I. Original Summary of Proposal of Current Award
Project DE-FG03-95ER 20196, begun August 1995

The anaerobic organism *Clostridium acetobutylicum* has been used for commercial production of important organic solvents due to its ability to convert a wide variety of crude substrates to acids and alcohols. Current knowledge concerning the molecular genetics, cell regulation and metabolic engineering of this organism is still rather limited. The objectives are to improve our knowledge of the molecular genetics and enzymology of Clostridia in order to make genetic alterations which will more effectively channel cell metabolism toward production of desired products. Two factors that limit butanol production in continuous cultures are: 1) The degeneration of the culture, with an increase in the proportion of cells which are incapable of solvent production. Currently isolated degenerate strains are being evaluated to analyze the molecular mechanism of degeneration to determine if it is due to a genetic loss of solvent related genes, loss of a regulatory element, or an increase in general mutagenesis. Recent studies show two general types of degenerates, one which seems to have lost essential solvent pathway genes and another which has not completely lost all solvent production capability and retains the DNA bearing solvent pathway genes. 2) The production of hydrogen which uses up reducing equivalents in the cell. If the reducing power were more fully directed to the reduction reactions involved in butanol production, the process would be more efficient. We have studied oxidation reduction systems related to this process. These studies focus on ferredoxin and rubredoxin and their oxidoreductases.

II. Summary of Progress on DE FG03-95ER 20196

The report covers the period 8/95-present and the latter portion of previous award DE-FG05-92ER 20092 to include coverage of the last eighteen months 1/95-6/96 as requested.

Several degenerate strains were isolated and characterized by sporulation, motility and growth properties. Cell appearance and colony morphology were also recorded. Enzymatic assays revealed reduced butyraldehyde dehydrogenase and Co-A transferase enzyme activities in the degenerates. DNA analysis revealed that in complete degenerate strains the genes of the solvent locus were absent. Gyrase inhibitors slightly reduced the growth rate and decreased acetone formation preferentially. In an effort to analyze the role of sporulation sigma factors in solvent gene expression, recombination experiments were conducted and led to strains with increased solvent production.

Analysis of redox systems has resulted in the sequence analysis of a cluster encoding formyl transferase proteins and an oxidoreductase-like gene. The genes for the two subunits of an apparent electron transfer flavoprotein were sequenced and suggest this factor acts to carry electrons to the butyryl-CoA dehydrogenase. Investigations related to ATP metabolism have led to the cloning and sequencing of the two key regulatory enzymes of glycolysis, phosphofructokinase and pyruvatekinase. The genes encoding the Fo subunits of the membrane ATPase have been sequenced.

III. List of Publications

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.
DISCLAIMER

Portions of this document may be illegible electronic image products. Images are produced from the best available original document.


IV. Progress Report

IV.1. Analysis of Degenerate Strains

Degenerate strains were isolated by transfer of C. acetobutylicum ATCC 824 cultures. An overnight freshly heat shocked Clostridial Growth Medium (CGM) culture was streaked onto Nutrient Yeast Glucose (NYG) or Reinforced Clostridial (RCM) agar plates. After 4-6 days cells from the edges of the colonies were restreaked onto new plates and the transfer repeated as many as nine times. Some were also transferred from broths to plates repeatedly. Colonies arising from the treatment were subjected to crude solvent testing by a color test and those appearing to have reduced solvent were tested by gas chromatography of small cultures grown for various intervals from 22-70 hrs. In contrast to wild-type cultures which produced 80 mM butanol and ~40 mM acetone, the levels of these products from the degenerate mutants were low, less than 10 mM butanol and <1 mM acetone for some mutants. Also found were mutants which produced low acetone but were less affected in butanol levels. In general, the NYG procedure was better in producing degenerate cultures. The colony morphology of the degenerates was similar to that described previously by Adler as being large, white, diffuse flat colonies with irregular edges and a less defined center and no outgrowths, as compared with the wild type. Examination of the degenerate mutants under the phase contrast microscope indicated the presence of non-motile long filaments and some short motile rods. The relative proportion of each type varied among the mutants. The motility of the mutants was studied using motility agar since earlier work had reported reduced solvent production from less motile stocks. These mutants seemed to have little or no motility compared to wild-type C. acetobutylicum. In addition to mutants isolated in the above fashion, degenerate strains have been obtained from long-term chemostatic cultures analyzed for solvent production.

IV.1.1. Enzyme analysis
Enzyme analyses of the mutants were performed at various times in batch cultures and cell extracts were made and analyzed from log phase cells or as the culture entered stationary phase. The degenerate mutants grew to a lower density (O.D. 2.7-3.0 vs. 5 for wild type) and achieved a lower final pH (3.2) in the CGM broth than the wild-type strain. Assays of phosphotransbutyrylase and butyrate kinase were near normal throughout the culture time. The levels of NADH dependent BDH were lower than wild type and declined to zero late in the culture, in contrast to the wild-type strains in which these activities are induced in the later stages of the fermentation. NADH- and NADPH-dependent butyraldehyde dehydrogenase levels were below normal in the mutants. In some mutants the enzyme was completely eliminated as had been found in the previously characterized mutant M5. The assays of Co-A transferase indicated that the complete degenerate mutants were completely deficient in this enzyme. Some partial degenerates seemed to still retain reduced levels of this enzyme. However, the lower growth of the degenerate cultures raised the possibility that conditions were not completely reached for optimal induction of the Co-A transferase. The Co-A transferase was the most indicative marker of the degenerate state. Acetoacetate decarboxylase, an enzyme involved in acetone production whose gene is located adjacent to those for the subunits of the Co-A transferase, has not been assayed in all mutants. The morphology, enzyme and metabolite data from the degenerate mutants provide us with detailed information which describes the properties of the strains.

IV.1.2. DNA analysis

In our DNA analysis we were particularly interested in the segment of DNA around the Co-A transferase gene. We wanted to see if there had been any specific type of deletion or other small rearrangement in the genome of these degenerate strains. In at least one strain exhibiting a zero solvent production phenotype, the appropriate PCR product was not found upon amplification. An analysis by Southern hybridization revealed that the genes encoding acetoacetate decarboxylase (adc), acetoacetyl-coenzyme A:acetate:butyrate:coenzyme A transferase (ctfAB), and the alcohol/aldehyde dehydrogenase (aad or adhE) were absent in the complete degenerates. The A and B genes encoding butanol dehydrogenases were retained, indicating they are not encoded on the large plasmid. This was also the case for the thiolase A and ptb genes. However, others (138) and collaborators (139) have shown that the loss of solvent genes correlates with the loss of a large plasmid which bears these solvent genes. The analysis of the partial degenerates could yield information about this process. The strains exhibiting partial loss of solvent production could represent another type of degeneration or could arise from a mixed culture which contains some cells that have lost the plasmid and some which retain it. Also possible is a deletion of part of the plasmid.

IV.1.3. Complementation of degenerates

Recent complementation experiments have introduced solvent stage regulated genes cloned on plasmids into degenerate cells by electroporation. In this experiment degenerate cells of the zero solvent production class were used. These experiments have indicated that the solvent-stage enzymes can be expressed in the degenerate cells of this class. This finding suggests that the degenerated state is not due to the lack of a particular required regulatory factor (sigma factor, gene activator, etc.) which is lost in the degenerate strains. This finding is consistent with the loss of the normal solvent gene cluster in the degenerate strains. The relatively higher expression of the solvent genes in the early stages of fermentation would correlate with the absence of a negative regulator in the degenerate cells. This finding would fit with the recent observation of a negatively acting protein encoded on the plasmid (93).

In another finding from complementation experiments on degenerate strains, evidence for an additional plasmid encoded butyraldehyde dehydrogenase was found. The degenerate strain M5, which has lost the 210 kb plasmid containing the solvent genes studied above, and a strain made by recombination to inactivate the aad (adhE) gene were used as hosts for plasmids bearing either the aad (adhE) gene or the artificial acetone operon (ace) (containing the acetoacetate decarboxylase gene as well as the genes for the two subunits of the Co-A transferase). When the metabolite profile of the cultures was examined and compared to that of
the wild-type culture, the introduction of the plasmid bearing aad restored ability to form butanol. It was also found that introduction of the ace operon plasmid into strain M5 yielded a strain which produced only acetone but no butanol, while introduction of the ace operon plasmid into the aad (adhE) mutant yielded a strain which produced butanol as well as acetone. This suggests that function of Co-A transferase and acetone production may lead to a condition which stimulates butyraldehyde dehydrogenase activity and it cannot achieve this state without the presence of the large plasmid. Therefore either an additional butyraldehyde dehydrogenase is encoded on the large plasmid and it is stimulated by events associated with the formation of acetone, or some regulatory factor required for expression of this enzyme is encoded on the large plasmid. Future analysis of this effect is planned.

IV.1.4. Supercoiling

The level of supercoiling in DNA is reduced in most species as they undergo entry into stationary phase. The level of supercoiling can be potentially critical for plasmids, affecting their replication and loss from cells. In this regard, our study of the effects of novobiocin, a DNA gyrase inhibitor, on the entry into solventogenesis may now be considered in a new context, that related to effects on plasmid encoded as well as chromosomal encoded genes. While these experiments did not reveal a very dramatic result, they are useful in showing that supercoiling does change in C. acetobutylicum as it does in other species during culture stage transitions and that other aspects may need to be taken into account for a full explanation of the mechanism of the metabolic shift.

In specific results, cells treated with novobiocin, a DNA gyrase inhibitor, produced higher butyrate levels and lower solvent levels with acetone being the most affected. Seven enzyme activities involved in acid and solvent production were analyzed; among them, CoA transferase required for acetone formation and acid uptake, experienced a decrease in activity. Like Escherichia coli and Bacillus subtilis, DNA from C. acetobutylicum became less negatively supercoiled in the early solventogenic stage (stationary phase), as shown by analysis of linking number of a reporter plasmid by agarose gel electrophoresis in the presence of chloroquine.

IV.1.5. Recombination experiments

Three sporulation-specific genes (orfA, sigE, sigG) from Clostridium acetobutylicum ATCC 824 are arranged in a cluster, encoding the putative σE-processing enzyme, σE, and σG respectively. We were interested in possible relations between solvent stage processes and sporulation and conducted some experiments to study the role of these genes. When they were transformed into Clostridium acetobutylicum while on a plasmid functional in this organism, transformants did not survive. Three kinds of recombinations were then attempted with nonreplicative plasmids: duplication of orfA and sigE, replacement of all of the three genes, and inactivation of orfA. While the wild-type strain ceased to grow and produce solvents in batch cultures after approximately 24 h, certain mutant strains were isolated that showed sustained growth for a much longer time and produced a threefold increase in acetone and butanol in test tube cultures. In addition, one of the derived strains showed a significantly higher growth rate. Features of the restriction maps of the recombinants did not correlate with expected maps, indicating possible complications occurring during the recombination events, so the basis for increased production cannot be determined. Further analyses of some of these strains are planned.

IV.2. Analysis of Oxidation-Reduction Processes in Clostridium acetobutylicum ATCC 824

Due to recent information available on genes of hydrogenase, rubredoxin and ferredoxin and their oxidoreductases from various organisms, we have begun efforts to clone these genes from C. acetobutylicum ATCC 824. The strategy is to make PCR primers to conserved regions in these genes and to amplify a segment from the C. acetobutylicum genome
and use this specific fragment as a probe to identify recombinant clones bearing the corresponding genes. For ferredoxin, two conserved regions were chosen and oligonucleotides of 23 nucleotides were synthesized. Amplification resulted in a fragment of ~100-120 bp, as expected. In the case of rubredoxin, oligonucleotides of 26 and 29 nucleotides were synthesized and a product of the expected size was found (~100-120 bp). For rubredoxin oxidoreductase, oligonucleotides of 30 nucleotides and 42 nucleotides were synthesized and, when used in a PCR experiment with genomic C. acetobutylicum DNA, a band of ~500 bp was observed, as expected. For work with the hydrogenase, the sequence of hydrogenase I from C. pasteurianum was very useful, as this enzyme has a 4Fe-4S active site and is different than the more widespread Ni-containing enzyme. The information on this enzyme was most useful for our approach. Oligonucleotides were made corresponding to the amino acid segments around positions 280 and 420 and another set at amino acids 300 and 500. Upon use of these sets in PCR amplification with C. acetobutylicum DNA products of ~350 bp and 600 bp were observed.

IV.2.1. Rubredoxin-ferredoxin related experiments

Of these recombinant phages isolated from the above screening, a phage isolated with a rubredoxin oxidoreductase probe was analyzed and found to encode a putative primosomal protein n' and genes for the formylation reactions connected with translation. Some rubredoxin oxidoreductase activity was observed in E. coli bearing a plasmid with this region. Structural similarity between the primosomal protein, n', and oxidoreductases have been observed. Whether such activity is part of the normal function of this protein or if it is just some nonspecific activity is unclear.

Most of a novel operon from C. acetobutylicum has been cloned and sequenced. The operon is composed of polypeptide deformylase, formyltransferase and an ORF with homology to both rubredoxin oxidoreductase and primosomal protein n'. Polypeptide deformylase (def) plays an essential role in the maturation of eubacterial, chloroplast and mitochondrial proteins. Deformylase removes the formyl group from the N-terminal methionine post-translationally. Formyl-tRNA transformylase (fmt) is responsible for transferring the formyl group to the initiator Met-tRNAfMet. Transformylase's activity enhances the rate of initiation in vitro, but may be dispensable under certain growth conditions or genetic contexts. Closely related on a functional level, it is not surprising to find def sharing an operon with fmt in Escherichia coli, and Thermus thermophilus. The first ORF's deduced 432 amino acid sequence has 58% homology to the rubredoxin oxidoreductase of Clostridium pasteurianum and 53% homology with the C-terminus of primosomal protein n' of Escherichia coli. The ORF continues upstream of the putative start site of the rubredoxin oxidoreductase and lacks a definitive ribosomal binding site.

The second ORF in the operon, def, encodes a deduced 150 amino acid (16,500) polypeptide deformylase. Deformylase has recently been found to be related to the zinc protease family. The zinc proteases are characterized by the amino acid motif HEXXH. This motif is found in the C. acetobutylicum deformylase. The functional activity of the C. acetobutylicum deformylase gene was shown by complementation experiments with E. coli def mutants.

The fmt gene, encoding formyltransferase, starts 16 nucleotides (nt) downstream of def in C. acetobutylicum. However, only 600 nt were sequenced before reaching the end of the insert. Approximately 300 nt remain to be sequenced. The incomplete sequence shows good homology with the N-terminal portion of formyltransferases from other organisms.

Work on the ferredoxin redox system has been undertaken by Petitdemange in France, and we have cooperated in their effort.
IV.2.2. Hydrogenase

Hydrogenase plays a pivotal role in controlling electron and ultimately carbon flow during the fermentation and regulation of hydrogenase activity, at the enzymatic level, has been used to manipulate carbon flow. For example, carbon monoxide inhibits hydrogenase and CO additions were found to inhibit cell growth and shift metabolism towards lactate and acetoin production. More recently, CO additions have been used to increase the concentrations of ethanol and butanol. Publications from other groups have recently reported the cloning and characterization of the gene encoding hydrogenase (hydA) from *C. acetobutylicum* strains P262 (58) and ATCC 824 (59). The two hydA genes are about 1700-bp and encode 64-KDa proteins. The genes are about 70% identical and display strong identity with the (Fe) hydrogenase family of genes from *Desulfovibrio* and Clostridium species.

We have worked toward increasing ethanol and butanol production in *C. acetobutylicum* by using a novel method, based on the homologous recombination, to inactivate hydA on the chromosome (112). Plasmids bearing the hydrogenase gene, PMFH1, or a portion of it, pSM1, have been obtained from P. Soucaille. The hydA inactivation plasmid pSM1 contains a 1.1 kb HindIII-EcoRI hydA fragment from pMFH1 (59). The plasmid DNA was methylated, in vivo, and then transformed into *C. acetobutylicum* by electroporation using a previously described method (105,106). Several transformants were selected for characterization. Chromosomal DNA was analyzed by Southern hybridization and preliminary experiments to analyze fermentation products by gas chromatography have been undertaken. Initial experiments to analyze the Em<sup>R</sup> recombinant strains suggested a difference in hybridization pattern using a hyd gene segment as probe and an increase in solvent levels in tube cultures. Subsequent experiments have indicated that the mapping is not as expected and hydrogenase activity remains. Thus there may be some instability in the hydA gene recombinants, or a strong selection for some hydrogenase activity in early growth. Alternative types of hydrogenase knockouts and routes to decrease hydrogenase activity through genetic means have been considered.

IV.2.3. Electron transfer flavoproteins

Between the genes encoding β-hydroxybutyryl-CoA dehydrogenase (hbd) and butyryl CoA dehydrogenase (bcd), two ORF's were found. The bcd proximal ORF encodes a polypeptide (etfB) homologous to FixA of *A. caulinaudans* and *R. meliloti* and the electron transfer flavoprotein β-subunit, while the ORF distal to bcd encodes a polypeptide (etfA) homologous to FixB of various rhizobia. Although no specific biochemical reaction has been identified for the FixAB gene products, it has been suggested they are involved in electron transport to nitrogenase. The electron transfer flavoproteins play an important role in the β-oxidation of fatty acids by coupling flavoprotein dehydrogenases to the electron transport chain. ETFs have also been isolated and characterized from bacterial systems such as *P. elsdenii*, *M. elsdenii*, and *P. denitrificans*. The proteins are generally heterodimers of α- and β-subunits.

The location of these genes adjacent to that of the butyryl-CoA dehydrogenase and the enzymatic coupling between electron transfer flavoproteins and the acyl-CoA dehydrogenases of other species suggest that these flavoproteins are likely to be the carriers in the reduction of crotonyl-CoA to butyryl-CoA by the butyryl CoA dehydrogenase.

The cloning of the etfAB genes provides some insight into important but previously uncharacterized elements in the clostridial butyryl-CoA synthesis pathway. Neither ETF nor the related enzyme BCD have been purified from *Clostridia acetobutylicum*. Only recently was the presence of BCD-related ETF in solvent-producing Clostridia suggested (247) and a FixB related ORF observed in *C. acetobutylicum* P262 (79). In P262, however, the gene arrangement is different than in *C. acetobutylicum* 824, the fixB (etfA) gene being adjacent to that of hbd and an alcohol dehydrogenase (adh).

IV.3. ATP Metabolism
The reactions involved in ATP metabolism are key factors in the energy levels of the cell and likely play an important role in the growth and metabolic efficiency of the cell. For this reason two areas have been investigated through cloning to begin to analyze how these systems play a role in defining ATP levels. The reactions of glycolysis provide the cells energy through substrate phosphorylation reactions. An important connection between ATP and membrane gradients is the membrane F$_{1}$F$_{0}$ ATPase. Genes corresponding to key control point enzymes in glycolysis and ATPase have been isolated and sequence analysis has been performed.

IV.3.1. ATPase

In aerobic organisms, the F-type ATP synthase (F$_{1}$F$_{0}$) catalyzes the production of ATP, driven by an electrochemical proton gradient formed by respiration. The enzyme complex can also generate an electrochemical proton gradient through the hydrolysis of ATP. A sodium pump, rather than a proton pump, is found with the F-type synthase in Propionigenium modestum. ATPases are composed of two distinct entities: a catalytic component (F$_{1}$) possessing 5 subunits (α, β, γ, δ and ε), and a structural component (F$_{0}$) possessing 3 subunits (a, b and c). An additional subunit (i) has recently been found upstream of the others. Its function is still unclear.

An operon possessing atp-i and the three genes composing the F$_{0}$ portion of F-type synthases has been cloned and sequenced from C. acetobutylicum. Within the operon: the 345 nucleotide (nt) atp-i is followed by the 663 nt atp-a; which is followed by the 243 nt atp-c; finally followed by the 477 nt atp-b. Thirteen nucleotides separate atp-i and atp-a. Seventy-seven nucleotides separate atp-a and atp-c. Fifty-two nucleotides separate atp-c and atp-b. There was no apparent transcriptional stop site. This is the common gene arrangement in F-type ATP synthases from non-photosynthetic bacteria. The five genes encoding the F$_{1}$ subunits are usually located downstream of the F$_{0}$ genes. This may also be the case in C. acetobutylicum, but the insert ends 20 base pairs beyond atp-b. Labelled oligos from the F$_{0}$ operon could be used to clone the F$_{1}$ operon.

Subunit I is a 115 amino acid (12,800 Da) protein that is very hydrophobic, containing 4 transmembrane domains. It shows modest homology (~50%) with atp-i from other organisms. Subunit a is a 221 amino acid (24,300 Da) protein that is extremely hydrophobic, containing 5 membrane-spanning helices. Its highest homology (66% identity and 35% identity) is found with the Na$_{2}^{+}$ motive Propionigenium modestum. Subunit c is an 81 amino acid (8300 Da) protein forming a hairpin-like structure with 2 transmembrane domains separated by a hydrophilic loop. Its highest homology (71% similarity and 44% identity) is with the cyanobacterium Synechococcus 6301. Subunit b is a 159 amino acid (18,200 Da) protein with a hydrophobic C-terminus anchoring it to the membrane. Its highest homology (58% similarity and 34% identity) is with Bacillus megaterium. The hydrophobicity profiles and homologies indicate that this operon encodes subunit I and the F$_{0}$ portion of an F-type synthase from C. acetobutylicum. However, the homology found with P. modestum's subunit a, and to a lesser extent between the other subunits, makes the issue of the type of gradient unclear.

IV.3.2. Analysis of genes of glycolysis

Phosphofructokinase (PFK) catalyzes the rate-limiting step of glycolysis and has a large number of metabolic regulatory effectors. This makes it an attractive candidate for being the crucial regulatory enzyme of glycolysis. Together, with pyruvate kinase (PYK, another enzyme allosterically controlled by glycolytic intermediates), PFK can mediate glycolytic flux. Gram-positive organisms (including the mollicutes), have a "regulatory operon" possessing pJk and pyk.

The nucleotide sequence of a 1.7-kb region of Clostridium acetobutylicum ATCC 824 DNA, including one complete and one incomplete open reading frame (ORF), has been determined. The deduced 319 amino acid sequence of the complete ORF encoded a protein of 34-kDa and was homologous to phosphofructokinases from other organisms. The deduced 39 amino acid sequence of the adjacent incomplete ORF was homologous to the N-termini of pyruvate kinases from other organisms. The partial pyk followed pfk by 40 nucleotides, and
there was no apparent transcriptional termination signal. This gene arrangement is similar to that found in other gram-positive organisms which utilize the EMP pathway for glycolysis. An experiment using reverse transcriptase-PCR to detect a transcript containing the coding region of PFK and PYF indicated that both were present on a single mRNA. The gene arrangement and function of the encoded enzymes suggest that this operon is important in defining glycolytic flux for C. acetobutylicum and likely plays an important role in growth and energetics of the cell. Complementation experiments with cloned C. acetobutylicum PFK and E. coli pfk mutants have shown the clostridial enzyme can function in E. coli, and enzyme assays have shown a corresponding increase in activity of PFK.