DOE/ER/20002--1 DOE/ER/20002-1

Conway; Grant No. DE-FG02-90ER20002 DE92 015013

PROGRESS REPORT

The funded research deals with the physiology and genetics of glycolytic flux control in *Zymomonas mobilis*. Two fundamental biological questions are being addressed: First, how do the enzymes of glycolytic pathways act in concert to regulate metabolic flux? Second, what is the role of gene expression in regulating high level synthesis of the glycolytic enzymes in a balance that allows proper glycolytic flux control? The specific objectives of the grant are as follows: **1.** To clone the structural and regulatory regions of the *Z. mobilis* genes encoding glucose-6-phosphate dehydrogenase, phosphoglucose isomerase, enolase, 6-phosphogluconate dehydratase, 2-keto-3-deoxy- 6-phosphogluconate aldolase, glucokinase and fructokinase. **2.** To characterize the structure of these genes with respect to nucleotide sequence, transcriptional initiation sites, promoter location, evolutionary relatedness to similar genes from other organisms, and organization of these genes on the genome. **3.** To investigate the effects of genetically engineered alterations in the levels of the cloned enzymes on metabolic flux and cell growth. **4.** To study transcriptional and post-transcriptional regulation of the genes encoding the enzymes of the Entner-Doudoroff pathway. The first two specific objectives have now been fully completed. Significant progress has been made on the fourth objective and work on the third objective is well underway.

Specific Objective 1.

The *Zymomonas mobilis* genes that encode glucose-6-phosphate dehydrogenase (*zwf*), 6-phosphogluconate dehydratase (*edd*), and glucokinase (*glk*) were cloned independently by genetic complementation of specific defects in *E. coli* metabolism. The identity of these three cloned genes was confirmed by various biochemical means (Barnell, et. al., 1990). The *Z. mobilis* gene encoding phosphoglucose isomerase (*pgi*) was cloned by genetic complementation of an *E. coli pgi* mutant. Enzyme assay and SDS-polyacrylamide gel electrophoresis confirmed the presence of excess amounts of phosphoglucose isomerase in *E. coli* clones carrying the *Z. mobilis pgi* gene (Hesman, et. al., 1991). The *eda* gene that encodes 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase of the Entner-Doudoroff pathway was cloned from *Z. mobilis* by genetic complementation of an *E. coli* mutant (Conway, et. al., 1991). The *Z. mobilis frk* gene, encoding fructokinase, was cloned in collaboration with Dr. Robert K. Scopes, LaTrobe University, Australia, by using a novel enzyme activity-based detection plate (Zembrzuski, et. al., submitted for publication). The *Z. mobilis eno* gene that encodes enolase was cloned by genetic complementation of an *E. coli* mutant (Burnett and Conway, unpublished data).

Specific Objective 2.

a. Evolutionary relatedness. Each of the cloned genes have now been completely sequenced. The 6-phosphogluconate dehydratase sequence was the first for this enzyme from any source. Interestingly, the deduced amino acid sequence was similar to that of a dihydroxyacid dehydratase. The glucose-6-phosphate dehydrogenase was the first sequenced from a bacterium and resembles other analogous enzymes. The glucokinase is extremely unusual in lacking a consensus glucose binding site and is dissimilar to other hexokinases. Likewise, the fructokinase is dissimilar to other hexokinases. The KDPG aldolase is similar to the *Pseudomonas* enzyme, as well as perhaps being identical to 2-keto-4-hydroxyglutarate aldolase. The phosphoglucose isomerase is similar to other analogous enzymes. Nucleotide sequence analysis established that the *zwf*, *edd*, and *glk* genes are clustered on the genome with an additional open reading frame in this region that has significant amino acid identity to the *E. coli* xylose-proton symporter and the human glucose transporter. On the basis of this evidence and structural analysis of the deduced primary amino acid sequence, this gene was believed to encode the *Z. mobilis* glucose facilitated diffusion protein, *glf* (Barnell, et. al., 1990). Genetic reconstitution of glucose transport in appropriate *E. coli* mutants has provided proof that *glf* does indeed

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

encode the transporter.

b. Gene Organization. All eight of the characterized genes are present in single copies on the genome. The *glf*, *zwf*, *edd*, and *glk* genes are clustered on the genome, in that order, and form an operon. The *glf* and *zwf* genes are separated by 146 base pairs. The *zwf* and *edd* genes overlap by 8 base pairs and their expression may be translationally coupled. The *edd* and *glk* genes are separated by 203 base pairs. The four genes of the operon are co-transcribed on a 6.14 kb mRNA (Barnell, et. al., submitted for publication). Such an arrangement of the genes that govern glucose uptake and the first three steps of the Entner-Doudoroff glycolytic pathway provides the organism with a mechanism for carefully regulating the levels of the enzymes that control carbon flux into the pathway. The *eda*, *pgi*, *frk*, and *eno* genes are all monocistronic. The *eda* gene is not tightly linked to the *edd* gene, as in other bacteria. No evidence for linkage of the other glycolytic genes has been obtained, although experiments specifically designed for this purpose have not been conducted.

c. Transcriptional Initiation Sites and Promoter Location. A single site of transcriptional initiation for the polycistronic *glf-zwf-edd-glk* message was mapped by primer extension and nuclease S1 protection analysis (Barnell, et. al., submitted for publication). The *eda* gene is transcribed from a single promoter that was mapped with respect to the transcriptional initiation site (Conway, et. al., 1991). The same is true for the *pgi* gene (Hesman, et. al., 1991). The *eno* promoter was also mapped (unpublished). Unfortunately, the *frk* promoter has yet to be cloned. The *glf* operon promoter region was found to share significant identity to the *eda*, *pgi*, *eno*, and other highly expressed *Z. mobilis* promoters, but not to consensus promoters from other bacteria. The highly expressed *Z. mobilis* promoter set contains two independent, overlapping, conserved sequences that extend from approximately -100 to +15 with respect to the transcriptional start sites (Barnell, et. al., submitted for publication). It is quite possible, perhaps even likely, that the structures of these highly expressed promoters identify them as members of a separate expression class that is characterized by high transcription rate.

d. Gene Structure. One of the primary aims of this work has been to identify features of gene structure in *Z. mobilis* that facilitate high level expression of the glycolytic enzymes. High level expression of the genes can be attributed, in part, to very efficient translational initiation due to the high quality of the ribosome binding sites. A comparison of the highly expressed *Z. mobilis* genes indicated that the relative quality of the ribosome binding sites of these genes might play an important role in determining the relative level of enzyme synthesis (Conway, et. al., 1991). Interestingly, the *eno* gene does not have a strong ribosome binding site (unpublished data), despite being the most abundant protein in *Z. mobilis* (Mejia, et. al., submitted for publication). This discrepancy is being pursued.

Specific Objective 3.

The available experimental evidence indicates that glucokinase, glucose-6-phosphate dehydrogenase and enolase are the most likely flux-generating steps in glucose fermentation by *Z. mobilis*. In accordance with metabolic control theory, changes in rate, even if caused by an enzyme inhibitor, can be considered to be equivalent to changes in the concentration of the enzyme. The levels of these enzymes of the Entner-Doudoroff pathway in *Z. mobilis* can be perturbed by expression of the cloned genes from multi-copy plasmids. Cotranscription of *zwf* and *glk* has greatly complicated the proposed work. In order to express these genes in recombinant strains, it was first necessary to subclone the individual structural genes onto appropriate broad host range plasmids under control of heterologous promoters. For this purpose, a *tac* promoter vector and the enteric *lacI* gene have been used to construct a vector that allows regulation of transcription rate, in *Z. mobilis*, by exogenous inducer (IPTG). Work with the *eno* gene has just now progressed to the point that equivalent constructions with this gene can be made.

This system shows promise for pursuit of objective 3, which shall commence in the coming months.

Specific Objective 4.

a. Transcriptional Regulation. Expression of *pgi*, *frk*, and the *glf* operon have been found to be subject to carbon source-dependent regulation. The level of the 1.8 kb *pgi* transcript was found to be influenced by the phase of growth and by the carbon and energy source (Hesman, et. al., 1991). Transcript levels increased with respect to total RNA during logarithmic growth and were 3-fold higher when growing on fructose as compared to glucose. These changes in transcript levels paralleled phosphoglucose isomerase activities in the cultures. The level of the full-length 6.14 kb *glf* operon message was three-fold higher in cells grown on fructose as compared to cells grown on glucose (Barnell, et. al., submitted for publication). Likewise, the 1.2 kb *frk* transcript was elevated two-fold when growing on fructose (Zembrzuski, et. al., submitted for publication). In each case, enzyme levels paralleled mRNA levels. This increase was not the result of differential mRNA processing when grown on the different carbon sources, since the patterns of mRNA bands on northerns was similar for RNAs prepared from either glucose- or fructose-grown cells. Nor was it the result of differential transcript stability, since the degradation rates for each of the mRNAs was similar in fructose-grown cells when compared with glucose-grown cells. Thus, the higher level of these mRNAs in fructose-grown cells is the result of an increased rate of transcription from the three respective promoters. The physiological role of increasing phosphoglucose isomerase and fructokinase activities is to increase flux of fructose into the pathway. The increase in *glf* expression when growing on fructose suggests that the same transporter is used for both glucose and fructose transport, since it is known that affinity of the transporter is far greater for glucose than for fructose. The remainder of the *Z. mobilis* genes that have been characterized to date, *eda*, *gap-pgk*, *eno*, *adhB*, and *pdc* do not show carbon source-dependent regulation. It seems likely that the promoters of the three regulated genes are somehow different than the others, but computer comparisons have failed to reveal "fructose-specific" promoter domains.

Other aspects of transcriptional regulation, such as the participation of DNA binding proteins, are being assessed by so called "promoter titration" experiments. This work involves examining the effects of having multiple copies of the various promoters on expression of the intact, genomic copies of the genes. The appropriate plasmids for constructing these recombinant strains are in various stages of completion and will be tested soon.

b. Relative Transcript Stability. It is likely that high level expression of the glycolytic genes is coordinated at the level of mRNA abundance. The relative abundance of specific mRNAs is regulated by the relative rates of transcription and message decay. In a collaborative study with Dr. L. O. Ingram, transcript stability and the abundance of the gene products have been determined for eight of the glycolytic enzymes (Mejia, et. al., submitted for publication). The transcripts were all unusually stable, with half-lives ranging between 8 and 18 minutes. Transcript stability is proposed to be a dominant feature in distinguishing between the highly expressed glycolytic genes and the vast majority of poorly expressed biosynthetic and house keeping genes. Among the eight glycolytic genes studied, differences in transcript stability were found to correlate directly with protein levels. Thus, differential transcript stability is hypothesized to be a primary mechanism for balancing the levels of individual enzymes of the *Z. mobilis* glycolytic pathway.

c. Post-transcriptional Regulation. The full-length 6.14 kb polycistronic *glf-zwf-edd-glk* mRNA from *Zymomonas mobilis* appears to be processed by endonucleolytic cleavage, resulting in the formation of several discrete transcripts (Liu, et. al., submitted for publication). Northern analysis and transcript mapping revealed that the processed transcripts correspond to func-

tional mono-, di-, or tri-cistronic messages. The relative abundance of the gene-specific, functional messages was measured. Expression of *zwf* and *edd* correlated well with functional message levels. Disproportionally high levels of the *glk*-specific mRNAs might compensate for the instability of glucokinase by allowing increased translation. The relative abundance of the discrete transcripts was shown to be a function of their respective decay rates. Northern analysis of the fate of the 6.14 kb transcript following inhibition of transcription by rifampicin showed that the abundance of shorter, more stable transcripts increased at the expense of longer, less stable transcripts. This is suggestive of endonucleolytic mRNA processing. The most abundant 5' and 3' transcript ends were found to lie within secondary structures that probably impart stability to the most abundant mRNAs. The exact chronology of these events remains to be elucidated.

d. Relative Glycolytic Gene Expression. The genetic mechanisms by which high level synthesis of the Entner-Doudoroff glycolytic pathway enzymes are balanced in *Z. mobilis* are beginning to be understood. A growing body of evidence indicates that several layers of genetic control, perhaps in a hierarchal arrangement, act in concert to determine the relative abundance of the enzymes. The genes all possess canonical ribosome binding sites and display high codon bias. Co-translation of the overlapping *zwf* and *edd* genes might be important for controlling the balance of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydratase. The relative stabilities of the mRNAs appear to play a role in coordinating expression of the glycolytic genes. Likewise, mRNA processing appears to be important for dictating an appropriate balance of the glycolytic enzymes. Promoter strength and the relative rates of transcription of the glycolytic genes, which appear to form a separate, high expression class, are likely to be important. One further level of control might be imposed by modulation of promoter activity by transcription factors.

CONTINUATION REQUEST

At this time, the research plans remain intact relative to the original proposal. To date, the first two of four specific objectives have been achieved and significant progress has been made on the fourth objective. The research has generated three published papers, with an additional four manuscripts in review. These reprints and preprints are enclosed. Several intriguing lines of investigation lie ahead. The fortuitous cloning of the glucose facilitated diffusion transport is a real stroke of luck. This is the only fully accepted example of a facilitator system in the prokaryotic world. Over the next year this picture will be developed more completely. The majority of effort will be devoted to fulfillment of the third specific objective, an investigation of potentially rate-limiting steps in glycolytic metabolism. *Z. mobilis* has proven to be a most interesting organism and gives every indication of harboring additional secrets that are worth pursuing.

papers removed and cycled separately -

DISCLAIMER

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

DATE
FILMED
8/12/92