Lars G. Ljungdahl
Progress Report (4/30/91 - 4/30/92
and Outline of work for the period
9/1/92 - 9/1/93

Clostridium thermocellum. We are continuing our efforts to partly dissociate the cellulolytic enzyme complex of C. thermocellum. This complex named cellulosome (also existing as polycellulosome) consists of perhaps as many as 26 different subunits. It is extremely resistant to dissociation and denaturation. Treatments with urea and SDS have little effect unless the latter treatment is at high temperature. Significantly, some of the subunits after SDS dissociation have CMCase (endoglucanase) activity but no activity toward crystalline cellulose. The only reported success of hydrolysis of crystalline cellulose by cellulosomal subunits is by Wu et al. (Biochemistry 27, 1703-1709, 1988) who isolated two protein fractions labeled SL and SS which when combined exhibit a low (about 1% of the original cellulosome) activity toward crystalline cellulose. Our long standing goal is still to determine the activities of the individual subunits, to characterize them, to find out how they are associated in the cellulosome, and to establish the minimum number of subunits needed for efficient hydrolysis of crystalline cellulose.

The cellulolytic activity of the cellulosome is stimulated by Ca²⁺ and mercaptaethanol or other sulfhydryl containing compounds. We have found that the cellulosome dissociates into subcellulosomal particles upon treatment with 2mM EDTA. The conditions for reproducible dissociation are incubation with EDTA at 37°C for 18 hrs in 50 mM Tris pH 7.5 and 0.1 M NaCl. Gel filtration of the incubation mixture on Superose 6 (range 5000 x 5 x 10⁶ Da) yields 5 distinct fractions of decreasing molecular weights. The first is undissociated cellulosomes. The other 4 are subcellulosomal particles differing both quantitatively and qualitatively in subunit compositions. The fractions have endoglucanase activity, and together they act on crystalline cellulose. The fractions contain carbohydrates and at least some of the subunits are glycoproteins. The removal of the carbohydrates with NaBH₄ simplifies the protein patterns of the subcellulosomes on SDS-PAGE. This indicates that at least some polypeptides differ in glycosylation but likely contain the same or similar protein moiety thus yielding complicated SDS-PAGE patterns before the NaBH₄-treatment. We are now isolating the different polypeptides of the cellulosome and the subcellulosomal particles. Amino acid analyses and partial (N-terminal) sequencing should allow us to identify the different subunits and aid in their identification with gene products obtained by cloning of cel genes from C. thermocellum into E. coli (Beguin Annu. Rev. Microbiol. 44, 219, 1990), and also to identify the different polypeptides in the subcellulosomal particles, yielding a map. The possibility of obtaining antibodies
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 in electronic image products. Images are produced from the best available original document.
against individual polypeptides of the cellulosome complex will also be investigated. In preparation of the antibodies we will first remove any carbohydrates from the polypeptides.

*Thermoanaerobacter ethanolicus.* We have published our results on the isolation and characterization of the FDP-dependent lactate dehydrogenase (Bryant, *J. Enzyme Inhibition.* 5, 235-248, 1991) and alcohol dehydrogenases from mutant strains (Bryant et al. In press). Present work is concentrated on the secondary alcohol dehydrogenase. It is a remarkable enzyme with very broad specificity. It reduces a number of both aliphatic and aromatic compounds containing carboxy group to corresponding secondary alcohols. The reaction is stereo-specific. This specificity depending on the temperature (Pham et al. *J. Am. Chem. Soc.* 111, 1935, 1989). The enzyme is very thermostable. The work on this enzyme which is being done in cooperation with Dr. Robert S. Phillips of the Department of Chemistry at UGA has led to the suggestions that its structure should be determined by X-ray crystallography. To compliment the structure work we are now cloning and sequencing the enzyme.

A second aspect of *T. ethanolicus* is an investigation of its cellobiase or β-glucosidase. The original goal was to obtain a thermostable cellobiase to be inserted into *C. thermoaceticum* to give it the ability to grow on cellobiose. It turned out that *T. ethanolicus* has two cellobiases, one is constitutive, the other is induced when the bacterium is grown on cellobiose. The first one is hydrolyzing xylobiose and its main functions may be that of a xylobiase. Both enzymes are being isolated and characterized. The induced cellobiase has a very high thermostability and we will attempt to clone it and insert it into *C. thermoaceticum* as originally was planned. However, presently we do not have a system for cloning into *C. thermoaceticum*. We propose to develop such a system. Dr. S. Kushner, Head of the Department of Genetics, UGA has kindly agreed to cooperate with us in this endeavor.

*Clostridium thermoaceticum.* We have finished the sequencing of carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH). (Morton et al. *J. Biol. Chem.* 266, 23824, 1991). The enzyme is a nickel-iron-sulfur protein consisting of two subunits in the form of (αβ). The α subunits consists of 729 amino acids (M, 81,730). The β subunit consists of 674 amino acids (M, 72,928). No significant homology was found to any known sequence by computer search of GenBank™. The sequence of CO dehydrogenase from *Methanothrix soengenii* has been published (Eggen et al. *J. Biol. Chem.* 266, 6883, 1991). This enzyme consists of two large (79, 4kDa) and two small (19kDa) subunits. No extensive homology exists between the two enzymes. However, two shorter sequences of the *C. thermoaceticum* β subunit have homology with two sequences of the large subunit of the *M. soengenii* enzyme. They are residues 495-500 (VватГС) and 548-551 (ГСВ) of the *C. thermoaceticum* enzyme and residues 544-549 and 583-586, respectively, of the *M. soehngennii* enzyme. Our long term goal of CODH is to identify the Ni and iron
sites of the enzyme. Recent cyanide inhibition studies indicate
two separate Ni sites and it is possible that there is a site for
CO dehydrogenase activity and a second site for acetyl-CoA synthase
activity. We plan extensive spectroscopic studies of the cyanide
inhibition to distinguish the apparent two Ni-sites.

Formate dehydrogenase remains to be sequenced. This enzyme
containing tungsten, selenium, and iron-sulfur centers is extremely
oxygen sensitive and is found in very low concentration in cells
grown on glucose. Recently Dr. T. Imai (Visiting Professor from
Japan) has done extensive studies in our laboratory and found that
the enzyme is induced when C. thermoaeticum is grown on methanol
and that its activity is stabilized by the presence of tungstate
during the purification. We feel we are now in the position to
partially sequence the enzyme to obtain AA-sequences to construct
some degenerate DNA nucleotide probes. Cloning and DNA sequencing
is planned using PCR methods.

Anaerobic Rumen Fungi. I developed an interest in anaerobic
fungi in 1982 when asked to write a review on the ecology of
cellulose hydrolysis. The literature search for the review
revealed that 3 isolates of strictly anaerobic fungi had been
reported. Only a few studies had been made with them. They were
isolated from rumens of herbivores and had been shown to produce
cellulases. Some of these studies had originated by Dr. D. Akin at
Russell Research Center in Athens, Georgia. Subsequently, D. Akin
and I together with R.D. Hartley and Scott Borneman (graduate
student) started work which led to the isolation of 5 different
anaerobic fungi from rumen of a cow. (Borneman et al. Appl.
Environ. Microbiol 55, 1066, 1989). They have been identified and
named Neocallimastix sp. MC-1, MC-2, and MC-3 (all monocentric),
and Orpinomyces PC-1 and PC-2 (both polycentric) (Wubah and Fuller,
Department of Botany, UGA, personal communication).

Physiological studies demonstrated that the fungi degraded
lignocellulosic tissues such as grasses and leaves. Enzyme studies
demonstrated the presence of cellulolytic, xylanolytic and other
polysaccharide degrading enzymes. The xylanolytic activity was
particularly high as was esterase activity. We have chosen to
study two of the fungal isolates, Neocallimastix MC-2 and
Orpinomyces PC-2.

Recent works include the isolation of a β-glucosidase and 4
different xylanases from Orpinomyces PC-2, and one p-coumaroyl
esterase and 2 different feruloyl esterases from Neocallimastix
MC-2. The work on the esterases involved the isolation and
characterization of specific substrate FAXX and PAXX from Coastal
Bermuda grass (CBG) (Borneman et al. Analytical Biochem. 190, 129,
1990). FAXX is O-[5-O(trans-feruloyl)-α-L-arabinofuranosyl]-(1→3)O-β-D-xylopyranosyl-(1→4)-D-xylopyranase and PAXX is the similar
compound but with the feruloyl moiety substantiated with a p-
coumaroyl moiety (see Fig. 1 for structures).
Of importance is the observation that the esterases when alone have very little activity toward CBG cell walls, although they rapidly hydrolyzed the purified substrates. Similarly the purified xylanases have little activity toward CBG. However, by combining the esterases with xylanases CBG is rather rapidly attacked as evidenced by the release of feruloylate or p-coumaroylate and xylose. We postulate that the enzymes act synergistically. First xylanases may produce esterified oligosaccharides, which then must be de-esterified before further hydrolysis to simple sugars occurs.

Our present goals are to purify several plant cell wall degrading enzymes from the anaerobic fungi, to study the relations of these enzymes with equivalent enzymes from aerobic fungi and anaerobic bacteria, to study the apparent synergism between purified enzymes involved in degrading plant tissues and to determine the primary structures of these enzymes for elucidating enzymatic mechanisms.

Fig. 1 Structures of PAXX and FAXX isolated from Coastal Bermuda Grass.
Ljungdahl, Lars G.

Publications since 1990


Abstracts


PROGRESS REPORT FOR THE D.O.E. GRANT
L.E. MORTENSON

GENETICS AND BIOCHEMISTRY OF THE AZOTOBACTER VINELANDII HOXC
(HYDROGENASE PROCESSING AND MATURATION) DNA REGION


1.1 We will study *Azotobacter vinelandii* cassette (1) insertion mutants of the putative HoxC (hydrogenase processing and maturation) region of DNA. Analysis of mutants for each individual open reading frame (ORF), ORF10 through ORF15, and deletion, cassette-replacement mutants for most of the same region should reveal whether or not each individual ORF, as well as the entire region of DNA, affects (a) hydrogenase activity and expression, (b) the rate of dihydrogen (H2) utilization, and (c) the synthesis of hydrogenase. In addition, these studies will establish whether each individual gene is transcriptionally active under given growth conditions (1). Techniques used will include western blots using antibodies against hydrogenase, northern blots, hydrogenase assays using the hydrogen electrode (with O2 and with methylene blue as electron acceptors), and tritium-exchange, mass spectroscopy, EPR, Ni incorporation and beta-galactosidase assays as a measure of expression. Since H2 is a by-product of dinitrogen (N2) fixation, and active hydrogenase is present under N2-fixing conditions (2, see our last year's report), we intend to run these tests by growing wild type and mutant cells under either N2-fixing or -nonfixing conditions. The roles of the HoxC genes and/or their expression under the two conditions will be compared.

1.2. Various parts of HoxC DNA will be introduced into the mutants in trans in order to determine the genetic organization (how many transcriptional units or operons exist) within this region of DNA. We also plan to define the boundaries of each individual operon using S1 mapping and additional mutagenesis and sequencing if needed.

1.3. We will identify and purify the gene products of HoxC ORFs generated using a T7 RNA polymerase expression system (3). Antibodies against the products will be made when high purity is established and then the products will be characterized biochemically.

1.4. We are in the process of obtaining the crystal structure of the clostridial bidirectional hydrogenase and the clostridial and azotobacter MoFe protein (nitrogenase). This is a collaborative study with Professor Jeff Bolin of Purdue University. We have already collected diffraction data on crystals of these proteins and expect to have the backbone of the MoFe protein shortly.

We also plan to obtain the relative positions of the Fe centers of the bidirectional hydrogenase during this next period. We have postulated five centers based on cluster extraction, Mossbauer, MCD, Resonance Raman Spectroscopy, electrochemistry and EPR analyses. Three of these centers are "typical" Fd-like centers, a forth appears to be a 2Fe-2S center and the fifth the H center where H2 is activated. Knowledge of the spacial relationships of these centers to one another will greatly enhance future research aimed at a complete structural and catalytic understanding of this and other hydrogenases.

2. Work done during the past year
2.1. We reported last year the identification of 13 complete ORFs and 1 incomplete ORF, ORF1 through ORF13 and ORF14', respectively. A gap of approximately 1.0 kb DNA, the only sequence gap between ORF1 and ORF14', between ORF9 and ORF10 was
subsequently determined and an ORF (new ORF10) was identified. In order to complete the sequence of (new) ORF15, cloning of the 2.7 kb BamHI DNA fragment from pALM21 (4) into pGEM-7Zf(+) (Promega, Co.) was carried out (part of this DNA fragment hybridizes to the *Alcaligenes eutrophus* HoxC (3) DNA probe, Menon, A.L., unpublished data). The resulting plasmids, pJCC143 and pJCC146, carried opposite orientations of the fragment with respect to the pGEM-7Zf(+) vector.

2.2. Sequencing of pJCC143 and pJCC146 using synthetic oligonucleotides and sequenase (U.S. Biochemicals Co.) methods gave us the remaining sequence of ORF15 ('ORF15) followed by another incomplete ORF (ORF16'). These ORFs were identified using computer programs reported earlier (4). The deduced amino acid sequences of (new) ORF10, 'ORF15 and ORF16' are shown in Figure 1.

2.3. Using the FASTA program of UWGCG software (6), we also identified that the deduced amino acid sequences of ORFs 11, 12, 14, 15, and 16' shared high homology with those of the hypA, B, C, D, and E genes of *Escherichia coli*, respectively. The latter five genes are required for hydrogenase activity in *E. coli* (7). Sequence data for ORF8 through ORF15, unique information for the DNA and deduced amino acid sequences, and a summary of homology comparisons with other genes required for H2 metabolism are in press in *Biochimica et Biophysica Acta* (8, 9; ORF1 in the paper is ORF8 in this report, numbering of the rest of ORFs follows the same order). The HoxC-hybridizing region (7.4 kb) begins at a *BglII* site in ORF8 and ends at an *EcoRI* site in ORF16' (Menon, A.L. unpublished data, refs. 4, 8, 9 and ORF16' data in this report. In conclusion, a continuous 13.6+ kb of *A. vinelandii* DNA has been sequenced and most of the DNA (genes) has been shown to be required for H2'ase synthesis or H2 metabolism (4, 8, 9, 10 and data of this report).

2.4. Preparation of cassette (1) insertion mutants of *A. vinelandii* for each individual ORF, ORF10 through ORF15, as well as deletion, cassette-replacement mutants of the same organism in which most of ORF10, the entire ORF11 through ORF14, and part of ORF15 were removed and replaced by the cassette, were completed via transformation. These mutants were kanamycin resistant as expected. Southern analysis using a non-radioactive labelling technique (Boehringer Mannheim Biochemicals Co.) verified the validities of the mutants with respect to the sites of cassette insertion or replacement into the chromosome.

2.5. The DNAs of ORF11 and ORF12 were amplified separately using a polymerase chain reaction protocol (Perkin Elmer Cetus Co.). The amplified DNA was then cloned into plasmid pT7-7 (3). Plasmid DNA sequencing of several clones of the *E. coli* strain BL21(DE3) (Novagen Co.) identified one clone for each ORF which had the correct deduced amino acid sequence in the coding region of the ORF. The expressions of the gene products of ORF11 and ORF12 were carried out in the same *E. coli* strain after IPTG induction. It was estimated that 3% to 10% of the total cellular proteins after induction were the ORFs' gene products, and the estimated molecular weight of the products on SDS gels were similar to the weights determined from the deduced amino acid sequence (9), i.e., 12.6 kDa and 33.2 kDa for ORF11 and ORF12, respectively.

2.6. We have established that the MoFe protein tetramer consists of two identical dimers each representing a nitrogenase active site. An active unit is an αβ dimer containing one 8Fe center (not two P clusters as originally thought) and a MoFe center. The two centers are 19-20 A apart and this two center unit is separated from a second identical unit on an αβ dimer by 70 A.
Good diffracting crystals of the bidirectional H2'ase of *Clostridium pasteurianum* have been obtained and diffraction data at varying wave lengths of X rays has been collected. This may be the first hydrogenase whose crystals allow a three dimensional solution since crystals of three other H2'ases did not diffract well. Analysis of this data is in progress and is part of our planned research for 1992/1993.

References


8. Chen, J.C. and Mortenson, L.E.: Two open reading frames (ORFs) identified near the hydrogenase structural genes in *Azotobacter vinelandii*, the first ORF may encode for a polypeptide similar to rubredoxins. *Biochimica et Biophysica Acta* (1992) in press.


Figure 1. Deduced amino acid sequence of (new) ORF10, 'ORF15 and ORF16' of the *Azotobacter vinelandii* DNA. Capital letters are deduced amino acid sequences determined from two strands (confirmed) DNA sequences and small letters are from single strand DNA sequences. The three possible initiation codons (Met) of ORF16' are underlined.

(New) ORF10

LAGRLHVEVRLQDGVIRA VDTRLQRPLPOISRLLVGGQTAEEAALRRLPLLFL GLCAAAQQVAALRALERAAGWAIAIÆEVEEGRTLGELESIRESLRLTVQQ WELPVPLERLKKALLALCRRAAARLQLALTAFRAGLPDAELEGTLAALAA AWADLQQPAPADWRPLDRWQEVALGGPPPPQAFTADELPALLAQLRASD ARAEIAGQPRLLGGPAASAGATAQAASQIEQHVGLALLRRTAQADLSLQSSP APPAVAGLAAGEGVGLARTARGWLLLHRVCLDDQAVGTWQL LAPTDWNFHADGPLRRRLCGVRVAAGEVEALLRELILALDPCVAFEVKIVHA

'ORF15

'REAFAEFDAERRFALEARSGLENKACECPAILRGAKSPRDCKLFGNCPSP DNPLGSCMVSSEGAACAWAYAYGRQRAVVAVAR

ORF16'

MSRRLDRNGSVEMVHGSGGRAGMQLIEELFARALRNEWLDQRNDQAQFEL PPGRVVMATDSSHVISPLFFPGDHSALAVHTINDVAMAGARPCYLAAGF ILEEGFPPLADRIVESAAMAREAGVPVTGDKVVEHKGDGVFITT TGVGVPPLGLHSQDRQARPDRILLSGSIGHGVVI SERIALREGLEFGFEADIGS D SQ ALHGLVAM LA AVEIRCRDPTRRGGLGNT L NELArqs gvgmqlver aipvrepvra aceflgldpl yvanegklia icpa'
Mortenson, Leonard E.

Publications since 1990


Wiegel, Juergen

Publications since 1990


Abstract


Invited Seminars