (1978a). J. Bact. 133, 1401-1411: Intracellular Serine Protease of B. subtilis; sequence homology with extracellular subtilisins.

Strongin, A.Ya., Izotova, L.S., Abramov, Z.T., Gorodetsky, D.I. and Stepanov, V.M. (1978b). Mol. Gen. Genet. 159, 337-339: Two related structural genes coding for two homologous serine proteases in the B. subtilis chromosome.

Strongin, A. Ya., Gorodetsky, D.I., Kuznetsova, I.A., Yanonis, V.V., Abramov, Z.T., Belyanova, L.P., Baratova, L.A. and Stepanov, V.M.(1979a). Biochem. J. 179, 333-339: Intracellular Serine Proteinase of B. subtilis strain Marburg 168: Comparison with the homologous enzyme from B. subtilis strain A-50.

Strongin, A. Ya., Abramov, Z.T., Yaroslavtseva, N.G., Baratova, L.A., Shaginyan, K.A., Belyanova, L.P. and Stepanov, V.M. (1979b). J. Bact. 137, 1017-1019: Direct comparison of the subtilisin-like intracellular protease of B. licheniformis with the homologous enzymes of B. subtilis.

Sun, D. and Takahashi, I. (1984). Can. J. Microbiol. 30, 423-429: A catabolite-resistant mutation is localized in the rpo operon of B. subtilis.

Sun, D. and Takahashi, I. (1985). Can. J. Microbiol. 31, 429-435: Suppressor mutations for crs mutants of B. subtilis.

Sutcliffe, J.G. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 77-90: Complete nucleotide sequence of the E. coli plasmid pBR322.

Svendsen, I. (1976). Carlsberg Res. Commun. 41, 238-291: Chemical modification of the subtilisins with special reference to the binding of large substrates. A review.

Svendsen, I., Genov, N. and Idakieva, K. (1986). FEBS lett. 196, 228-232: Complete amino acid sequence of alkaline mesentericopeptidase.

Szekeley, M. (1977). Nature 269, 754-755: Rapid RNA sequencing.

Takahashi, I. (1979). Can J. Microbiol. 25, 1283-1287: Catabolite repression-resistant mutants of B. subtilis.

Takahashi, W., Yamagata, H., Yamaguchi, K., Tsukagoshi, N. and Udaka, S. (1983). J. Bact. 156, 1130-1134: Genetic transformation of B. brevis 47, a protein-secreting bacterium, by plasmid DNA.

Takeda, Y., Ohlendorf, D.H., Anderson, W.F. and Matthews, B. (1983). Science 221, 1020-1026: DNA-binding proteins.

Talkington, C. and Pero, J. (1978). Proc. Nat. Acad. Sci. USA: Promoter recognition by phage SP01-modified RNA polymerase.

Talkington, C. and Pero, J. (1979). Proc. Nat. Acad. Sci. USA 76, 5465-5469: Distinctive nucleotide sequences of promoters recognized by RNA polymerase containing a phage-encoded "sigma-like" protein.

Talmadge, K., Kaufman, J. and Gilbert, W. (1980). Proc. Nat. Acad. Sci. USA 77, 3988-3992: Bacteria mature preproinsulin to proinsulin.

Tatti, K.M., Kenney, T.J., Hay, R.E. and Moran, C.P. Jr. (1985). Gene 36, 151-157: Promoter specificity of a sporulation-induced form of RNA polymerase from B. subtilis.

Tatti, K.M. and Moran, C.P. Jr. (1984). J. Mol. Biol. 175, 285-297: Promoter recognition by sigma-37 RNA polymerase from B. subtilis.

Tatti, K.M. and Moran, C.P. Jr. (1985). Nature 314, 190-192: Utilization of one promoter by two forms of RNA polymerase from B. subtilis.

Taylor, W.E., Straus, D.B., Grossman, A.D., Burton, Z.F., Gross, C.A. and Burgess, R.R. (1984). Cell 38, 371-381: Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of E. coli RNA polymerase.

Thomas, P.G., Russell, A.J. and Fersht, A.R. (1985). Nature 318, 375-376: Tailoring the pH dependence of enzyme catalysis using protein engineering.

Thorne, C.B. and Stull, H.B. (1966). J. Bact. 91, 1012-1020: Factors affecting the transformation of B. licheniformis.

Tinoco, I. Jr., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973): Nature New Biology 246, 40-41: Improved estimation of secondary structure in ribonucleic acids.

Tjian, R. and Pero, J. (1976). Nature 262, 753-757: Bacteriophage SP01 regulatory proteins directing late gene transcription in vitro.

Toma, S., Del Bue, M., Pirola, A. and Grandi, G. (1986). J. Bact. 167, 740-743: nprR1 and nprR2 regulatory regions for neutral protease expression in B. subtilis.

Tomioka, N., Honjo, M., Funahashi, K., Manabe, K., Akaoka, A., Mita, I. and Furutani, Y. (1985). J. Biotechnol. 3, 85-96: Cloning, sequencing, and some properties of a novel B. amyloliquefaciens gene involved in the increase of extracellular protease activities.

Trach, K.A., Chapman, J.W., Piggot, P.J. and Hoch, J.A. (1985): Proc. Nat. Acad. Sci. USA 82, 7260-7264: Deduced product of the stage 0 sporulation gene spo0F shares homology with the Spo0A,OmpR and SfrA proteins.

Travers, A.A. (1969). Nature 223,1107-1110: Bacteriophage sigma factor from RNA polymerase.

Travers, A.A. (1970). Nature 225, 1009-1012: Positive control of transcription by a bacteriophage sigma factor.

Travers, A.A. and Burgess, R.R. (1969). Nature 222, 537-540: Cyclic reuse of the RNA polymerase sigma factor.

Trempy, J.E., Bonamy, C., Szulmajster, J. and Haldenwang, W.G. (1985a). Proc. Nat. Acad. Sci. USA 82, 4189-4192: B. subtilis sigma factor sigma-29 is the product of the sporulation-essential gene spoIIG

Trempy, J.E., Morrison-Plummer, J. and Haldenwang, W.G. (1985b). J. Bact. 161, 340-346: Synthesis of sigma-29, an RNA polymerase specificity determinant, is a developmentally regulated event in B. subtilis.

Trempy, J.E. and Haldenwang, W.G. (1985). J. Bact. 164, 1356-1358: Sigma-29 -like protein is a common sporulation-specific element in bacteria of the genus Bacillus.

Trowsdale, J., Chen, S.M.H. and Hoch, J.A. (1979). Mol. Gen. Genet. 173, 61-70: Genetic analysis of a class of polymyxin resistant partial revertants of stage 0 sporulation mutants of B. subtilis: map of the chromosomal region near the origin of replication.

Truitt, C.L., Luann Ray, G., Trempy, J.E., Da-Jian,Z. and Haldenwang, W.G. (1985). J. Bact. 161, 515-522: Isolation of B. subtilis mutants altered in expression of a gene transcribed in vitro by a minor form of RNA polymerase (E-sigma-37).

Tsuru, D., Kira, H., Yamamoto, T. and Fukumoto, J. (1966a). Agric. Biol. Chem. 30, 856-862: Studies on bacterial protease XIV. Zinc as an essential metal component of neutral protease of B. subtilis var. amylosaccharíticus.

Tsuru, D., Yamamoto, T., and Fukumoto, J. (1966b). Agric. Biol. Chem. (Tokyo) 30, 651-658: Studies on bacterial protease Part XIII. Purification, crystallization and some enzymatic properties of neutral protease of B. subtilis var. amylosaccharíticus.

Tucker, W.T., Miller, C.A. and Cohen, S.N. (1984) Cell 38, 191-201: Structural and functional analysis of the par region of the pSC101 plasmid.

Tullis, R.H. and Rubin, H. (1980). Anal. Biochem. 107, 260-264: Calcium protects DNaseI from Proteinase K: a new method for the removal of contaminating RNase from DNaseI.

Twigg, A.J. and Sherratt, D. (1980). Nature 283, 216-218: Trans-complementable copy-number mutants of plasmid ColE1.

Uehara, H., Yoneda, Y., Yamane, K. and Maruo B. (1974). J. Bact. 119, 82-91: Regulation of neutral protease productivity in B. subtilis: transformation of high productivity.

Uehara, H., Yamane, K. and Maruo, B. (1979). J. Bact. 139, 583-590: Thermosensitive extracellular neutral proteases in B. subtilis: isolation, characterization, and genetics.

Uhlen, M., Flock, J-I., and Philipson, L. (1981). Plasmid 5, 161-169: recE- independent deletions of recombinant plasmids in B. subtilis.

Ulmanen, I., Lundstrom, K., Lehtovaara, P., Sarvas, M., Ruohonen, M. and Palva, I. (1985). J. Bact. 162, 176-182: Transcription and translation of foreign genes in B. subtilis by the aid of a secretion vector.

Vasantha, N. and Freese, E. (1980). J. Bact. 144, 1119-1125: Enzyme changes during B. subtilis sporulation caused by deprivation of guanosine nucleotides.

Vasantha, N. and Thompson, L.D. (1986). J. Bact. 165, 837-842: Secretion of a heterologous protein from B. subtilis with the aid of protease signal sequences.

Vasantha, N., Thompson, L.D., Rhodes, C., Banner, C., Nagle, J. and Filpula, D. (1984). J. Bacteriol. 159, 811-819: Genes for alkaline protease and neutral protease from B. amyloliquefaciens contain a large open reading frame between the regions coding for signal sequence and mature protein.

Verbruggen, R. (1975). Biochem. J. 151, 149-155: Subtilopeptidase A Isoenzyme system: interaction with serum components and its importance for quantitative immunoelectrophoresis.

Vitkovic, L. and Sadoff, H.H. (1977). J. Bact. 131, 891-896: Purification of the extracellular protease of B. licheniformis and its inhibition by bacitracin.

Vlasuk, G.P., Inouye, S., Ito, H., Itakura, K., and Inouye, M. (1983). J. Biol. Chem. 258, 7141-7148:Effects of the complete removal of basic amino acid reidues from the signal peptide on secretion of lipoprotein in E. coli.

Vold, B.S. (1973). J. Bact.113, 825-833: Analysis of isoaccepting transfer RNA species of B. subtilis.

von Heijne, G. (1983). Eur. J. Biochem. 133, 17-21: Patterns of amino acids near signal sequence cleavage sites.

von Heijne, G. (1984a). J. Mol. Biol. 173, 243-251: How signal sequences maintain cleavage specificity.

von Heijne, G. (1984b). EMBO J. 3, 2315-2318: Analysis of the distribution of charged residues in the N-terminal region of signal sequences.

von Heijne, G. (1986).Nucl. Acids Res. 14, 4683-4690: A new method for predicting signal sequence cleavage sites.

de Vos, W.M. and Venema, G. (1981). Mol. Gen. Genet. 182, 39-43: Fate of plasmid DNA in transformation of B. subtilis protoplasts.

de Vos, W.M., Venema, G., Canosi, U. and Trautner, T.A. (1981). Mol. Gen. Genet. 181, 424-433: Plasmid transformation in B. subtilis: Fate of plasmid DNA.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T. and Itakura, K. (1979). Nucl. Acids Res. 6, 3543-3557: Hybridization of synthetic oligonucleotides to ØX174 DNA: the effect of single base pair mismatch.

Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981). Nucl. Acids Res. 9, 879-894: The use of synthetic oligonucleotides as hybridization probes II. Hybridization of oligonucleotides of mixed sequence to rabbit beta-globin DNA.

Walter, P. and Blobel, G. (1981). J. Cell Biol. 91, 557-561: Translocation of proteins across the endoplasmic reticulum III. Signal Recognition zparticle (SRP) causes signal sequence dependent and site-specific arrest of chain elongation that is released by microsomal membranes.

Walter, P. and Blobel, G. (1983). Cell 34, 525-533: Disassembly and reconstitution of signal recognition particle.

Wang, L.F. and Doi, R.H. (1986). Nucl. Acids Res. 14, 4293-4307: Nucleotide sequence and organization of B. subtilis RNA polymerase major sigma (sigma-43) operon.

Wang, P-Z., and Doi, R.H. (1984). J. Biol. Chem. 259, 8619-8625: Overlapping promoters transcribed by B. subtilis sigma-55 and sigma-37 RNA polymerase holoenzymes during growth and stationary phases.

Warburton, N., Boseley, P.G. and Porter, A.G. (1983). Nucl. Acids Res. 11, 5837-5854: Increased expression of a cloned gene by local mutagenesis of its promoter and ribosome-binding site.

Warren, S.C. (1968). Biochem. J. 109, 811-818: Sporulation in B. subtilis: Biochemical changes.

Watson, M.E.E. (1984). Nucl. Acids Res. 12, 5145-5164: Compilation of published signal sequences.

Waye, M.M.Y. and Winter, G. (1986). Eur. J. Biochem. 158, 505-510: A transcription terminator in the 5′ non-coding region of the tyrosyl tRNA synthetase gene from B. stearothermophilus.

Weinrauch, Y. and Dubnau, D. (1983). J. Bact. 154, 1077-1087: Plasmid marker rescue transformation in B. subtilis.

Weir, J., Dubnau, D., Ramakrishna, N. and Smith, I. (1984). J. Bact. 157, 405-412: B. subtilis spo0H gene.

Welker, N.E. and Campbell, L.L. (1967a). J. Bact. 94,

1124-1130: Unrelatedness of B. amyloliquefaciens and B. subtilis.

Welker, N.E. and Campbell, L.L. (1967b). J. Bact. 94, 1131-1135: Comparison of the alpha-amylase of B. subtilis and B. amyloliquefaciens.

Wells, J.A., Ferrari, E., Henner, D.J., Estell, D.A. and Chen, E.Y. (1983). Nucl. Acids Res. 11, 7911-7925: Cloning, sequencing and secretion of B. amyloliquefaciens subtilisin in B. subtilis.

Wells, J.A. and Powers, D.B. (1986). J. Biol. Chem. 261, 6564-6570: In vivo formation and stability of engineered disulfide bonds in subtilisin.

Westpheling, J., Ranes, M. and Losick, R. (1985). Nature 313, 22-27: RNA polymerase heterogeneity in Streptomyces coelicolor.

Wiggs, J.L., Bush, J.W. and Chamberlin, M.J. (1979). Cell 16, 97-109: Utilization of promoter and terminator sites on bacteriophage T7 DNA by RNA polymerase from a variety of bacterial orders.

Wiggs, J.L., Gilman, M.Z. and Chamberlin, M.J. (1981). Proc. Nat. Acad. Sci. USA 78, 2762-2766: Heterogeneity of RNA polymerase in B. subtilis: evidence for an additional sigma factor in vegetative cells.

Wilkinson, A.J., Fersht, A.R., Blow, D.M., Carter, P. and Winter, G. (1984). Nature 307,187-188: A large increase in enzyme-substrate affinity by protein engineering.

Williams, D.C., van Frank, R.M., Muth, W.L. and Burnett, J.P. (1982). Science 215, 687-689: Cytoplasmic inclusion bodies in E. coli producing biosynthetic human insulin proteins.

Wilson, C.R. and Morgan, A.E. (1985). J. Bact. 163, 445-453: Chromosomal-DNA amplification in B. subtilis.

Wong, H.C., Schnepf, H.E. and Whiteley, H.R. (1983). J. Biol. Cchem. 258, 1960-1967: Transcriptional and translational start sites for the B. thuringiensis crystal protein gene.

Wong, S-L. and Doi, R.H. (1982). J. Biol. Chem. 257, 11932-11936: Peptide mapping of B. subtilis RNA polymerase sigma factors and core-associated polypeptides.

Wong, S-L. and Doi, R.H. (1984). J. Biol. Chem. 259, 9762-9767: Utilization of a B. subtilis sigma-37 promoter by E. coli RNA polymerase in vivo.

Wong, S-L-. Price, C.W., Goldfarb, D.S. and Doi, R.H. (1984): Proc. Nat. Acad. Sci. USA 81, 1184-1188: The subtilisin E gene of B. subtilis is transcribed from a sigma-37 promoter in vivo.

Wright, C.S., Alden, R.A. and Kraut, J. (1969). Nature 221, 235-242: Structure of subtilisin BPN' at 2.5 A resolution.

Yamada, H., Yamagata, H. and Mizushima, S. (1984). FEBS lett. 166, 179-182: The major outer membrane lipoprotein and new lipoproteins share a common signal peptidase that exists in the cytoplasmic membrane of E. coli.

Yamaguchi,K., Nagata, Y. and Maruo, B. (1974). J. Bact. 119, 416-424: Isolation of mutants defective in alpha-amylase from B. subtilis: Genetic analyses.

Yamamoto, S. and Lampen, J.O. (1976). Proc. Nat. Acad. Sci. USA 73, 1457-1461: Membrane penicillinase of B. licheniformis 749/C: sequence and possible repeated tetrapeptide structure of the phospholipopeptide region.

Yamane, K. and Maruo, B. (1978). J. Bact. 134, 100-107: Purification and characterization of extracellular soluble and membrane-bound insoluble alkaline phosphatase possessing phosphodiesterase activities in B. subtilis.

Yamazaki, H., Ohmura, K., Nakayama, A., Takeichi, Y., Otozai, K., Yamasaki, M., Tamura, G. and Yamane, K. (1983). J. Bact. 156, 327-337: Alpha-amylase genes (amyR2 and amyE+) from an alpha-amylase-hyperproducing B. subtilis strain: molecular cloning and nucleotide sequence.

Yanagida, N., Uozumi, T. and Beppu, T. (1986). J. Bact. 166, 937-944: Specific excretion of Serratia marcescens protease through the outer membrane of E. coli.

Yang, M., Galizzi, A. and Henner, D. (1983). Nucl. Acids Res. 11, 237-249: Nucleotide sequence of the amylase gene from B. subtilis.

Yang, M.Y., Ferrari, E. and Henner, D. (1984). J. Bact. 160, 15-21: Cloning of the neutral protease gene of B. subtilis and the use of the cloned gene to create an in vitro-derived deletion mutant.

Yang, M., Ferrari, E., Chen, E. and Henner, D.J. (1986). J. Bact. 166, 113-119: Identification of the pleiotropic sacQ gene of B. subtilis.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Gene, 33, 103-119: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and mp19 vectors.

Yansura, D.G. and Henner, D.J. (1984). Proc. Nat. Acad. Sci. USA 81, 439-443: Use of the E. coli lac repressor and operator to control gene expression in B. subtilis.

Yoneda, Y. and Maruo, B. (1975). J. Bact. 124,48-54: Mutation of B. subtilis causing hyperproduction of alpha-amylase and protease, and its synergistic effect.

Yoneda, Y., Yamane, Y. and Maruo, B. (1973). Biochem. Biophys. Res. Commun. 50, 765-770: Membrane mutation related to the production of extracellular alpha-amylase and protease in B. subtilis.

Yoneda, Y., Graham, S. and Young, F.E. (1979). Biochem. Biophys. Res. Commun. 91, 1556-1564: Cloning of a foreign gene coding for alpha-amylase in B. subtilis.

Yoneda, Y., Yamane, K., Yamaguchi, K., Nagata, Y. and Maruo, B. (1974). J. Bact. 120, 1144-1150: Transformation of B. subtilis in alpha-amylase productivity by DNA from B. subtilis var. amylosaccharlticus.

Yoshimura, K., Miyazaki, T., Nakahama, K. and Kikuchi, M. (1986). Appl. Microbiol. Biotechnol. 23, 250-256: B. subtilis secretes a foreign protein by the signal sequence of B. amyloliquefaciens neutral protease.

Young, M. (1984). J. Gen. Microbiol. 130, 1613-1621: Gene amplification in B. subtilis.

Yudkin, M.D. (1986). Mol. Gen. Genet. 202, 55-57: The sigma-like product of sporulation gene spoIIAC of B. subtilis is toxic to E. coli.

Zeigler, D.R. and Dean, D.H. (1985). FEMS Microbiol. Rev. 32, 101-134: Revised genetic map of B. subtilis 168.

Zuber, P. and Losick, R. (1983). Cell 35, 275-283: Use of a lacz fusion to study the role of the spo0 genes of B. subtilis in developmental regulation.