MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

2

JMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A261 021

2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
10/10/92	

4. TITLE AND SUBTITLE
Plasmid stabilization to insure gene expression

5. FUNDING NUMBERS
DAAL03-88-K-0064

6. AUTHOR(S)
Arnold L. Demain

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Dept. of Biology
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 25126.11-LS

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

An unstable (structurally) plasmid in *Bacillus subtilis* was stabilized by modifying the nutritional environment and reducing the size of the plasmid. The stabilization was explained solely by the effect of the growth rate ratio between cells containing modified and parental plasmids. The recombinant protein produced by the *B. subtilis* system was β-galactosidase from *Escherichia coli*. With the success achieved here, attention was shifted to cellulase production in *Clostridium thermocellum* including enzyme isolation and gene cloning.

14. SUBJECT TERMS
Plasmid; Gene expression; *Bacillus subtilis*; *Clostridium thermocellum*; cellulase; β-galactosidase

15. NUMBER OF PAGES
10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

" PLASMID STABILIZATION TO INSURE GENE EXPRESSION "

FINAL REPORT

ARNOLD L. DEMAIN

OCTOBER 10, 1992

U. S. ARMY RESEARCH OFFICE

FUNDING DOCUMENT DAAL03-88-K-0064

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

APPROVED FOR PUBLIC RELEASE;

DISTRIBUTION UNLIMITED.

The main objectives of this study were to quantify the parameters that govern plasmid loss; to explore genetic and nutritional conditions that affect these parameters; to investigate plasmid-host interactions in order to stabilize plasmid-encoded genes in B. subtilis cultures. The B. subtilis-E. coli bifunctional plasmid pCED3 was used as a model of a recombinant plasmid. The principal findings and conclusions are as follow:

1) Plasmid pCED3 is structurally unstable in B. subtilis cultures grown in the presence of kanamycin. Analysis of 96 modified plasmids indicated that deletions in the plasmid occur at many different sites suggesting a deletion mechanism that involves short recognition sequences.

2) The presence of plasmid pCED3 slowed the growth rate of the B. subtilis host. Cells that contain modified plasmids grow faster than the parental cells and take over the cell population.

3) Two different methodologies were developed to reduce the cultural instability of the plasmid-directed lacZ$^+$ phenotype:

a) By growing the cells in media that supported low growth rates, the growth rate ratio between modified and parental cells was reduced resulting in a partial stabilization of the LacZ$^+$ phenotype in the population.

b) Removal of a 4.77 kb EcoRI fragment (that consisted primarily of the pBR322 replicon) from plasmid pCED3 produced a more stable plasmid derivative, designated pYS1. Cells harboring plasmid pYS1 grew faster than pCED3 bearing cells, although the level of activity of ß-galactosidase was comparable in both strains.

By combining the two approaches (i.e. growth of pYS1 bearing cells in medium that supports low growth rate), the LacZ$^+$ phenotype was stably maintained in the cell population for over 170 generations. Under these conditions there was no detectable difference between the growth rates of cells bearing parental and modified plasmids.

4) The stabilization of the LacZ$^+$ phenotype can be explained solely by the effect on the growth rate ratio between cells containing modified and parental plasmids. By using modified stability experiments (where a single cell rather than a suspended colony was used to initiate growth), independent growth rate measurements and a simple mathematical model, the kinetics of the loss of the LacZ$^+$ phenotype was described in terms of the two variables α and p (where α is the ratio of growth rates between modified and parental cells, and p is the probability of obtaining modified cells from parental cells). Under the conditions tested the average values of α were: 1.52 for cultures growing in medium that supports high growth rates, 1.28 for cultures growing in medium that supports low growth rates, and 1.18 for cultures containing the modified plasmid pYS1 grown in medium that supports high growth rates. The calculated p values ranged between 10^{-8} to 10^{-10} under all conditions. These results indicate that although plasmid deletions may be rare events in B. subtilis, plasmid instability in these cultures is readily observed due to the growth advantage of cells bearing the deleted plasmids.

5) Plasmid pYS137 was developed as a tool for direct estimation of deletion rates in B. subtilis and may prove useful in comparing deletion rates in various B. subtilis

strains. Deletions in plasmid pYS137 occur at a rate of 1.0 X 10^{-8} deletions per cell per generation in B. subtilis strain BD393.

6) There are no detectable differences in growth rates between E. coli strains harboring plasmid pCED3 and plasmid-free cells. This fact explains the observed stability of E. coli populations bearing plasmid pCED3, and may provide a general and novel explanation for the observed differences in plasmid instability between B. subtilis and E. coli cultures.

7) The ß-lactamase gene product (pre-Ap) does not mediate the growth inhibition exerted by the pBR322 portion of pCED3.

8) Efficient lacZ directed transcription from the tms promoter resulted in reduction in growth rate and plasmid stability without an increase in the activity of the lacZ gene product. The plasmid copy number varied between 660-1300 with no clear correlation to the strength of the tms promoter. This result indicates that reduction in copy number is not responsible to the growth inhibition exerted by tms directed transcription.

9) Transcription from the tms promoter inhibited the expression of the plasmid-encoded kanamycin resistance gene (KNT) resulting in the reduction of both ß-galactosidase expression and cell growth rate in the presence of kanamycin. This effect was diminished by changing the orientation of the KNT gene on the plasmid. The negative effect of tms-directed transcription on the cell resistance to kanamycin and the high copy number (660-1300) of pUB110-derived plasmids in B. subtilis can account for the different growth effects that plasmid pCED3 exerts on B. subtilis and E. coli.

With the success achieved in stabilizing the model heterologous gene (encoding *E. coli* ß-galactosidase) in *B. subtilis*, we became interested in exploiting our findings with genes of greater economic potential. We have thus centered our attention on the genes and enzymes of the thermophilic anaerobe, *C. thermocellum*, an organism which degrades cellulose and hemicellulose at high temperature and carries out a direct fermentation to ethanol, a potential liquid fuel for replacement or supplementation of gasoline.

Prior to these studies, we had noted (Demain and Wu, 1989) that extensive aggregation of extracellular cellulases occurs with *C. thermocellum*. The protein aggregate has a molecular weight higher than 3 million Da and contains a multiplicity of subunits. The true* cellulase (Avicelase) and carboxymethylcellulase (CMCase) activities of the culture fitrate were predominantly associated with this protein aggregate. Treatment with sodium dodecyl sulphate (SDS) was effective in dissociating the aggregate. Upon fractionation of the dissociated subunits or subcomplexes, at least four components of this aggregate could degrade CMC. However, the Avicelase activity was restored only by combining components, indicating that degradation of crystalline cellulose requires the cooperative action of different subunits. Two components, S_S (M_r = 82,000) and S_L(M_r = 250,000) are necessary and sufficient for Avicelase activity. S_L was purified by elution from an SDS-polyacrylamide (PAGE) gel whereas S_S was purified by gel chromatography of the dissociated protein aggregate.

S_L is a glycoprotein. The hydrolysis product from Avicel by the action of S_S and S_L was mainly cellobiose. The Avicelase activity displayed by S_L and S_S was inhibited by cellobiose, required Ca^{2+} for maximum effect and was inhibited by EDTA. These properties were consistent with those which we had reported earlier for the crude enzyme complex suggesting that S_S and S_L contributed in a major way to overall cellulose degradation.

S_L displayed no activity on CMC or any other synthetic substrate tested (but see below for more recent data). Conversely, S_S released reducing sugar from CMC and was therefore a ß-glucanase. The binding of S_S to the cellulose surface was dependent on S_L. A model was proposed to explain the cooperative action of S_S and S_L. In this model, S_L functions as an anchorage subunit on the cellulose surface to mediate the binding of S_S, a catalytic subunit, to the cellulose surface where degradation occurs.

Present on the extracellular surface of *C. thermocellum* vegetative cells are protuberances which contain high molecular weight, multienzyme, multifunctional complexes called cellulosomes (Lamed *et al.*, 1983); these not only are cell-bound but are found free in the extracellular medium. Cellulosomes have molecular weights from 2 to 6.5 x 10^6 Da (Lamed and Bayer,

1988; Mayer et al., 1987; Wu and Demain, 1988). These aggregated macromolecular glycoprotein complexes are arranged into polycellulosomes with sizes of 5 to 8 x 10^7 Da which line the surface protuberances. Several hundred cellulosomes appear to be present in each protuberance. Cellulosomes have been estimated to contain from 14 to 50 proteins detectable by SDS-PAGE ranging in molecular weight from 20 to 250 KDa and, like the extracellular culture fluid, they exhibit true cellulase activity which is activated by Ca^{++} and thiols and inhibited by cellobiose (Johnson and Demain, 1984; Johnson et al., 1982; Johnson et al., 1982; Lamed and Bayer, 1988, Mayer et al., 1987).

It has been extremely difficult to purify individual enzymes from undenatured cellulosomes since these structures have thus far been dissociable only by SDS. The cloning of 22 or so individual genes (encoding cellulases, xylanases and ß-glucosidases) from C. thermocellum into Escherichia coli and Saccharomyces cerevisiae (Hazlewood et al., 1988; Kadam et al., 1988; Schwarz et al., 1988; Sakka et al., 1989) has been useful but no enzyme produced by recombinant DNA technology has been able to degrade crystalline cellulose. Although we (Wu et al., 1988) had showed that the combination of two cellulosome proteins, named S_L and S_S, could completely solubilize Avicel, this occurred only at very low rates since the proteins were isolated using denaturing conditions. S_L, the largest cellulosome component, is glycosylated, highly antigenic, has a molecular weight of 210 to 250 KDa and apparently is the same as S_1 in the terminology of Lamed and Bayer (1988). S_L contains 25 to 40% carbohydrate (Gerwing et al., 1989; Wu and Demain, 1988) which is predominantly galactose (6) and binds strongly to Griffonia simplicifolia lectin GS-1 and to its homotypic isolectin B_4 (Lamed et al., 1987).

Due to the complexity of the cellulosome and the difficulties involved in dissociating it without denaturing its components, we were interested in its simplification. With the aid of a lectin-affinity material ("Jacalin"), preferentially binding galactosyl carbohydrates, we were able to isolate a simpler complex which we call the "sub-cellulosome" preparation (Kobayashi et al., 1990). This preparation was obtained by column chromatography on CM-Biogel A and then on a Jacalin column. The sub-cellulosome preparation is a macromolecular complex, composed of six main protein subunits (MW 210,000 to 58,000) as revealed on SDS-PAGE. The specific CMCase and Avicelase activities are 15- and 8-fold higher respectively than those of crude extracellular cellulase. We could not fractionate this preparation any further without denaturing it. The optimum pH and temperature of the sub-cellulosome preparation are 5.5 to 7.0 and 70°C for CMCase and 5.5 to 7.0 and 65°C for Avicelase, respectively. The subcellulosome preparation acts on various types of CMC, cellulose and para-nitrophenyl (PNP)-cellobioside but not on PNP-glucoside. Sulfhydryl reagents and N-bromosuccinimide inhibit both the CMCase and Avicelase activities whereas ethylene diamine tetraacetic acid and o-phenanthroline inhibit Avicelase activity only.

We have constructed a genomic library of C. thermocellum DNA in a lambda gt11 phage vector. Genes coding for subunits of the sub-cellulosome fraction (which is composed of six main protein subunits, named J1 to J6) were screened by an immunological reaction. Four of the genes (encoding J2, J4, J5 and J6) from the recombinant lambda phages were cloned on vectors pBR325 and pUC8 to form recombinant plasmids which were transformed into E. coli. Two of the four genes, coding for subunits J2 and J5, were found to be identical by Southern blot hybridization and restriction mapping. Each recombinant protein was prepared by ion-exchange column chromatography from heat-treated crude E. coli cell-free extract. Recombinant proteins, J2, J5 and J6 showed CMCase activity; J2 and J5 were also able to hydrolyze PNP cellobioside and xylan. Recombinant protein J4 did not show any enzymatic activities.

We also have cloned the gene (or part of the gene) encoding protein S_L. Antibodies raised against the S_L subunit of the cellulosome were used to screen a library of C. thermocellum chromosomal DNA fragments constructed in the vector lgt11. A DNA fragment was isolated which encoded a polypeptide which cross-reacted with anti-S_L antibodies. The DNA fragment was isolated and its restriction map elucidated. No similarity with other cloned DNA fragments was found. The anti-S_L cross-reacting polypeptide was isolated from recombinant E. coli and found to have a molecular weight of 37,000 Da and to possess low levels of CMCase and Avicelase activity in contrast to our earlier finding that C. thermocellum S_L has no enzyme activity. Using CMCase as the substrate, a temperature optimum of 55°C and a pH optimum of pH 6.6 were observed. These properties were compared to those of the 210,000 Da C. thermocellum S_L subunit which was isolated from an SDS gel. We found it to also possess low levels of CMCase and Avicelase activities. In addition, these two S_L-type proteins were able to positively interact against Avicel with an endoglucanase (S_S) purified from the C. thermocellum crude cellulase preparation, and a recombinant protein which cross-reacted with anti-S_S antibodies.

By using a combination of ion exchange and hydrophobic interaction chromatography, it was possible to isolate endoglucanase S_S without the use of denaturants. This endoglucanase was identified as the cellulosomal subunit S_S by the use of specific antibodies. The enzyme has a molecular weight of 83,000 Da, an isoelectric point of 3.55, optimum pH of 7.0 and optimum temperature of 70°C. It hydrolyzes CMC and at a faster rate, the cellodextrins, cellotetraose and cellopentaose, but does not hydrolyze a crystalline cellulose such as Avicel. Cellobiose and cellotriose are also immune to attack. The S_S endoglucanase differs from endoglucanases previously isolated by others and from a new 80,000 Da endoglucanase recently isolated in this laboratory and described in the following paragraph. We also have been able to clone and sequence a gene which reacts with antibodies derived from C. thermocellum subunit S_S.

The new endoglucanase from C. thermocellum was purified to apparent homogeneity without the use of denaturants. It contains no carbohydrate. A molecular weight of 80,000 was determined by SDS-PAGE. The optimal pH is pH 6.6 and the enzyme is isoelectric at pH 5.05. The enzyme has a temperature optimum of 70°C and retains approximately 50% of its activity after 48 h at 60°C. Hydrolysis of CMC takes place with a rapid decrease in viscosity but a slow liberation of reducing sugars, indicating endoglucanase type activity. It shows little ability to hydrolyze highly ordered cellulose. Cellobiose inhibits whereas magnesium and calcium ions stimulate the activity. The new endoglucanase is completely inactivated by 1 mM Hg^{2+} and is inhibited by a thiol reagent.

LIST OF PUBLICATIONS FROM THIS WORK

Demain, A. L., and J. H. D. Wu. 1989. The cellulase complex of Clostridium thermocellum. In Bioprocess Engineering, The First Generation (T. K. Ghose, ed). pp. 70-86. Ellis Horwood Ltd., Chichester, U. K.

Kobayashi, T., M. P. M. Romaniec, U. Fauth and A. L. Demain. 1990. A sub-cellulosome preparation with high cellulase activity from Clostridium thermcellum. Appl. Environ. Microbiol. 56: 3040-3046.

Shoham, Y. and A. L. Demain. 1990. Stabilization of a plasmid-encoded LacZ phenotype in Bacillus subtilis. Curr. Microbiol. 20: 373-379.

Shoham, Y., Stephanopoulos, G. and A. L. Demain. 1991. Effects of the β-lactamase gene and orientation of the kanamycin-resistance gene in plasmid pCED3 on the growth of Bacillus subtilis. J. Ferm. Bioeng. 72: 244-248.

Shoham, Y. and A. L. Demain. 1991. Kinetics of loss of a recombinant plasmid in Bacillus subtilis. Biotechnol. Bioeng. 37: 927-935.

Shoham, Y., E. Israeli, A. L. Sonenshein and A. L. Demain. 1991. Inhibition of growth of Bacillus subtilis by recombinant plasmid pCED3. Arch. Microbiol. 156: 204-212.

Fauth, U., M. P. M. Romaniec, U. Fauth, P. J. Barker and A. L. Demain. 1991. Purification and characterization of endoglucanase S_S from Clostridium thermocellum. Biochem. J. 279: 67-73.

Romaniec, M. P. M., T. Kobayashi, U. Fauth, U. T. Gerngross and A. L. Demain. 1991. Cloning and expression of a Clostridium thermocellum DNA fragment that encodes a protein related to cellulosome component S_L. Appl. Biochem. Biotechnol. 31: 119-134

Kobayashi, T., M. P. M. Romaniec, U. Fauth, P. J. Barker and A. L. Demain. 1992. Cloning and expression in Escherichia coli of Clostridium thermocellum DNA encoding sub-cellulosomal proteins. Enzyme Microb. Technol. 14: 447-453.

Romaniec, M. P. M., U. Fauth, T. Kobayashi, N. S. Huskisson, P. J. Barker and A. L. Demain.1992. Purification and characterization of a new endoglucanase from Clostridium thermocellum. Biochem. J. 283: 69-73.

PARTICIPATING SCIENTIFIC PERSONNEL

Arnold L. Demain
Yuval Shoham (Ph.D. degree granted)
Marek P. M. Romaniec
Ursula Fauth
Tohru Kobayashi
U. Tillmann Gerngross
Eitan Israeli

BIBLIOGRAPHY

Demain, A. L. and J. H. D. Wu. 1989. In Bioprocess Engineering. The First Generation (T. K. Ghose, ed.) pp 70-86. Ellis Horwood Ltd., Chichester, U. K.

Gerwing, G. J., J. P. de Waard, J. P. Kamerling, J. F. G. Vliegenthart, F Morgenstern, R. Lamed and E. A. Bayer. 1989. J. Biol. Chem. 264: 1027-1035.

Hazlewood, G. P., M. P. M. Romaniec, K. Davidson, O. Grapinet, P. Beguin, J. Millet, O. Raynaud and J. -P. Aubert. 1988. FEMS Microbiol. Lett. 51: 231-236.

Johnson, E. A. and A. L. Demain. 1984. Arch. Microbiol. 137: 135-138.

Johnson, E. A., E. T. Reese and A. L. Demain. 1982. J. Appl. Biochem. 4: 64-71.

Johnson, E. A., M. Sakajoh, G. Halliwell, A. Madia and A. L. Demain. 1982. Appl. Environ. Microbiol. 43: 1125-1132.

Kadam, S., A. L. Demain, J. Millet, P. Beguin and J. -P. Aubert. 1988. Enzyme Microb. Technol. 10: 9-13.

Kobayashi, T., M. P. M. Romaniec, U. Fauth and A. L. Demain. 1990. Appl. Environ. Microbiol. 56: 3040-3046.

Lamed, R. and E. A. Bayer. 1988. Adv. Appl. Microbiol. 33: 1-46.

Lamed, R., J. Naimark, E. Morgenstern and E. A. Bayer. 1987. J. Bacteriol. 169: 3792-3800.

Lamed, R., E. Setter and R. Kenig. 1983. Biotechnol. Bioeng. Symp. No. 13: 163-181.

Mayer, F., M. P. Coughlan, Y. Mori and L. G. Ljungdahl. 1987. Appl. Environ. Microbiol. 53: 2785-2792.

Sakka, K., S. Furuse and K. Shimada. 1989. Agric. Biol. Chem. 53: 905-910.

Schwarz, W. H., S. Schimming, K. P. Rucknagel. S. Brugschwaiger, G. Kreil and W. L. Staudenbauer. 1988. Gene 63: 23-30.

Wu, J. H. D. and A. L. Demain. 1988. In Biochemistry and Genetics of Cellulose Degradation (J. -P. Aubert, P. Beguin and J. Millet, eds.) pp. 17-131. Academic Press, NY

Wu, J. H. D., W. H. Orme-Johnson and A. L. Demain. 1988. Biochemistry 27: 1703-1709.