APPENDIX

LIDSTROM DEFG03-87ER13753

PROGRESS REPORT + 5 REPRINTS
7/1/92 / 6/30/95
1993 Annual Reports
94 95

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
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.
SUMMARY

The long-term thrust of this DOE project has been to identify and characterize mox genes and other methylotrophy genes in both methane- and methanol-utilizing bacteria, and study expression of those genes. In the last three years of support, the project has focused on identifying methylotrophy genes and the regions involved in their expression for comparative purposes, and has begun the process of analyzing the genes involved in transcriptional regulation of the Mox system in the strain for which we have the most information, M. extorquens AM1.

In order to carry out comparative studies of the transcription of methylotrophy genes, we have cloned and characterized genes involved in methanol oxidation (mox genes) from two Type I methanotrophs, Methylobacter marinus A45 (formerly Methylomonas sp. A45) and Methylobacter albus BG8 (formerly Methylomonas albus BG8). In both cases, the organization of the genes was found to be identical, and the transcriptional start sites upstream of the moxF genes were mapped. Other methylotrophy genes have been cloned and characterized from these methanotrophs, including mxaAKL and fdh.

The rest of this project has focused on the regulatory network for the mox system in M. extorquens AM1. We have sequenced two mox regulatory genes, mxbD and mxbM and they show identity with a specific group of sensor-kinase/response regulator pair systems. We have shown that these genes are required for expression of mox genes and of pqq genes, and also for repression of a PQP-linked dehydrogenase of unknown function. A second set of mox regulatory genes, mxcQE also have similarity to sensor-kinase/response regulator pairs, but to a different subfamily than mxbDM. We have studied the regulation of pqq genes in more detail, and have shown that the major regulatory step is post-transcriptional. Finally, we have shown that an RNA polymerase preparation purified from methanol-grown M. extorquens AM1 contains two polypeptides with similarity to known sigma factors, and this preparation initiates transcription at the moxF transcription start site.

Progress

Each of the publications resulting from this project period are cited by author names in the text below, and are listed at the end of this Progress Report. The numbered references are found in the list at the end of the main proposal.

1. Promoter regions of methylotrophic genes
   a. Methanotrophs. At the beginning of this project period, we continued a series of studies to characterize promoter regions of genes involved in methylotrophic growth in both methane- and methanol-utilizing bacteria, for comparative purposes. We had previously cloned the region containing mxaf from two methanotrophs, Methylobacter albus BG8 and Methylobacter marinus A45. We first determined the order of the mox genes and found they were the same as in M. extorquens AM1 (mxafJGIR). Next, we mapped the transcriptional start sites for the two moxF genes. The regions immediately upstream of both transcriptional start sites did not show substantial similarity to each other or to the conserved -10, -35 sequence for moxF from Methylobacterium strains, nor to any recognizable promoters. Transcriptional fusions to a lacZ reporter gene confirmed the presence of promoter activity in the case of M. albus BG8. This work has been published (Waechter-Brulla et al., 1993; Chistoserdova et al., 1994). Another gene involved in methylotrophy encoding a formaldehyde dehydrogenase (fdh) was cloned and sequenced from M. marinus A45 (Speer et al., 1994) and a gene cluster involved in the Mox system (mxakKL) was also cloned from M. albus BG8 and partially sequenced (Arps et al., 1995). An analysis of the sequence upstream of these genes did not show resemblance to the Methylobacterium moxF promoter, to the regions upstream of the methanotroph moxF genes, nor to any of the known promoters noted above. This effort has resulted in the identification of a number of methylotrophic genes from methanotrophs, and has shown that the promoter sequences for methylotrophic genes in methanotrophs and methylotrophs must be different. Since promoter mutation studies are difficult to carry out in methanotrophs, we have focused our efforts for promoter characterization and genetic regulation on genes from M. extorquens AM1, a strain more amenable to genetic analysis.
b. *M. extorquens AM1*. We had previously mapped the transcription start site for *mxaf* from *M. extorquens AM1* (11). We cloned and characterized a number of genes involved in the serine cycle (assimilatory metabolism) from this organism (Arps et al., 1993) and generated *xyle* transcriptional fusions to two of these (*hprA*, encoding hydroxypyruvate reductase, and *sga4*, encoding serine-glyoxylate aminotransferase). These enzymes are induced about 3-fold during growth on C1 compounds (30). An analysis of upstream sequences did not reveal any recognizable promoter sequences.

The next system we studied was that for PQQ synthesis. We focused on the first gene of a five-gene cluster (*ppqDGCBA*; see Fig. 3, main proposal), which turned out to be highly transcribed. We mapped the transcriptional start site and detected two transcripts that initiated at that site, one encoding only *ppqD* and a second encoding *ppqDG* (Ramamoorthi and Lidstrom, 1995). We were unable to detect a transcript for the entire 5-gene cluster, but it may have been present at too low a level for us to detect. Once again, the -10, -35 region of *ppqD* did not show similarity to these regions from *mxaf*, but it did show significant similarity to the -10, -35 sequences of the gene equivalent to *ppqD* from *Klebsiella pneumoniae* (Ramamoorthi and Lidstrom, 1995). *K. pneumoniae* does not grow on methanol, but it has a PQQ-linked glucose dehydrogenase and therefore, has *ppq* genes (21). We have also sequenced two other *ppq* genes (*ppqEF*), unlinked to the first gene cluster, and have shown that *ppqE* is equivalent to *ppqF* of *K. pneumoniae* (Springer et al., in prep.; see Fig. 3, main proposal). Both of these genes appear to encode a protease, but its role in PQQ synthesis is still unknown.

2. *M. extorquens AM1* RNA polymerase.

The data we have obtained so far suggest that methylotrophic genes in *M. extorquens AM1* may be transcribed using a novel sigma factor, since we have not identified promoter sequences indicative of known promoters. In order to assess this question, we have purified RNA polymerase from *M. extorquens AM1* cells grown on succinate or on methanol, and in both cases, have obtained a preparation that has 6 polypeptides. Three of these are the correct size to be α, β, and β', and in the case of the putative α polypeptide, N-terminal sequencing of the polypeptide after blotting to a nylon membrane has revealed high identity to *E. coli* α subunit of RNA polymerase. The putative β and β' subunits have not been sequenced. The other three polypeptides are 97, 40 and 24 kDa in mass. These have been blotted to nylon membranes and N-terminal sequences have been determined. The sequence from the 24 kDa polypeptide has 50% identity to the N-terminal sequence of the 24 kDa α polypeptide found in RNA polymerase preparations of *E. coli* (Fig. A). The sequence from the 97 kDa polypeptide has 58% identity to the N-terminal sequence of the 77 kDa vegetative sigma factor from *Agrobacterium tumefaciens* (Fig. A), which keys into the large class of RpoD/SigA sigma factors (31). The sequence of the 40 kDa polypeptide has 33% identity to the Group 2 sigma factor from *Streptomyces coelicolor*, HrdD (Fig. A), including 10 residues that are highly conserved in this group (31). Neither of the putative sigma factors was detected in immunoblots using antibody against *E. coli* σ70.

In vitro runoff transcription assays were carried out to determine whether this RNA polymerase preparation initiates transcription from *mxaf* at the same start site measured using bulk RNA. The template used was a 392 bp fragment generated by PCR that contains the *mxaf* transcription start site within it, and is predicted to produce an RNA of 133 nucleotides. A major RNA species was observed of approximately this size (between 123 and 147 nucleotides), and this species was not produced with a control DNA fragment from upstream of *mxaf* that lacked the *mxaf* promoter. These data suggest that this RNA polymerase preparation does initiate transcription at the correct site. This work is currently being written up for submission to *J. Bacteriol.* (Davagnino et al., in prep.) The sequence comparisons suggest that the 97 kDa polypeptide may be the σ70 equivalent, and the 40 kDa polypeptide may be the sigma factor responsible for transcribing *mxaf*, but further work will be necessary to define the role of each.
A low level methanol-inducible activity was observed when the fusion was in the incorrect orientation (Fig. B), as was reported for vectors described earlier, pGD590. The transcriptional start site for mxaF, mxaG and mxaI polypeptides (Fig. B) were studied in wild-type and in the mxaZ mutant. This work has involved two promoter fusion vectors described earlier, pGD500 (lacz reporter) and pHX200 (xylE reporter). With pGD500, mxaF-lacz fusions were generated using a 1.55 kb M. extorquens AM1 fragment, which contained the transcriptional start site for mxaF, about 1 kb of upstream DNA, and a small portion of mxaF (Fig. B). The fusions (both orientations) were studied in wild-type and in the mxaB mutant. This work was supported by a previous DOE funding period and has been published (11), but it is summarized here for completeness. B-galactosidase activity was at a high constitutive level in wild-type, and not detectable in the mutant. A chromosomal construction, pCM301 that regenerates a complete mxaF in the chromosome (see Fig. B), was generated, and in this case, the regulation of B-galactosidase activity correlated well with the regulation of the MxaF, MxaG and MxaI polypeptides and more generally with MedH activity (11).

We have subsequently used the same 1.55 kb promoter-containing fragment to generate a mxaF-xylE fusion in pHX200, and in this case, the expected regulation was observed in the plasmid construction (Fig. B), as was reported for M. organophilum XX promoter fusions using this vector (12). The background activity for pHX200 alone in M. extorquens AM1 is very low, often not detectable. A low level methanol-inducible activity was observed when the fusion was in the incorrect orientation (Fig. B), as was reported for vectors described earlier, pGD590. The transcriptional start site for mxaF, mxaG and mxaI polypeptides (Fig. B) were studied in wild-type and in the mxaZ mutant. This work has involved two promoter fusion vectors described earlier, pGD500 (lacz reporter) and pHX200 (xylE reporter). With pGD500, mxaF-lacz fusions were generated using a 1.55 kb M. extorquens AM1 fragment, which contained the transcriptional start site for mxaF, about 1 kb of upstream DNA, and a small portion of mxaF (Fig. B). The fusions (both orientations) were studied in wild-type and in the mxaB mutant. This work was supported by a previous DOE funding period and has been published (11), but it is summarized here for completeness. B-galactosidase activity was at a high constitutive level in wild-type, and not detectable in the mutant. A chromosomal construction, pCM301 that regenerates a complete mxaF in the chromosome (see Fig. B), was generated, and in this case, the regulation of B-galactosidase activity correlated well with the regulation of the MxaF, MxaG and MxaI polypeptides and more generally with MedH activity (11).

We have subsequently used the same 1.55 kb promoter-containing fragment to generate a mxaF-xylE fusion in pHX200, and in this case, the expected regulation was observed in the plasmid construction (Fig. B), as was reported for M. organophilum XX promoter fusions using this vector (12).
orientation (pCMSau4), suggesting the possibility of a divergently transcribed low-level promoter (Springer et al., submitted), which is probably analogous to the mxaW promoter studied in M. organophilum XX (12). An analysis of this region in M. extorquens AM1 has revealed that it contains a potential partial open reading frame of 220 amino acids, transcribed in the direction opposite to mxaF that shows good correlation with the expected codon usage of M. extorquens AM1. The N-terminal 34 amino acid overlap with the putative mxaW from M. organopolium XX shows 52% identity, but neither sequence shows substantial similarity to any entries in GenBank. We have not yet constructed a mutant in this putative gene, and so its potential function is unknown.

An analysis of this region in M. extorquens AM1 has revealed that it contains a potential partial open reading frame of 220 amino acids, transcribed in the direction opposite to mxaF that shows good correlation with the expected codon usage of M. extorquens AM1. The N-terminal 34 amino acid overlap with the putative mxaW from M. organopolium XX shows 52% identity, but neither sequence shows substantial similarity to any entries in GenBank. We have not yet constructed a mutant in this putative gene, and so its potential function is unknown.

Figure B. The M. extorquens AM1 mxaF promoter fusions and activities from them. The fragment cloned is shown above, and the two xyle fusions (both orientations) are shown next. The mxaF-lacZ chromosomal insertion construction is shown at the bottom.

For the mxaF-xyle fusion, the increase in catechol dioxygenase activity measured in cells grown on methanol and succinate (10-14 fold) is similar to that observed for MxaF levels, as measured by immunoblots, suggesting that this system faithfully reports transcription from this promoter (Springer et al., submitted). We have used this fusion to assess transcription from the M. extorquens AM1 mxaF promoter in each of the 6 regulatory mutant classes noted above. In all cases, no detectable catechol dioxygenase activity (above the very low vector background) was found, regardless of growth substrate (succinate, methylamine, or methanol + methylamine) (Springer et al., submitted). These results confirm with a M. extorquens AM1 promoter fusion the phenotype of the MxbB, MxbD, and MxcE mutant classes as measured by Xu et al. using M. organophilum XX promoter fusions and cells grown on C1 substrates (12). This also extends the information to the MxbM, MxcQ and MxcB mutant classes and a broader range of growth substrates. We have not yet tested the expression of the putative mxaW promoter in these mutant classes. However, the mxaW promoter of M. organophilum XX was not expressed in the M. extorquens AM1 MxbD mutant class, although it was expressed in the MxbA and MxcE mutant classes (12).

We have also studied regulation of PQQ synthesis and related this to transcription of pqqD, which encodes the peptide thought to be the precursor of PQQ. This work is described in a paper published in J. Bacteriol. (Ramamoorthi and Lidstrom, 1995). A single transcriptional start site for pqqD was mapped, and RNA blots identified two transcripts, a major one that encoded only pqqD, and a minor one that encoded both pqqD and the next gene, pqqG (Fig. C). PQQ levels were found to be about 5-fold higher in methanol-grown M. extorquens AM1 than in succinate-grown cells. However, both transcription of pqqD (as measured using a pqqD-xyle fusion; see Fig. C) and the steady-state levels of the pqqD transcripts were the same under both growth conditions. Therefore, the methanol regulation of PQQ synthesis must occur either at the level of transcription of different pqq genes, or at the post-transcriptional level. However, some regulation of pqqD does occur, since a 2-3 fold induction of pqqD transcription was observed in cells grown on methanol plus methylamine, but not in cells grown on single substrates (methylamine, methanol or succinate). This induction was
abolished in the \textit{mox} regulatory mutant classes MxbM, MxbD and MxaB, and in addition, activity from the \textit{pqqD-xyIE} fusion was decreased 5-10 fold in MxbM and MxbD mutants, but not affected significantly in any of the other regulatory mutants (Ramamoorthi and Lidstrom, 1995). These data are suggestive that MxbM and MxbD are involved either directly or indirectly in transcription of \textit{pqqD}, and that \textit{mxaB} mediates the induction of \textit{pqqD} by \textit{C}_1 compounds. However, MxcQE do not seem to be involved in transcription or transcriptional regulation by \textit{C}_1 compounds of \textit{pqqD}.

The \textit{mox} regulatory genes also affect a third system. We had previously shown that MxbD mutants contained very high levels of a cytochrome that is normally present in only trace amounts, which was later shown to be cytochrome \textit{c-553} (14). This cytochrome is now thought to be the electron acceptor for an unknown PQQ-linked dehydrogenase that has high identity with methanol dehydrogenase (23). We have now cloned and sequenced the genes for cytochrome \textit{c-553} and the dehydrogenase (\textit{mxaF}'\textit{G}') from \textit{M. extorquens} AM1 (see Fig. 3, main proposal), and have generated mutants defective in these genes by allelic exchange. As in \textit{P. denitrificans} (23), mutants in these genes grow normally on methanol and all other tested substrates, and the role of this dehydrogenase is still unknown. However, we have shown by immunoblot that cytochrome \textit{c-553} is derepressed in all six of the regulatory mutant classes in cells grown on methylamine + methanol, but not in cells grown on succinate (Springer et al., submitted). Therefore, the derepression phenomenon is dependent upon the presence of methanol. Although we do not know the function of this dehydrogenase, it is possible that it is a detoxifying, periplasmic formaldehyde dehydrogenase, which might be expected to be regulated in the opposite manner to the formaldehyde-producing system (methanol dehydrogenase) in response to the presence of methanol. We are attempting to determine whether or not this is so.

In summary, the data available so far suggest that in \textit{M. extorquens} AM1, \textit{mxbD} and \textit{mxbM} encode a sensor kinase/response regulator system that is involved in the expression of several sets of genes (Fig. D). These include positive regulation of several \textit{mox} genes including at least one more sensor kinase/response regulator pair and a third linked gene (\textit{mxcBQE}), as well as \textit{pqqD} and \textit{mxaW} \textit{mxcBQE} in turn are involved in positive regulation of \textit{mxaF} and negative regulation of the genes for a PQQ-linked dehydrogenase of unknown function, \textit{mxaF}'\textit{G}'\textit{J}'\textit{G}'. It is not yet known whether \textit{mxaB} expression is dependent upon \textit{mxbDM}, but current data suggest it is not dependent upon \textit{mxcBQE}. \textit{mxaB} apparently works with the \textit{mxcBQE} system in the regulation of \textit{mxaF} and \textit{mxaF}'\textit{G}'\textit{J}'. One of the goals of this project is to determine which of the regulatory genes is required for expression of the others, in order to determine the hierarchy of expression in this system.

This regulatory system is a little unusual, in that the products of all six genes are required for any detectable expression of \textit{mxaF}, regardless of the presence or absence of inducer. Therefore, at least one of them must encode a protein (for convenience, I will call it \textit{R}_{mox} ) that is absolutely required for transcription of \textit{mxaF}. \textit{R}_{mox} may be a DNA-binding activator protein, it may be a
protein that binds to RNA polymerase and changes its activity (a sigma factor or other binding protein), it may be a protein that acts with other mox gene products in a complex to either bind DNA and/or bind RNA polymerase, or it may be an anti-sigma factor or a protein that inactivates a repressor protein. The other 5 regulatory genes may be involved in generating the active form of Rmox, either by transcription of required genes or direct activation of Rmox, and/or be required with Rmox to generate an active complex. In addition, one or more of these genes may be required for the induction phenomenon, or it is possible the induction may be mediated by a separate set of genes. Data from Xu et al., (12) show that transcription of the mxaW promoter from M. organophilum XX does not require mxE or mxB in M. extorquens AM1 and we have shown that transcription of the pqd promoter does not require mxBQE or mxB. However, induction of the pqD promoter by C1 compounds does require mxB. It is not yet known whether induction of the mxaW promoter requires mxB. We also do not know whether intermediary regulators exist in the pathway between mxbDM and pqD or mxaW, but they may exist, as indicated in Fig. D. It is also possible that another set of regulators exists in the pathway between mxBQEmxaB and the target genes, that has not yet been identified. One of the goals of this project is to obtain evidence as to whether other regulatory functions are operating in this system. What is the identity of Rmox? It could be the product of mxE, which is predicted to encode a DNA-binding protein, the product of mxB, or the product of an unknown gene. The experiments proposed in this project will provide clues as to the identity of Rmox, but definitive identification will probably require protein purification and in vitro studies.

**Figure D.** Summary of the data involving Mox regulatory genes in M. extorquens AM1. +, activation; -, repression; I, required for induction only; *, expression decreases 5-10 fold (for all other positive regulation, expression is not detectable in the absence of the required gene product).

Substantial MeDH activity (0.02 µmol/min/mg protein) and activity from the mxaF-xyIE fusion (0.25 µmol/min/mg protein) are found in cells grown on succinate, in the absence of any external inducer. Since the products of all six regulatory genes are required for any detectable transcription of the mxaF promoter, these data indicate that these gene products must be present in an active state at some level in the absence of inducer. How might induction work in such a system? Two main possibilities exist, depending on whether the lower transcription in the absence of inducer is due to low levels of Rmox protein and/or other regulatory proteins, or whether it is due to presence of Rmox in a less active state. If the level of Rmox is limiting in the absence of the inducer, then the inducing signal might be sensed and transduced by one of the sensor kinase/response regulatory systems, with the result being increased levels of Rmox. Alternatively, if the steady-state levels of Rmox are sufficient for induction, the signal may be involved in generating a more active form of existing molecules of Rmox. This could occur through a phosphorylation cascade, in which case the signal would probably be sensed by one of the sensor kinase/response regulatory systems. Alternatively, Rmox itself might bind the inducer, creating a more active form. That scenario leaves open the
question of what the sensor kinases sense, but it is possible that they are involved in more subtle regulation, such as coordination of formaldehyde-producing and formaldehyde-consuming systems. In that case they might sense a ratio of, for instance, formaldehyde/methylene THF or methylene THF/formate. It is also possible that a combination of these mechanisms (increased levels of $R_{\text{max}}$ plus activation of $R_{\text{max}}$) is operating. A final possibility, noted above, is that another set of unidentified regulatory genes is involved in induction. We do know however, that $mxaB$ is involved in induction of $pqqD$. One of the goals of the new project period is to obtain information on the mechanism of the induction process and the inducing signal.

Publications from this project (starred papers are included in this Appendix):

Research Papers


Book Chapters
