2

DOE/PC/89902--T14

DOE/PC/89902--T14

DE93 005814

FINAL TECHNICAL REPORT

on

MOLECULAR BIOLOGICAL ENHANCEMENT OF COAL BIODESULFURIZATION

Contract No. DE-AC-22-89PC89902

to

U. S. DEPARTMENT OF ENERGY
PITTSBURGH ENERGY TECHNOLOGY CENTER

by

J.H. Litchfield, T.J. Zupancic, J.D. Kittle Jr.,
B. Baker, D.T. Palmer, C.G. Traunero, R.E. Wyza,
A. Schweitzer, and H.N. Conkle

(Battelle)

and

L. Chakravarty and O.H. Tuovinen (The Ohio State University)

October 8, 1992

U.S. DOE Patent Clearance
is <u>not</u> required prior to publication of this document

BATTELLE
505 King Avenue
Columbus, Ohio 43201-2693

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLES

FIGURES

FIGURES
(Continued)

EXECUTIVE SUMMARY

The purpose of this project, "Molecular Biological Enhancement of Coal Biodesulfurization" (Contract Number DE-AC-22-89PC89902), was to develop improved technology for the biodesulfurization of coal. The ultimate goal was to combine genes for both organic and inorganic (pyritic) sulfur utilization into one organism. To pursue this goal, research activities in several different areas were performed. These areas included: 1) characterization of heterotrophic bacteria potentially capable of releasing sulfur from organosulfur compounds while leaving the carbon skeleton intact; 2) characterization of autotrophs capable of metabolizing inorganic sulfur as potential host organisms (*Thiobacillus ferrooxidans*); and 3) development of genetic engineering technology for introduction of heterologous genes into *T. ferrooxidans*.

While organisms such as *T. ferrooxidans* were known to be capable of removing pyritic sulfur from coal, limited information was available concerning organisms capable of removing organic sulfur. In order to define the metabolic capabilities of candidate heterotrophic organisms, methods were needed to evaluate their metabolism of the model organic sulfur compound, dibenzothiophene (DBT). The Lehigh plate fluorescence assay was established in our laboratory to screen strains for DBT metabolism, and thin layer chromatography (TLC) and high performance liquid chromatography techniques (HPLC) were developed for detailed analysis of selected organisms.

Heterotrophic organisms from Lehigh University, which were reported to contain at least part of the desired DBT metabolic pathway, were obtained and analyzed for their ability to metabolize DBT. The results indicated that the Lehigh strains were quite variable with respect to DBT metabolism. Degradation of DBT to the end product o,o'-biphenol could not be established. However, DBT sulfoxide and DBT sulfone were sometimes produced. The primary metabolites produced from DBT were found to be the so

called "Kodama" compounds, in which the carbon skeleton of DBT is cleaved.

Genetic engineering of *T. ferrooxidans* required technology for the introduction of foreign genes into these bacteria. The necessary elements for this technology included: identification of a suitable host strain; construction of a functional plasmid vector which can be selected and maintained in *T. ferrooxidans*; and, a method for the transformation of *T. ferrooxidans* cells with recombinant vectors. Significant progress was made in developing a plasmid vector system for transformation of *T. ferrooxidans*.

A series of *T. ferrooxidans* strains obtained from The Ohio State University were purified and analyzed to define their sensitivity to various antibiotics, heavy metals, organic sulfur compounds and coal cleaning agents. Through these analyses, potential selectable marker genes were identified for genetic transformation. The data also indicated that the viability of the strains was not adversely affected by organic sulfur compounds or coal cleaning agents.

To develop plasmid vectors, native plasmids from the *T. ferrooxidans* strains were isolated and characterized. A family of circular DNA plasmids ranging in size from 9,800 to 20,000 base pairs was identified. The smallest of these, PTFI91, was cleaved with restriction enzymes and joined to an *E. coli* cloning vector to create a potential shuttle vector. This vector as well as other broad host range plasmids were used in experiments to demonstrate electroporation of *T. ferrooxidans*. In these experiments no transformants were detected among the electroporated cells, probably because the proper combination of host/vector and selection system had not yet been developed. Data from these electroporation experiments suggested that DNA was entering the electroporated cells, and that a different choice of genetic marker and selection media might lead to stable transformation.

Southern hybridization analysis was used to identify a

segment of the cloned pTFI91 plasmid which was present in all of the *T. ferrooxidans* plasmids. Sequencing of this segment (a 2.2 kb SacI restriction fragment) resulted in identification of the plasmid's vegetative replication origin. Detailed sequence analysis showed that this plasmid replication origin was similar in general structure to broad host range plasmids such as RSF1010 and pSC101, but represented a novel class of DNA replication origin.

In addition to the plasmid origin, the sequenced segment also contained a new type of insertion sequence (IS), designated IS3091. This IS element was related to the previously reported sequences IS30 and Tn4551, and served to define a family of replicative IS elements. This insertion sequence may prove valuable as a tool for future engineering of recombinant *Thiobacillus*.

As part of this work, the mercury resistance genes of the *T. ferrooxidans* strain DSM583 were cloned by hybridization screening of a genomic library. The *T. ferrooxidans* mercury resistance genes are potentially valuable as genetic markers, since: 1) mercury resistance genes have been well characterized as genetic markers for the transformation of other species; and 2) mercury naturally contaminates some environments where recombinant *Thiobacillus* may be useful, so that enhanced resistance to mercury would be a desirable trait. The expression of the cloned *mer* genes in *E. coli* appears to be inducible by low levels of mercury, suggesting that the regulatory elements of the *mer* gene operon are present, intact and functional in at least some of these clones.

The most significant achievements of this project can be divided into two general areas: First, a knowledge of organic sulfur metabolism and its analysis was gained which helps further define the task of engineering a *Thiobacillus* capable of removing organic sulfur from coal. Second, the cloning of a *T. ferrooxidans* plasmid, the discovery of a novel plasmid origin, the identification of a new IS element and the cloning of a

functional *mer* gene operon all advanced the understanding of the molecular biology of *T. ferrooxidans* and define what is required for genetic engineering of this species.

FINAL TECHNICAL REPORT

on

MOLECULAR BIOLOGICAL ENHANCEMENT

OF COAL BIODESULFURIZATION

Contract No. DE-AC-22-89PC89902

to

U.S. Department of Energy

Pittsburgh Energy Technology Center

from

BATTELLE

October 8, 1992

1.0 INTRODUCTION

The U.S. Department of Energy (DOE) is investigating the microbially mediated release of sulfur from coal as a viable industrial process for coal desulfurization. The overall DOE goal is to develop an optimized microbiological process for removing both organic and inorganic sulfur from coal that is cost-efficient, operates under mild conditions and is simple to engineer.

Bacteria of the genera *Thiobacillus* and *Sulfolobus* are capable of removing inorganic sulfur from coal. Organic sulfur in coal exists in the forms of sulfides, thiophenes and thiols

and is comparatively more recalcitrant to removal by microorganisms. Research results reported by DOE contractors at the Fifth Annual Coal Preparation, Utilization and Environmental Control Contractors Conference, 1989, indicated that bacteria isolated from natural environments degrade model coal organic sulfur compounds such as dibenzothiophene (DBT) and phenylsulfide (Kilbane and Bielage, 1989; Wyza and Isbister, 1989). Mono- and dihydroxy-biphenyl and sulfate, the ultimate products of degradation of DBT by the proposed 4S pathway, were detected in these studies. The bacteria involved include members of the genera *Pseudomonas* and related organisms such as *Rhodococcus* and *Bacillus* (Kilbane, 1989). Also, a European patent application describes a novel mutant microorganism reported to be useful in removing organic sulfur from coal (Isbister and Doyle, 1985).

It would be desirable to have improved strains that would be capable of removing both inorganic and organic sulfur from coal. This program was directed toward achieving this goal.

The specific initial technical objectives of the project were to:

- Clone and characterize the genes encoding the enzymes of the "4S" Pathway (sulfoxide/sulfone/sulfonate/ sulfate) for release of organic sulfur from coal;

- Return multiple copies of these genes to the original host to enhance the biodesulfurization activity of that organism;

- Transfer this pathway into a fast-growing chemolithotrophic bacterium;

- Conduct a batch-mode optimization/analysis of scale-up variables.

This report summarizes the results of research

performed on this contract during the period beginning July 14, 1989. Bacterial strains which potentially contained the "4S" pathway were obtained from Lehigh University.

By the letter of September 3, 1991, from the Project Manager at Department of Energy, Pittsburgh Energy Technology Center, these objectives of the project were redirected toward finding and developing suitable vectors for *Thiobacillus* strains. All work on bacterial strains obtained from Lehigh University was terminated since these strains did not appear to contain desulfurization traits associated with the "4S" pathway.

2.0 DBT UTILIZATION BY CHEMOORGANOTROPHS

2.1 Analytical Methods to Monitor DBT Utilization

In order to achieve the primary initial objectives of the program, characterization of strains in terms of "4S" pathway activity was necessary. In this pathway, dibenzothiophene (DBT) is thought to be converted through a series of single oxidation steps to DBT sulfoxide, DBT sulfone, DBT sulfonate, and finally to o,o'-biphenol, with the release of sulfate (see Figure 1). A method was desired which could be used to rapidly screen organisms which may contain the "4S" pathway. A spray test method was available for screening and was developed further to define conditions which would yield optimal results (Dutt, et al. 1988; Krawiec, 1990). Other analytical methods were needed to confirm the presence of the "4S" pathway in the organisms. Both thin layer chromatography (TLC) and high performance liquid chromatography (HPLC) were developed for this purpose. This section provides a summary of the development of these methods.

2.1.1 4S Pathway Compounds

Dibenzothiophene was purchased from Aldrich Chemical Co., Eastman Kodak, and Lancaster Synthesis.

Dibenzothiophene sulfoxide was purchased from K&K Laboratories. Dibenzothiophene sulfone was purchased from Aldrich Chemical Co. The compound o,o-'biphenol was purchased from Aldrich Chemical Co., Lancaster Synthesis, and Sigma Chemical Co. It should be noted that o,o'biphenol is listed by suppliers as 2,2' dihydroxybiphenyl (Lancaster Synthesis), 2,2' biphenol (Aldrich), and o,o'-biphenol (2,2' dihydroxybiphenol) (Sigma). Dibenzothiophene sulfonate was not commercially available. Each chemical, except for DBT sulfoxide, yielded a single peak on HPLC chromatograms, indicating that the compounds were of high purity. DBT sulfoxide, however, was grossly contaminated with DBT sulfone regardless of the supplier.

2.1.2 Development of a Spray Test to Screen Organisms

A screening assay was developed by Dr. Steven Krawiec at Lehigh University for the presumptive detection of o,o'-biphenol, a possible end product of the putative "4S" pathway (Dutt el. al. 1988). In this assay, organisms to be tested for "4S" activity are grown out on agar medium. The resulting colonies are then sprayed with a DBT/ether solution. After a short incubation period (approximately 1-2 hours), the plate is viewed under a hand held mineral (short wave UV) light. Under excitation of short wave UV light, DBT and DBT sulfoxide do not fluoresce, DBT sulfone fluoresces dark purple, and o,o'-biphenol fluoresces a vivid purple. Thus, the fluorescing material indicates the possible presence of the product, o,o'-biphenol. Control plates (no DBT) are included to verify that the colonies do not themselves fluoresce.

Based on this assay, a method was developed for the screening of recombinants for DBT degradation to o,o'-biphenol. Details of the experiments were given in the 3rd Quarterly Report March 15, 1990. In summary, the limit of detection of o,o'-biphenol on an agar plate was determined to be between 0.05-0.5 μg in a 1 cm diameter spot. Best results were obtained using plates which were incubated for 6 days (colony

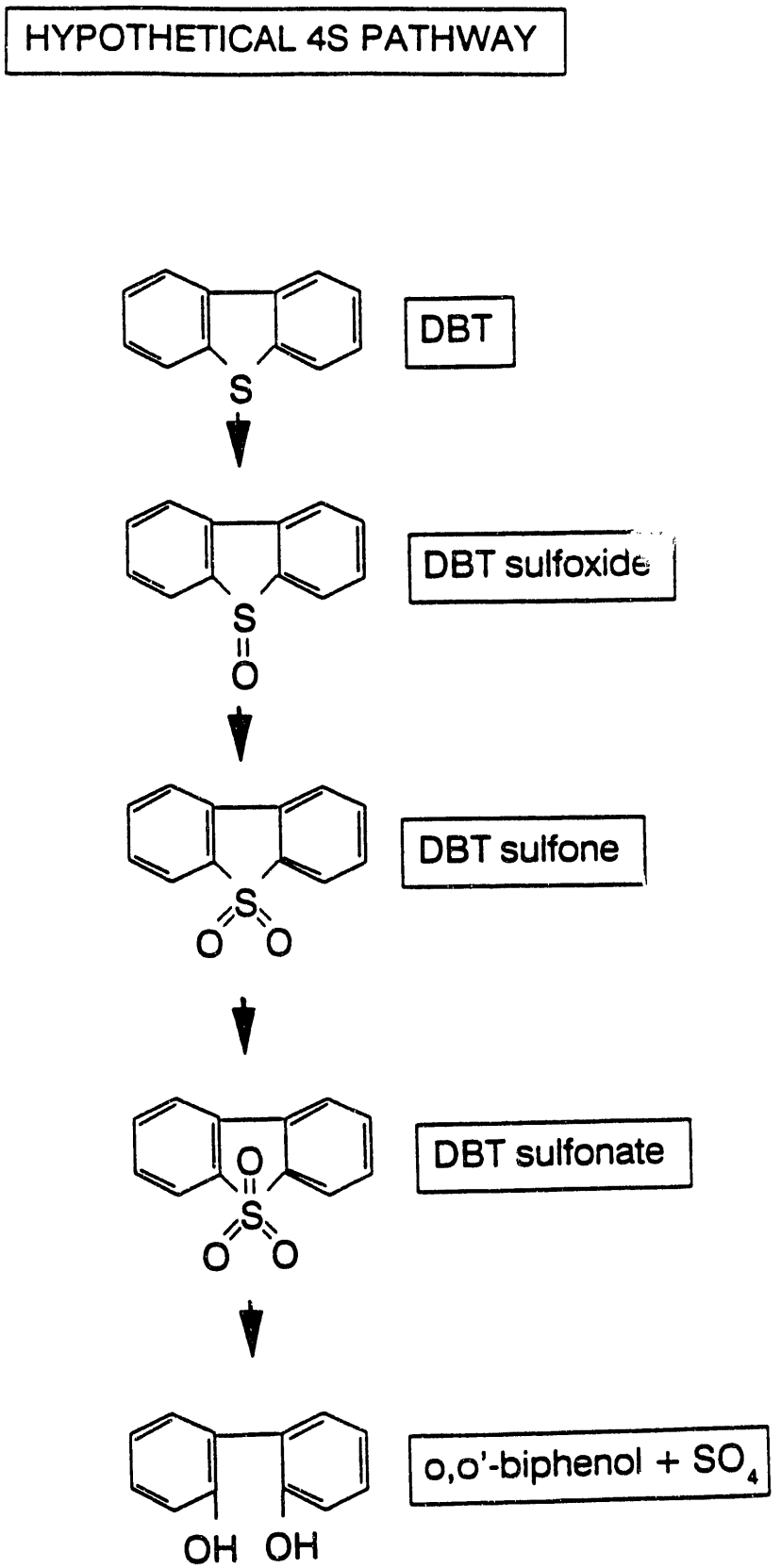


Figure 1

Degradation of Dibenzothiophene by the "4S" Pathway

diameter approximately 3-4 mm) and sprayed with a 0.15% (w/v) DBT/ether solution for 1 second. The spray rate was approximately 750 µl/sec, or 1.1 mg DBT/sec. Plates could be visualized within 15-60 minutes. Fluorescence was detected only when organisms were grown on glycerol plates. Additional colored compounds sometimes appeared depending on the strain, the carbon source, and the incubation time. Approximately 200 strains could be screened on a single plate using this assay.

2.1.3 Development of Thin Layer Chromatography

A thin layer chromatographic (TLC) assay system for detection of DBT metabolites (Brendlyn Faison, Oak Ridge National Laboratories, personal communication) was investigated. This assay separated all the "4S" metabolites on sheet-backed silica gel with a fluorescent indicator. Satisfactory resolution of DBT, DBT sulfoxide, DBT sulfone, o,o'-biphenol and 2-phenylphenol on silica gel was obtained using the hexane:acetone (5:1) mobile phase. Two additional solvent systems were evaluated for their ability to resolve the "4S" standard compounds. These systems, acetonitrile:hexane (1:4) and methylene chloride were not as suitable for the TLC analyses as the hexane-acetone system used by Oak Ridge National Laboratories.

Experiments using standards of the "4S" pathway indicated that the sample extraction procedure of Kodama et al. (1970) was compatible with TLC analysis of DBT metabolites. The intermediates separated by TLC were detected by using an adaptation of the Gibbs reagent (Gibbs, 1927). This visualizing agent reacted selectively with the phenolic standards used (o,o'-biphenol and 2-phenylphenol) and formed indophenol complexes. This complex formation resulted in the detection of blue spots on the chromatographs that were more easily visualized than the fluorescent indicator in the TLC silica (10 ng). Indophenol formation was stable, and the background color for this reagent

was low. The Gibbs reagent could also be used to detect the formation of Kodama intermediates, since most of them are also phenolic compounds (Kodama, *et al.*, 1970, 1973; Kodama, 1977, Monticello *et al.*, 1985).

The reactions of the Gibbs reagent with DBT sulfoxide, DBT sulfone, o,o'-biphenol, and 2-phenylphenol were also evaluated in alkaline buffer and these compounds were easily detectable at 100 ppm. The UV absorbance maximums of o,o'-biphenol and 2-phenylphenol indophenol complexes in aqueous solution were determined. These values were quite similar (311 nm and 317 nm, respectively), but could be distinguished from one another by the lack of reactivity of o,o'-biphenol with the Gibbs reagent in 100% ethanol.

2.1.4 Development of HPLC Analytical Technique

The spray test screening method described above was used to identify organisms appearing to have a complete "4S" pathway, that is, the metabolism of DBT to o,o'-biphenol. Further analysis of the fluorescing material was necessary to confirm the presence of o,o'biphenol. An HPLC technique was chosen for detailed analysis of the strains and in this program.

The HPLC setup was based on the use of a reverse-phase, isocratic solvent system developed at ARCTECH, Inc. The mobile phase, which was sparged with helium, consisted of 0.01 M phosphoric acid (46%), tetrahydrofuran (18%), acetonitrile (18%), and dioxane (18%). The column selected was a Dupont Zorbax ODS 4.6 mm x 25 cm C-18 analytical column, fitted with a Zorbax ODS 4mm x 1.25 cm guard column. Components of the HPLC system concluded a 10 μl injection loop, Beckman model 110A pump, Altex controller set at 1 to 1.2 ml/min, LCD/Milton Roy spectromonitor set at 242 nm, and a Hewlett Packard 3396A Integrator.

Excellent sensitivity and low baseline noise was achieved using a small (10 μl) injection loop. Baseline resolution of the "4S" pathway intermediates was achieved. All of

the "4S" compounds were identified by comparison of the retention times of the mixture and retention times of individual standards.

A gradient elution method was also developed to effect separation of metabolites which were poorly retained using the isocratic solvent system. Work by other investigators demonstrated that gradient systems were effective in analysis of "4S" compounds (Stoner et al., 1990). With our HPLC system, the more polar metabolites did not move appreciably from the solvent front when using the elution scheme of Stoner et al., 1990. The following elution scheme was tested: 30% acetonitrile/70% water for 2.5 minutes, followed by linear ramping to 100% acetonitrile over 10 minutes, holding at 100% for 8.5 minutes, and then returning to 30% acetonitrile/70% water. The flow rate was 1 ml/minute. This solvent program resulted in low solvent front response and good resolution of the 4S standards, except for o,o'-biphenol and o-phenylphenol, which were only partially resolved. Minor adjustments in the elution program would be needed to effect separation of these species. DBT sulfoxide eluted well after the solvent front, at approximately 11 minutes, as compared to about 3 minutes using the complex isocratic solvent system. Thus, relatively fast-moving metabolites could be better resolved using the gradient program. The late eluting species gave somewhat broad but acceptable peaks using this elution scheme; the height-to-width ratios did not appear to be improved significantly by the addition of 0.01% (final concentration) phosphoric acid to the water.

Sample preparation consisted of addition of acetonitrile to the culture to a final concentration of 60%. The samples were then vortexed and centrifuged to remove particulates. The supernatant containing unreacted DBT and any intermediates were then injected onto the HPLC column.

2.2 Characterization of Lehigh Strains

In this research effort, the overall objective was to transfer organic sulfur utilization traits present in chemoheterotrophs to an inorganic sulfur utilizing chemoautotroph such as *T. ferrooxidans*. Characterization of the donor chemoheterotrophs was necessary to provide baseline information on these organisms and to understand the nature of their metabolism with respect to the utilization of organic sulfur.

2.2.1 Sources of Microorganism and Basic Characterization

At the initiation of this program, four chemoheterotrophs (strains A4, A15, B24, and C18) reported to contain the 4S pathway were obtained from Dr. Steven Krawiec of Lehigh University. All of these strains were reported to show the capacity to degrade dibenzothiophene (DBT) by Krawiec's plate assay. These bacterial cultures were gram-negative motile rods all having reported optimum growth temperatures in the range 30 to 32°C and maximum growth temperatures of 34 to 35°C. The strain *Pseudomonas putida* (ATCC 33015) was used as a control. All strains were streaked on Stainer's agar to check for purity (Stainier et. al. 1966).

2.2.2 Biolog™ Characterization

Biolog™ (Biolog, Inc., 3447 Investment Blvd., Suite 3, Hayward, CA, 94545) plates were used to characterize the metabolic capability for the strains. The results of this test indicated that strains A4, A15, and B24 closely resembled the *P. putida* control. Strain C18 was distinctly different from the other strains. The results of the Biolog™ plate analysis of C18 grown on both trypticase soy broth (TSB) (Difco) and Stanier's medium showed that C18 utilized many of the carbon and energy sources to a moderate extent. The carbon sources that were

utilized to a significant degree using a Stanier's plus 0.2% glycerol medium are shown in Table 1.

The carbon source utilization profile of C18 was analyzed using Micrologg2N™ (Biolog) software. The only strain with a similar profile was *Klebsiella pneumoniae* subgroup A. Although the match was quite good (84%), specific physical characteristics of the *Klebsiella* genus indicated that strain C18 was not in this genus. For example, all *Klebsiella* are non-motile and encapsulated. C18 is motile and is not encapsulated. Because no other bacteria in the database provided

Table 1. UTILIZATION OF CARBON AND ENERGY SOURCES BY STRAIN C-18 AS DETERMINED BY BIOLOG™ ANALYSIS

Carbon Source	% Utilization
L-arabinose	79
D-galactose	80
α-D-glucose	72
m-inositol	61
cis-aconitate	95
citric acid	87
D-galacturonic acid lactone	66
D-galacturonic acid	85
D-glucosaminic acid	78
D-glucuronic acid	78
α-ketoglutaric acid	74
D-saccharic acid	100
D,L-lactate	85
L-serine	66
glycerol	77

% Utilization - values reported in Table 1 are expressed in terms of the percent positive reaction in the Biolog™ test relative to the carbon source best utilized (D-saccharic acid).

a logical match, it was concluded that C18 was not among the >400 bacteria contained in the Microlog2N database.

2.2.3 Production of Kodama-Like Products

A variety of colored products were produced by the Lehigh strains during incubation with DBT, depending on the medium employed. Distinctive colors appeared on Stanier's plates after prolonged (several hours) incubation of the Lehigh strains with DBT. A dark red color appeared with cultures grown on Stanier's agar plus 0.2% glycerol, an orange color was observed with glucose grown cultures, and a bright yellow color was seen with cultures grown on succinate. The DBT plates using Mormile and Atlas (1988) medium (glycerol as the carbon source) showed orange Kodama metabolites after 1 day of incubation at 30°C. (Kodama, et al. 1970).

Kodama-like compounds were not produced when incubating cultures with DBT sulfoxide or DBT sulfone, suggesting that DBT entered directly into the pathway(s) leading to the formation of colored compounds.

The actual chemical compositions of the Kodama-like products were not determined. In addition, it was not known whether the orange pigment produced was made up of a combination of the red and yellow compounds. In one study on the bioconversion of DBT, Hou and Laskin (1976) isolated and purified an orange pigment, and identified it as 4[2-(3-hydroxy)-thionaphthenyl]-2-hydroxy-3-butanoic acid. These investigators also partially purified two related red compounds. All of these materials appeared to contain an intact sulfur moiety; these compounds would not be present in a sequential pathway containing o,o'-biphenol as an intermediate or end product.

2.2.4 Requirement For Trace Metals

It was consistently observed that Kodama-like products were produced when the Lehigh strains were incubated with DBT in a complete medium. Production of colored products was not evident after incubation of C18 with DBT in minimal phosphate-glycerol medium. Thus, it appeared likely that some cofactors (i.e., metals) may be required in the medium for DBT metabolism. A sulfate-free trace metals solution was added to the culture (at 24 hours) (Table 2). Production of brownish-red material was evident in the culture after overnight incubation, indicating that the DBT was probably being metabolized and that certain metals were likely involved in a Kodama-like DBT degradation pathway. The colored compounds disappeared after extended incubation (5 days).

2.2.5 Production of Fluorescing Material

The spray test assay was used to screen for the possible production of o,o'-biphenol. Results indicated that the production of this material was dependent on the use of glycerol as the carbon source used to grow the organisms. The effort to identify the fluorescing material using TLC and HPLC, as discussed in the following section.

2.2.6 Identification of "4S" Intermediates Using TLC and HPLC

Attempts were made to identify the products of DBT oxidation by C18; the presence of DBT sulfone and o,o'-biphenol would have been of greatest interest. Thin layer chromatography (TLC), high performance liquid chromatography (HPLC) with UV and fluorescence detection and thermospray mass spectroscopy coupled to HPLC (HPLC/TSMS) were used in various analyses. In early studies, a DBT oxidation product was produced by C18 which had an R_f value similar to that

Table 2. SULFATE-FREE TRACE METALS SOLUTIONS FOR DBT DEGRADATION EXPERIMENT

Solution A		Solution B	
NH_4Cl	53.45 g	$MnCl_2 \cdot 4H_2O$	0.5 g
$CaCl_2 \cdot 2H_2O$	1.5 g	H_3BO_3	0.05 g
$MgCl_2$	0.4 g	$ZnCl_2$	0.05 g
$FeCl_2 \cdot 4H_2O$	0.4 g	$CuCl_2 \cdot 2H_2O$	0.03 g
H_2O[a]	1 liter	$NaMoO_4 \cdot 2H_2O$	0.01 g
		$CoCl_2 \cdot 6H_2O$	0.5 g
		$Na_2S_2O_3$	0.05 g
		$NiCl_2 \cdot 6H_2O$	0.05 g
		H_2O[a]	1 liter

(a) deionized water

of DBT sulfone. Using TLC, o,o'-biphenol was not detected. In a later study, HPLC analysis of the samples suggested that DBT sulfone was produced by C18 during DBT degradation (Fifth Quarterly report, September 14, 1990).

In this study the putative sulfone peak was shown to co-elute with DBT sulfone standard when the sample was spiked with DBT. Fluorescence detection was employed to further analyze the samples. The detector (Kratos Spectroflow 980 Programmable Fluorescence Detector) was set at 254 nm excitation, with a 418 nm cutoff emission filter. One fluorescing compound that eluted near or at the point of DBT sulfoxide was observed in both DBT spiked and control samples. The putative sulfone peak seen by UV detection in DBT amended samples was also observed by fluorescence detection. An additional fluorescing compound was observed in DBT spiked samples which eluted between DBT sulfoxide and DBT sulfone. This compound may be of importance relative to the plate screening assay since it may be responsible along with other compounds for the characteristic fluorescence.

Additional experiments were conducted in a final attempt to identify putative "4S" metabolites produced during the biological oxidation of DBT by C18. In one case, DBT (100 ppm) was added to the growth medium (Staniers medium with 0.2% glycerol) before inoculation with C18. In another case, DBT was added to the culture after growth. A medium control containing DBT (no organisms) was included as a control. After incubation with DBT for several days, acetonitrile was added directly to the cultures (60% final concentration). Selected samples were centrifuged to remove particulates and analyzed by HPLC with UV detection at 242 nm and HPLC/TSMS.

HPLC analysis showed that more chemical species were formed when DBT was added to C18 before growth than when added after growth. Most of these species were not identified. However, DBT sulfoxide appeared to be present (by comparison with retention times of known standards) at about 30 ppm when DBT was added before growth; somewhat less appeared to be present when DBT was added after growth. No DBT sulfone was detected in either case. The absence of DBT sulfone was confirmed by spiking

a C18 sample with DBT sulfone, in which case a new peak was observed (i.e., the sulfone did not coelute with any compounds in the C18 sample). Again, no o,o'-biphenol was detected.

Analysis of the samples by HPLC/TSMS provided additional evidence that DBT sulfoxide was present at relatively high concentration. Interestingly, a large amount of material corresponding to the same ion indicative of DBT sulfone was also observed, but at a different retention time. Since the elution pattern of this unidentified compound was different than that of DBT sulfone in the preliminary HPLC analysis (no sulfone peak was detected), it was believed that a different compound having the same representative ion for DBT sulfone was produced. One compound that can produce an ion identical to DBT sulfone via TSMS is hydroxy-DBT.

In conclusion, DBT metabolites produced during the oxidation of DBT by C18 were variable over time. A compound eluting near or at DBT sulfone appeared to have been produced in earlier studies but not in the later experiments. DBT sulfoxide was produced in relatively high concentrations in later studies.

2.2.7 Location of Fluorescent Metabolites

To determine if the fluorescing compound(s) observed on plates in the spray test assay was located intracellularly or extracellularly, samples of cells and agar surfaces were analyzed using UV fluorescence and HPLC. In the first experiment, colonies were isolated from plates which had shown fluorescence on a nitrocellulose filter. No UV-excited fluorescence was observed on the plates from which the colonies had been removed, suggesting that the material was either inside or bound to the cells. In a second experiment, solid cultures were incubated with DBT for 4 hours. The cell and agar fractions were first extracted with methylene chloride, and then with acetonitrile. The methylene chloride fractions were evaporated, and the residuals analyzed by HPLC. The agar samples appeared to be relatively free of any extractable metabolites. Control experiments indicated that pure o,o'-biphenol (1 μg) could be

extracted from agar using 60 % acetonitrile. It was concluded from these data that the fluorescent material was present in the cells, or in the cell walls but not in the culture medium.

2.2.8 Antibiotic Screening

The four Lehigh strains were screened for antibiotic resistance to ampicillin (Amp), chloramphenicol(Cm), kanamycin(Kan) and tetracycline(Tet). The results were as follows (r = resistant, s = sensitive):

A4 (Cm^r, Kan^s, Amp^r, Tet^s)
A15 (Cm^r, Kan^s, Amp^s, Tet^s)
B24 (Cm^r, Kan^s, Amp^s, Tet^s)
C18 (Cm^s, Kan^r, Amp^r, Tet^s)

Concentrations of antibiotics tested were as follows: Cm, 5 to 20 μg/mL; Kan, 50 to 200 μg/mL; Amp, 100 μg/mL; Tet, 20 μg/mL. Antibiotic resistance profiles of A15 and B24 were very similar. All strains were sensitive to tetracycline at 20 μg/mL. (Fry. et al., 1990).

2.2.9 Plasmid Screening

The four strains A4, A15, B24, and C18 were examined for the presence of plasmids using the method of Martin et al. (1981). Plasmids were present in A15, B24, and C18. No plasmid DNA was detected in the A4 strain.

2.3 Characterization of ALA and SDP Strains

Battelle bacterial soil isolates obtained from fuel-contaminated soil (ALA strains) and a sewage treatment plant (SDP isolates) were screened for possible utilization of sulfur from various organosulfonate and organosulfone sources.

Battelle ALA and SDP isolates were grown in sulfate-free medium (Hanna and Taylor, 1989) using DBT and DBT sulfone as the

sole sulfur source. Cultures were then analyzed for DBT metabolites by TLC and HPLC. Initial results suggested that several of the cultures make compounds which co-migrated with DBT sulfone when grown in DBT by TLC analysis (ALA9, ALA8, ALA3 and ALA1).

More rigorous testing was done to confirm the metabolism of organic sulfur compounds. While some of the strains were capable of using some organic sulfur sources, DBT and DBT sulfone were not used. The preliminary results were most likely the result of the presence of trace amounts of sulfate in the cells.

3.0 MOLECULAR GENETICS OF *THIOBACILLUS FERROOXIDANS*

Thiobacillus ferrooxidans is an acidophilic, gram-negative, chemolithotroph which derives energy from the oxidation of iron and sulfur and is potentially useful for removing sulfur from fossil fuels such as coal. Efficient removal of inorganic sulfur from high sulfur coal has been demonstrated on a laboratory scale. In principal, if genes encoding enzymes for metabolism of organic sulfur compounds could be introduced and made to function in *T. ferrooxidans*, a recombinant organism capable of removing both organic and inorganic sulfur from coal could be constructed. Development of genetic engineering technology for *T. ferrooxidans* strains is a necessary prerequisite for constructing such recombinant strains.

A functional system for genetic manipulation of *T. ferrooxidans* requires several components. Suitable host strains which are amenable to culture under laboratory conditions and which have a phenotype which allows selection of introduced marker genes must be identified. A functional vector for introduction of cloned genes must be developed. The vector must be capable of replication in *T. ferrooxidans*, and, ideally, in a second strain in which recombinant DNA manipulations can be performed. A functional marker gene for detection in *T. ferrooxidans* must be identified and inserted into vector and appropriate culture conditions for selection of recombinant strains determined. Finally, a method for introduction of

recombinant vectors into *T. ferrooxidans* must be developed.

To address each of these issues, the experimental plan for developing genetic engineering technology for *T. ferrooxidans* was divided into a series of tasks which proceeded in parallel. These included: purification and characterization of *T. ferrooxidans* strains; isolation, cloning and characterization of *T. ferrooxidans* plasmids; cloning of a potential marker gene for selection of recombinant *T. ferrooxidans* (the mercury resistance encoding *mer* operon from *T. ferrooxidans* strain DSM583); and development of electroporation techniques for *T. ferrooxidans*. The results of each of these experimental tasks are summarized in the following sections of the report.

3.1 Strain Purification and Characterization

Strains of *T. ferrooxidans* were purified and characterized to identify a suitable host strain of *T. ferrooxidans* for developing a recombinant plasmid vector system. Characterization included testing the relative sensitivities of the strains to various antibiotics and heavy metals. Also, the effects of organic sulfur compounds and coal cleaning agents on *T. ferrooxidans* were assessed. This section of the report summarizes the results of these analyses.

3.1.1 Strain Purification

T. ferrooxidans strains from the culture collection of Dr. O. Tuovinen of The Ohio State University were analyzed in detail during this program. In addition, a similar analysis was performed on a facultative heterotrophic *Thiobacillus* species, *T. cuprinus* (Huber and Stetter, 1990), which could be useful as an intermediate host for transferring recombinant vectors from *E. coli* to *T. ferrooxidans*. Finally, additional *Thiobacillus* strains were obtained from the American Type Culture Collection (ATCC) to provide a potential source of strains with alternative genotypes (ATCC strains 13598, 13661, 14119, 19859, 21834, 23270, and 33020). The ATCC strains were

only briefly examined during this program.

Seven strains of *T. ferrooxidans*, TFI29, 35, 70, 85, 91, 92 and DSM583, were characterized in detail. In preparation for antibiotic and heavy metal sensitivity testing, purified cultures were prepared from the starting mixed cultures by two rounds of single colony isolation and subculture, using both iron and tetrathionate media. In addition, the *T. ferrooxidans* strains were selected by culturing in a copper sulfate-containing medium. This procedure insures that the cultures isolated are *T. ferrooxidans* bacteria since the usual co-contaminants of *T. ferrooxidans* will not tolerate high concentrations of copper sulfate, while *T. ferrooxidans* is resistant to copper toxicity. Following purification, the strains were maintained throughout the program as active cultures in a mineral salts medium with ferrous sulfate.

Sensitivity testing for antibiotics and heavy metals was also performed with the facultative heterotrophic *T. cuprinus*. This organism oxidizes inorganic sulfur substrates, but not ferrous ion, and has metal leaching characteristics similar to *T. ferrooxidans*.

3.1.2 Antibiotic Sensitivity Screening of Thiobacillus ferrooxidans Strains

The seven *T. ferrooxidans* strains which had been purified from the Ohio State culture collection were tested for sensitivity to a variety of antibiotics to identify substances which could be used for selection of recombinant clones. Antibiotics tested included: ampicillin, gentamicin, chloramphenicol, neomycin, kanamycin, tetracycline, carbenicillin, and spectinomycin.

Initial studies used a plate sensitivity disc-test to measure the relative resistance of each strain to each of the antibiotics tested (Visca et al. 1989). To perform this analysis, antibiotics of a known concentration were added in a constant volume to porous discs and each disc placed in the center of a freshly inoculated plate. The growth medium used was

ferrous iron (9K) (Silverman and Lundgren, 1959) or tetrathionate. Antibiotic susceptibility was demonstrated by a zone of clearing around the disk after the plate cultures were incubated at 30°C for one week. Three concentrations of each antibiotic were used as follows: ampicillin (180, 360, 720 μg/disc), gentamicin (150, 300, 600), chloramphenicol (255, 510, 1020), neomycin (150, 300, 600), kanamycin (180, 360, 720), tetracycline (105, 210, 420), carbenicillin (180, 360, 720), and spectinomycin (375, 750, 1500). The results of the antibiotic sensitivity testing are presented in Table 3 (see also the Sixth Quarterly Report, December 14, 1990). In general, good antibiotic susceptibility was detected for ampicillin, carbenicillin (on tetrathionate only) and chloramphenicol. Moderate growth inhibition was obtained for spectinomycin (9K and tetrathionate) and for kanamycin and tetracycline on the tetrathionate medium. Antibiotics which did not seem to be appropriate for these strains were gentamicin and neomycin on both media. Strain TFI-35 did not grow well on solid media and results were not conclusive.

In subsequent experiments, selected *T. ferrooxidans* strains were tested for susceptibility to antibiotics in liquid TSM medium using a microtitre plate format (Visca et al. 1989). Results of these experiments defined specific concentrations of antibiotics required to inhibit growth of the various strains. Results were consistent with the disk plate assay. Resistance of TFI70 and TFI92 to chloramphenicol was further analyzed by plating on solid TSM medium with various concentrations of this antibiotic (Visca et al, 1989). The results were consistent with the preliminary testing and indicated that TFI70 was resistant to concentrations of chloramphenicol as high as 400 μg/ml while growth of TFI92 was inhibited when the chloramphenicol concentration was raised to 150 μg/ml. While chloramphenicol had originally appeared to provide a potential antibiotic for selection of recombinant clones, the resistance of the TFI92 strain to 100 μg/ml chloramphenicol was found to make clear selection of resistant transformants difficult (Fry et al., 1991, a, b).

Table 3. Antibiotic resistance of *T. ferrooxidans* TFI strains.

Strain	Zone of inhibition											
	Ampicillin			Gentamicin			Chloramphenicol			Neomycin		
	180	360	720	150	300	600	255	510	1020	150	300	600
TFI-29	55	62	68	10	10	10	27	40	51	-	-	-
TFI-35	55	58	65	-	-	-	20	37	42	-	10	10
TFI-70	50	57	61	-	10	15	20	29	40	-	-	-
TFI-85	58	61	64	-	-	-	37	45	56	-	-	-
TFI-91	56	58	64	-	-	-	30	43	50	-	-	-
TFI-92	55	58	67	-	-	-	47	54	67	30	5	5
DSM-583	55	57	86	-	-	9	40	33	60	5	-	-

Strain	kanamycin			Tetracycline			Carbenicillin			Spectinomycin			Alcohol
	180	360	720	105	210	420	180	360	720	375	750	1500	
TFI-29	8	-	8	8	32	36	-	20	10	10	13	33	-
TFI-35	-	-	-	-	-	11	60	15	11	-	-	23	-
TFI-70	-	-	12	-	-	5	-	-	11	-	18	15	-
TFI-85	-	-	9	18	46	84	-	-	9	15	11	26	-
TFI-91	-	-	-	34	8	13	-	-	9	-	11	24	-
TFI-92	-	-	-	NG	62	84	-	-	9	30	28	50	40
DSM-583	-	-	-	54	56	28	-	-	-	23	35	30	-

The concentrations are given as μg/disc. The results are expressed as the diameter of zone of inhibition (in mm); - = no inhibition.

Results of antibiotic sensitivity testing showed that the various strains tested had specific sensitivity profiles. In general, all of the strains were found to be relatively resistant to antibiotics when cultured in the ferrous sulfate medium. This result could be due to an instability of antibiotics in general in this acidic medium. Alternatively, it may be relatively difficult for a variety of antibiotics to enter into *T. ferrooxidans* cells when they are cultured in the ferrous sulfate medium. It seems less likely that gene encoded antibiotic resistance determinants are widespread in *T. ferrooxidans* species. It appears that tetrathionate containing media are more likely to be useful for antibiotic resistance selection of recombinant clones compared with ferrous sulfate medium.

3.1.3 Heavy Metal Sensitivity Screening of *Thiobacillus ferrooxidans* Strains

Resistance to heavy metals provides an alternative to antibiotic resistance as a means for selecting recombinant *T. ferrooxidans* clones. A variety of plasmid-encoded metal resistance functions, such as resistance to mercury and arsenic have been well characterized. To identify potential heavy metal sensitive host strains for genetic manipulations as well as sources of metal resistance genes for use as genetic markers, the sensitivity profiles of the seven purified *T. ferrooxidans* strains described above were examined by plate testing using the disc growth inhibition assay.

Specific metal ions tested and concentrations used were: Zn^{2+}, 46 μmol (3000 μg)/disc; Cu^{2+}, 18 μmol (1200 μg); Co^{2+}, 51 μmol (3000 μg); Ni^{2+}, 51 μmol (3000 μg); Cd^{2+}, 4.3 (480 μg); Ag^{+}, 0.14 μmol (15 μg); $SeO4_4^{2-}$, 0.21 μmol (30 μg); SeO_3^{2-}, 0.24 μmol (30 μg); AsO_4^{3-}, 0.86 μmol (420 μg); MoO_4^{2-}, 0.09 μmol (15 μg); and Hg^{2+}, 0.075 μmol (15 μg). The results are presented in Table 4 (see also Sixth Quarterly Report, December 14, 1990).

Table 4. Metal resistance of *T. ferrooxidans* TFI strains

Strain	Zone of inhibition										
	Zn	Cu	Co	Ni	Cd	Ag	SeO_4	SeO_3	AsO_4	MoO_4	Hg
TFI-29	-	27	-	-	20	86[a]	16	18	-	12	40
TFI-35	-	26	5	-	12	86	-	20	-	12	45
TFI-70	-	-	-	-	-	82	-	18	-	10	37
TFI-85	-	45	-	-	15	86	nd[b]	86	-	13	43
TFI-91	-	48	-	-	12	86	-	21	-	20	52
TFI-92	-	30	-	-	13	83	-	25	-	12	45
DSM-583	-	35	-	-	16	86	-	28	-	15	48

Zone of inhibition evaluated with a disc assay. The results are expressed as the diameter of zone of inhibition (in mm).
- = no inhibition.
[a]86 mm = 100% inhibition.
[b] = data not available.

All strains were resistant to zinc, cobalt, nickel, selenite and arsenate. A less pronounced resistance was evident with cadmium, molybdate and selenite for all strains with the exception of TFI85, which showed sensitivity to selenite. While all other strains were sensitive to copper, mercury, and silver (100% inhibition at the concentration used), strain TFI70 was resistant to these ions. The strain DSM583 was found to be relatively resistant to mercury in these tests and was subsequently used as a source of genomic DNA for cloning of the mercury resistance determinant, as will be described later. (Fry, et al. 1991 a, b).

Further analysis of the mercury resistance phenotypes of TFI70, TFI92, and DSM583 was carried out to clarify results of electroporation experiments (discussed later in this report). This analysis revealed that all three of these strains contained chromosomally located mercury resistance encoding genes. It was found that expression of the mercury resistance phenotype was inducible and full resistance required preliminary exposure to sub-inhibitory concentrations of mercury. These findings explained the relative sensitivity of TFI70 and TFI92 to mercury in the preliminary disc plate assay. However, the presence of mercury resistance genes in both TFI70 and TFI92, the two strains tested as hosts for transformation with the broad host range plasmid RSF1010::Tn501, prevented an unambiguous interpretation of electroporation experiments (see section 3.6).

3.1.4 Compatibility with Coal Cleaning Chemicals

To be useful for an industrial biodesulfurization process, it is important that the bacterial strains employed be fully viable when exposed to coal processing chemicals. Therefore, several strains of *T. ferrooxidans* were screened for sensitivity to chemicals used in the processing of coal slurries. These strains included TFI70, TFI29, TFI35, TFI85, TFI91, TFI92 and DSM583. Each strain of *T. ferrooxidans* was tested for sensitivity to coal chemicals at levels ranging from 0.1X to 10X of that used in the coal industry (kerosene was tested only with

TFI70). The chemicals used were supplied by PETC and are noted below:

Flocculents (Superfloc 16, Superfloc 214, Accoal-Floc 355)
Use Range: 1-10 ppm
Testing Range: 0, 0.1, 1, 10, 100 ppm

Frothers (Aerofroth 65, MIBC)
Use Range: 2-20 ppm
Testing Range: 0, 0.2, 2, 20, 200 ppm

Collector (Kerosene)
Use Range: 10-100 ppm
Testing Range: 0, 1, 10, 100, 1000 ppm

None of the coal chemicals irreversibly inactivated any of the strains, even at the highest concentrations tested. The result indicate that these thiobacilli suspended in the bulk fluid of a coal slurry should not be inactivated by coal preparation agents. Further work using coal slurries is needed to determine if these chemicals, which are expected to concentrate on the surface of the coal particles, affect depyritization activity.

3.1.5 DBT Utilization

The ultimate goal of DOE's biodesulfurization program is to develop an industrial bioprocess for removing both organic and inorganic sulfur from coal. While the ability of *T. ferrooxidans* to remove inorganic sulfur from coal has been well documented, the possibility that this species might also be able to metabolize organic sulfur compounds was a question. In addition, it is known that certain organic compounds can be toxic to *T. ferrooxidans* strains and it was conceivable that metabolites derived from biodesulfurization of organic sulfur compounds by a recombinant *T. ferrooxidans* strain could be toxic to this species. Therefore, *T. ferrooxidans* bacteria were

examined to determine whether they could utilize DBT or DBT metabolites, and also to assess the relative toxicity of these compounds to *T. ferrooxidans*.

Using the strain TFI70, aliquots (10 ml) of pre-grown *T. ferrooxidans* cells or sterile medium (as a control for abiotic changes) were added to separate tubes containing 100 ppm (final concentration) of DBT, DBT sulfoxide, DBT sulfone, o,o'-biphenol, biphenyl, or 2-phenylphenol. The cultures were analyzed by HPLC for metabolites after approximately one week of incubation at 26° C. No products appeared to be formed, and no obvious loss of the added compound was evident, indicating a lack of metabolism of these chemicals.

To determine whether organic sulfur compounds were toxic to *T. ferrooxidans* strains, cell viability in the presence of the organic sulfur compounds dibenzothiophene (DBT), DBT sulfone, DBT sulfoxide, and o,o'-biphenol was analyzed using the disc assay and the various *T. ferrooxidans* described previously. All but one strain had displayed similar sensitivities to the test compounds, as shown in Table 5 (Sixth Quarterly Report, December 14, 1990). Strain TFI 70 was resistant to DBT sulfone and DBT sulfoxide. Results indicated that the organic sulfur compounds tested could be tolerated by *T. ferrooxidans* strains (Fry et al., 1991 b).

3.1.6 Summary and Conclusions

Seven *T. ferrooxidans* from The Ohio State *University Thiobacillus* culture collection were purified and characterized by testing sensitivity to various antibiotics, heavy metals, coal cleaning agents, and organic sulfur compounds. The results defined sensitivity profiles for the strains analyzed and provided a basis for identifying potential antibiotic resistance or metal resistance determinants to be used as selectable marker genes with the various *T. ferrooxidans* strains.

Three of the strains tested were found to encode mercury resistance determinants, DSM583, TFI70, and TFI92. Based on these results the DSM583 strain was used as a source of

Table 5. Resistance of *T. ferrooxidans* TFI strains to organic sulfur compounds.

Strain	Zone of Inhibition												
	DBT			DBT sulfone			DBT sulfoxide			o,o'-biphenol			DMF
	75	150	300	75	150	300	75	150	300	75	150	300	
TFI-29	16	14	12	11	23	39	28	25	35	39	47	55	-
TFI-35	11	12	9	10	15	42	13	13	19	30	36	54	11
TFI-70	7	10	12	10	-	10	-	11	11	30	46	53	-
TFI-85	14	13	16	20	21	36	37	39	45	50	51	61	18
TFI-91	12	13	15	22	15	12	34	46	42	39	52	62	15
TFI-92	20	9	17	72	10	44	48	55	40	54	50	55	40
DSM-5083	13	15	11	32	16	13	NG	27	54	39	52	56	12

Zone of inhibition evaluated with a disc assay. The results are expressed as the diameter of zone of inhibition (in mm).
- = no inhibition.

genetic material for cloning and isolation of a *T. ferrooxidans* mercury resistance operon. The mercury resistance trait was found to be inducible by mercury and full expression of mercury resistance was found to require preliminary exposure to sub-inhibitory concentrations of this metal.

The strains tested were found to be resistant to various coal cleaning agents and organic sulfur compounds. None of the strains tested showed an ability to metabolize organic sulfur compounds.

The results of these analyses provide a basis for further development of genetic engineering technology for *T. ferrooxidans*.

3.2 Plasmid Isolation, Cloning, and Characterization

Genetic manipulation of *T. ferrooxidans* species requires the development of techniques for the introduction of genetic material into *T. ferrooxidans* cells. We have focussed on the development of plasmid vectors for this purpose. The effort in isolation, cloning and characterization of plasmids from *T. ferrooxidans* completed during this program is described in the following section. Additional research related to development of genetic engineering technology for *T. ferrooxidans* is described in subsequent sections. This includes, the isolation of a potential marker gene for selection of recombinant *T. ferrooxidans* cells (the *T. ferrooxidans mer* gene) and development of electroporation techniques for *T. ferrooxidans*.

Plasmid molecules were isolated from a series of purified strains of *T. ferrooxidans*. DNA restriction and Southern hybridization analyses of these plasmids indicated that they share substantial sequence homology, and appear to form a family of related plasmids. DNA sequence analysis has revealed a region which has structural features common to a number of plasmid origins of replication. However, this region has no strong homology to other sequenced *T. ferrooxidans* plasmids, to the broad host range IncP and IncQ plasmids, or other known Gram negative replicons.

3.2.1 Plasmid Isolation

In order to develop plasmid vectors for performing genetic manipulation, plasmids were isolated from a series of purified strains of *T. ferrooxidans*. Strains examined included; TFI29; TFI35; TFI85; TFI70; TFI91; TFI92; and DSM583.

Following plasmid isolation, DNA restriction analysis was performed on material isolated from each strain. The results indicated that four of the strains, TFI29, TFI85, TFI91, and TFI92, contain homologous 9.8 kb plasmids with identical restriction maps for EcoRI, SalI, and SacI. The more complete restriction map of the plasmid from TFI91 (pTFI91, Figure 2) is virtually identical to that of three plasmids previously isolated from copper mines in Chile (Sanchez et al., 1986).

Two of the *T. ferrooxidans* strains examined, TFI35 and DSM583, were found to contain somewhat larger plasmid molecules of 15.0 and 20.0 kb respectively. The relationship between these plasmids and pTFI91 was investigated by Southern hybridization analysis.

3.2.2 Plasmid Cloning

The complete DNA sequence of the *T. ferrooxidans* plasmid PTFI91 as well as smaller portions, were cloned into the pUC19 derivative pHSG398 to locate the minimal region required for replication and to construct a potential shuttle vector for *T. ferrooxidans*. Four of the clones constructed are shown in Figure 3. A complete list of the clones constructed is given in Table 6.

Genetic analysis of pTFI91 encoded replication functions would be facilitated if a system in which DNA replication directed by the pTFI91 origin of replication could be developed. To test for replication from the pTFI91 origin in *E. coli* the clones shown in Figure 2b were introduced into *polA* mutant strains. Results indicated that, while control IncQ plasmids could replicate in the *polA* strains, none of the pTFI91

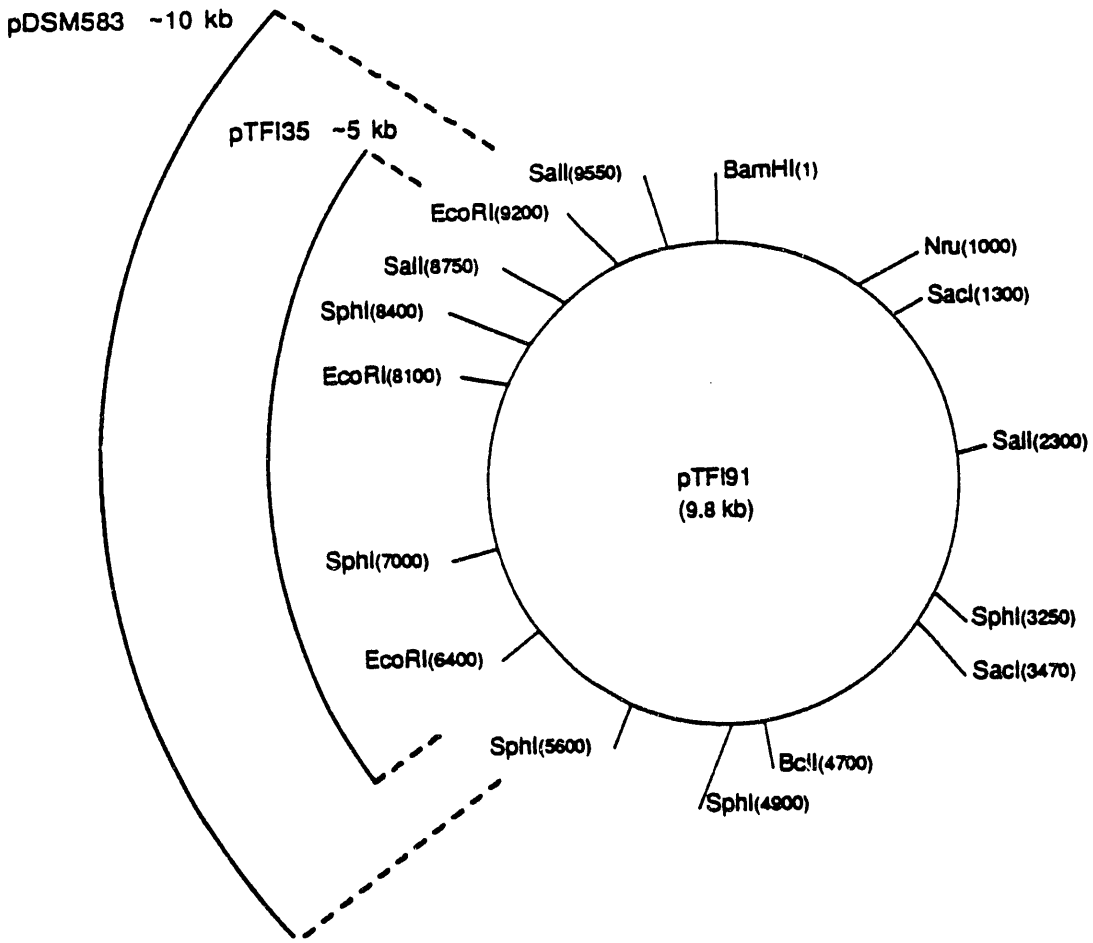


Figure 2: Restriction maps of the *T. ferrooxidans* plasmids pTFI91, pTFI35, and pDSM583.

The approximate location of the additional DNA sequences present in pTFI35 and pDSM583 is indicated on the left of the pTFI91 map.

clones could replicate in these strains. Either construction of each of the pTFI91 clones disrupts some function essential for plasmid replication, DNA polymerase I is required for pTFI91 replication, or the pTFI origin of replication does not function in *E. coli*.

Potential plasmid shuttle vectors which contained all or large portions of the pTFI91 plasmid inserted into pHSG398 were also used to test for transformation of *T. ferrooxidans* strains by electroporation. These experiments are summarized in the discussion of T. ferroxidans electroporation experiment (section 3.6).

3.2.3 Southern Hybridization Analysis

Southern hybridization analysis was performed to analyze DNA sequence homology between the various *T. ferrooxidans* plasmids mentioned earlier. These experiments confirmed that, of the seven strains tested, at least six harbored related plasmids which could be divided into three types based on their restriction patterns. Plasmid molecules isolated from six purified strains of *T. ferrooxidans* were analyzed (Figures 4-6). Results confirmed that the plasmids in strains TFI29, 85, 91, and 92 shared significant homology (Figure 4).

Strains TFI35 and DSM583 contain plasmids of 15 kb and 20 kb respectively. These plasmids also share substantial sequence homology to the pTFI91 plasmid. It appears that the larger size of these plasmids is the result of DNA insertions into a plasmid closely related to pTFI91. It is possible that pTFI35 and pDSM583 contain *T. ferrooxidans* transposons.

Based on DNA restriction and Southern hybridization analyses, tentative restriction maps for the pTFI35, and DSM583 plasmids were prepared (Figure 2b). The analysis of these plasmids defined a region of DNA which may contain the minimal replicon. Significant homology between pTFI91 and both pTFI35 and pDSM583 was found centered around the 2.2 kb SacI restriction fragment present in each plasmid. Sequence homology appeared to diminish at more distal regions. DNA sequence analysis,

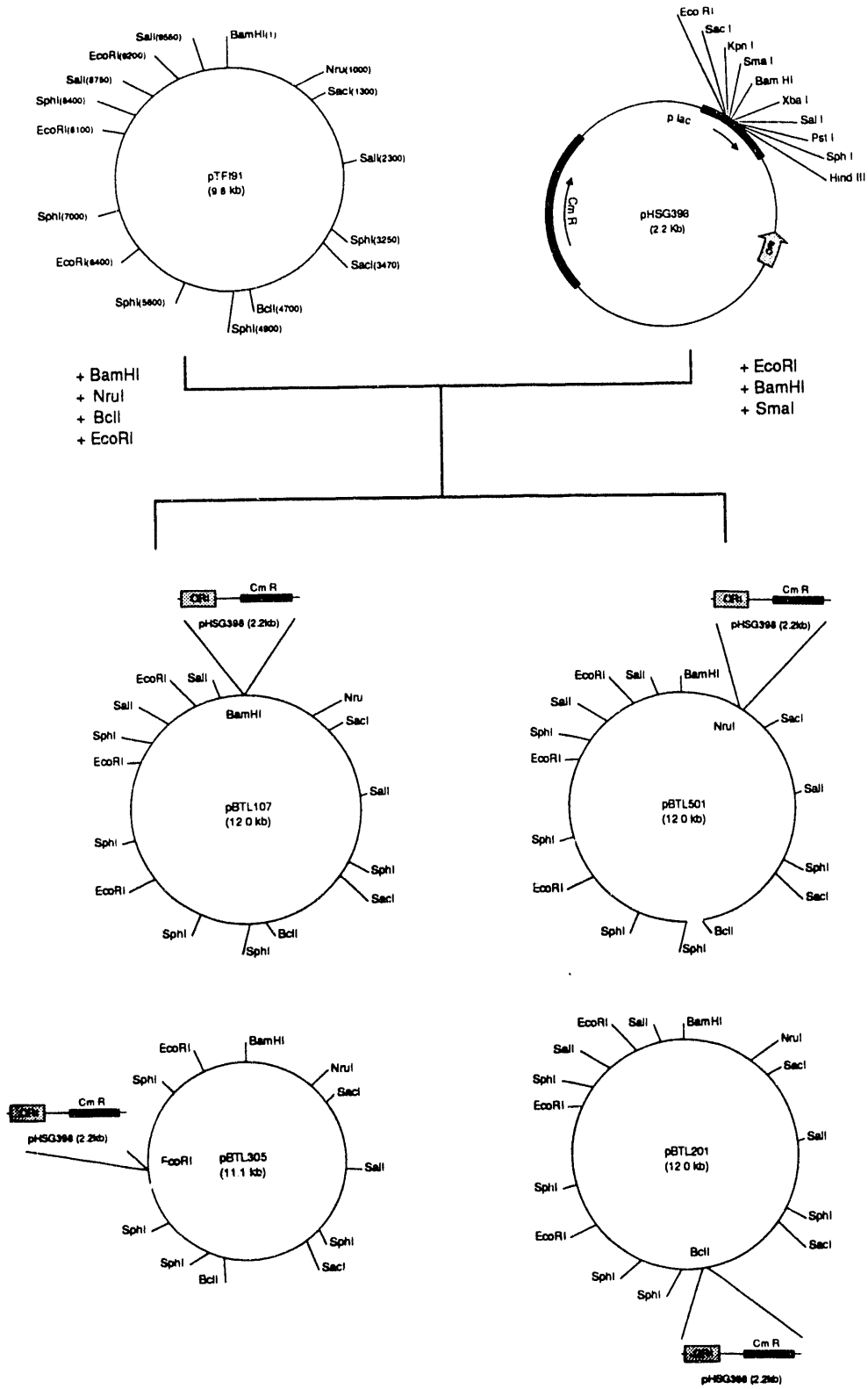


Figure 3: Construction of the pTFI91 clones, pBTL107, pBTL201, pBTL305, and pBTL501.

pTFI91 DNA was cleaved separately with BamHI, BclI, EcoRI, or NruI and ligated to the pUC19 derivative pHSG398 which had been cleaved with the appropriate restriction enzyme.

Table 6. Clones of pTFI91 in pHSG398

Restriction Fragment Cloned	Plasmid Designation
0.8 kb SalI	pBTL009
2.5 kb SalI	pBTL004
6.5 kb SalI	pBTL015
2.5 & 6.5 kb SalI	pBTL017
9.8 kb BamHI	pBTL107
9.8 kb BclI	pBTL201
1.1 kb EcoRI	pBTL306
1.6 kb EcoRI	pBTL303
1.6 & 7.2 kb EcoRI	pBTL305
1.1 & 1.6 & 7.2 kb EcoRI	pBTL302
0.8 kb SphI	pBTL421
1.4 kb SphI	pBTL414
1.7 kb SphI	pBTLB423
4.5 kb SphI	pBTL413
1.4 & 1.7 kb SphI	pBTL415
1.4 & 4.5 kb SphI	pBTLB429
1.7 & 4.5 kb SphI	pBTLB427
0.8 & 1.4 & 4.5 kb SphI	pBTLB424
1.4 & 1.7 & 4.5 kb SphI	pBTLB422
9.8 kb NruI	pBTL501
2.2 kb SacI	pBTL614
7.6 kb SacI	pBTL613

The clones were obtained by inserting restriction fragments of pTFI91 into pHSG398. Refer to the restriction map of pTFI91 (Figures 2 and 3) for the location of the various restriction fragments in the pTFI91 plasmid. In some cases clones were found to contain more than one pTFI restriction fragment. In these cases, the inserted fragments are not always in the same relative orientation with respect to each other as in the pTFI91 plasmid.

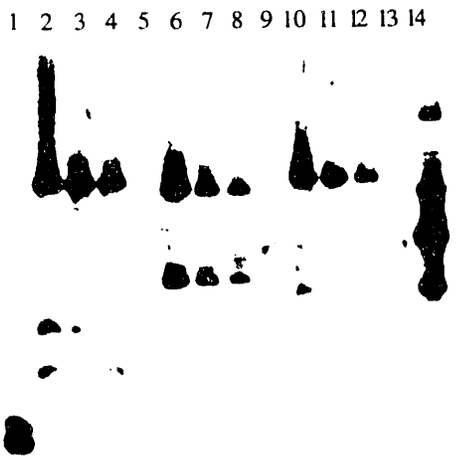


Figure 4: Southern hybridization of *Thiobacillus ferrooxidans* plasmids

Restriction patterns of plasmids isolated from *T. ferrooxidans* strains TFI29, 35, 85, 91, 92 and DSM583 were analyzed by agarose gel electrophoresis and transferred to nylon membranes. The membrane bound fragments were hybridized to a ^{32}P labeled probe prepared from cloned pTFI91 plasmid (pBTL107). Filters were then stringently washed and autoradiographed over night.

Lane 1: BRL size markers (visible on ethidium stained gel); low molecular weight bands contain pBR322 sequences hybridizing to the pHSG398 cloning vector. Restriction digests for lanes 2 through 4 are: pTFI29, pTFI85, and pTFI92 cut with EcoRI; lanes 6 through 8 are: plasmids pTI29, pTFI85, and pTFI92 cut with SalI; lanes 10 through 12 are: pTFI29, pTFI85 and pTFI92 cut with SacI; lane 14: cloned pTFI91 plasmid (construct pBTL107) cut with Sac I. Note that the three *Thiobacillus* plasmids appear to have identical restriction patterns. Additional data confirms their near identity to each other and pTFI91(not shown).

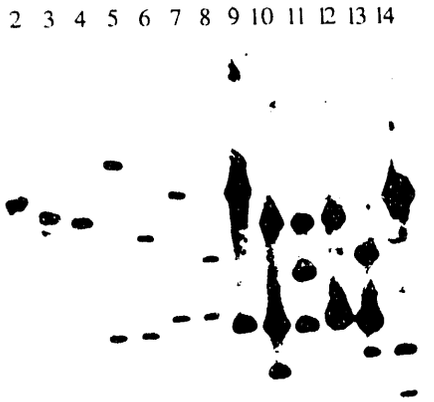


Figure 5: Restriction digest and Southern blot of the plasmid from DSM583 and cloned pTFI91 plasmid (pBTL107).

Lane 1: size markers; Lane 2-8: pDSM583 cut with BamHI (2), EcoRI (3), BamHI plus EcoRI (4), SacI (5), SacI plus BamHI (6), SalI (7), and SalI plus EcoRI (8) respectively. Lanes 9-14, pBTL107 cut with BamHI (9), BamHI plus SacI (10), SacI (11), SalI (12), SalI plus EcoRI (13) and EcoRI (14). Note that the restriction patterns indicate that the DSM plasmid is similar to, but larger than the pTFI91 plasmid. Additional, non-hybridizing pDSM583 plasmid fragments were visible on the ethidium bromide stained gel.

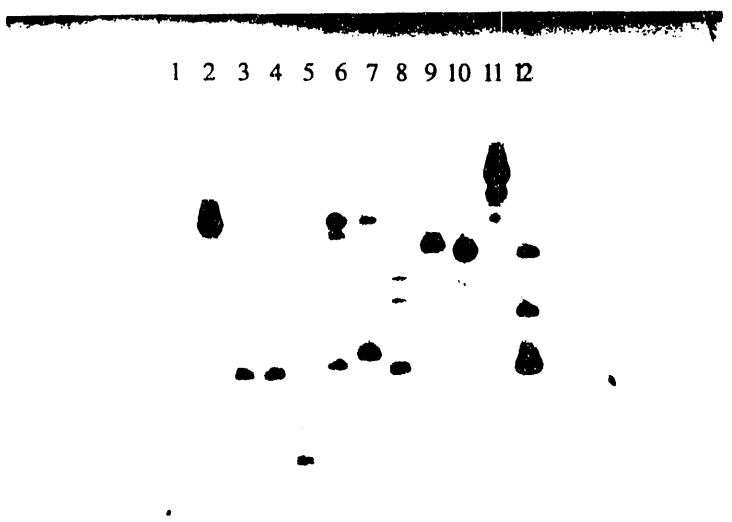


Figure 6: Restriction digest and Southern blot of the plasmid from TFI35 and pBTL107.

Lane 1: pBTL107 plus SacI. Lanes 2 through 11: pTFI35 cut with: no enzyme (2), EcoRI and BamHI (3), EcoRI (4), EcoRI and SacI (5), SalI (6), SalI and BamHI (7), SacI and SalI (8), SacI and BamHI (9), SacI (10), and BamHI (11). Lane 12: BRL size markers. The pTFI35 plasmid is about 15 kb in size. Restriction fragments which map to the additional DNA sequence present in this plasmid were visible on the ethidium bromide stained gel but did not hybridize with the pBTL107 probe. pTFI91 plasmid which has structural features common to bacterial plasmid origins of replication. This region is likely to be the vegetative origin of replication in this plasmid. In addition, a DNA sequence which appears to be a *Thiobacillus* insertion sequence also has been identified.

described below, revealed that the 2.2 kb SacI fragment contained structural features common to plasmid origin of replication regions. It is hypothesized that additional plasmid encoded replication functions are located in the conserved region between the SacI site at 1300 base pairs on the pTFI91 restriction map and the unique BamHI site.

Because previously cloned and sequenced *T. ferrooxidans* plasmids had been found to share DNA sequence homology with IncP or IncQ plasmids, the pTFI91 plasmid was compared with a series of plasmid vectors from various bacterial species (Drolet, Dorrington and Rawlings, 1990). Hybridization experiments did not show any homology between the T. ferrooxidans plasmids and incompatibility groups P, Q, N, and W, as well as the *Streptomyces* plasmid pHN1. This result is supported by the results of DNA sequence analysis, described below, which indicated that pTFI91 does not share sequence similarity with known plasmid origin of replication regions.

An additional *T. ferrooxidans* strain, TFI70, was also analyzed for the presence of plasmid molecules. While no plasmid could be clearly detected in this strain, Southern hybridization using pTFI91 DNA to probe total cellular DNA from TFI70 revealed the presence of an approximately 8.0 kb EcoRI restriction fragment with substantial sequence homology to pTFI91. It is not clear whether TFI70 actually contains a plasmid which has not yet been isolated or has a sequence homologous to pTFI91 inserted into its chromosome.

3.3 DNA Sequence Analysis of the Putative pTFI91 Origin of replication pTFI91

Our work on the development of genetic engineering technology for *T. ferrooxidans* focussed on the construction of plasmid vectors. The approach was to isolate natural, cryptic plasmids from *T. ferrooxidans* strains and use recombinant DNA manipulations to construct potential plasmid shuttle vectors. This effort has been complicated by the lack of known genetic markers on the cryptic *T. ferrooxidans* plasmids being used, as

well as the lack of an experimental system for identifying the location of plasmid encoded replication functions. Knowledge about the location and activities of the replication functions specified by plasmid DNA sequences would facilitate the construction of useful plasmid vectors for this bacterium. Therefore, DNA sequence analysis was used to analyze a portion of the *T. ferrooxidans* plasmid, pTFI91. Results reported in this section describe the identification of a region of the pTFI91 plasmid which has structural features common to bacterial plasmid origins of replication. This region is likely to be the vegetative origin of replication in this plasmid. In addition, a DNA sequence which appears to be a *Thiobacillus* insertion sequence also has been identified.

3.3.1 DNA Sequence Determination

Southern hybridization analyses of the plasmids present in three of the *T. ferrooxidans* strains, TFI91, TFI35, and DSM583, identified a region centered around the 2.2 kb SacI restriction fragment of pTFI91 that appeared to be highly conserved in all three plasmids. The hypothesis that this conserved region contained the plasmid origin of replication was tested by determining the DNA sequence of the 2.2 kb SacI restriction fragment and analyzing this DNA sequence.

The complete nucleotide sequence of the small SacI restriction fragment of pTFI91 is shown in Figure 7. Analysis of this sequence revealed that it contains structural features common to plasmid origins of replication (Bramhill, and Kornberg, 1988b; Kues, and Stahl, 1989). In addition, a second sequence, immediately adjacent to this putative vegetative origin of replication, was found to contain an open reading frame which appears to encode a protein with significant homology to several bacterial transposase enzymes and could represent of *T. ferrooxidans* insertion sequence.

The structural features identified in the putative origin of replication of pTFI91 are displayed graphically in Figure 8. The origin sequence has been divided into three

regions, termed Region I, Region II, and Region III. These are: Region 15; a series of tandem direct repeats, Region II is an AT-rich region of complex structure. Region III is a GC-rich region. A more detailed description of each region is provided below.

3.3.2 Region I: Tandem Direct Repeats

DNA sequence analysis of a number of plasmid origin of replication regions from gram-negative bacteria has shown that many origin sequences contain a series of tandem, directly repeated sequence elements (Kues, and Stahl, 1989). Molecular analysis of these sequences indicates that the repeated elements are binding sites for plasmid encoded replication initiator proteins (rep proteins). Binding of rep proteins to the repeated elements is essential for initiation of DNA replication. In many cases, the repeated elements are organized with respect to the replication initiator protein gene so that binding to certain repeated elements also controls expression of the rep protein itself. Thus, the repeat elements control both rep protein synthesis and the initiation of plasmid replication.

The pTFI91 plasmid was found to contain a 282 base pair sequence which consisted of a series of nine tandem direct repeats. Figure 9 shows the homology between these nine repeated sequence elements (Figure 9, repeats 1-9). These elements all share 23 base pairs of homology (referred to as the core element in this discussion). In addition, two versions of the repeated element show more extensive sequence homology. There are four copies of a 33 base pair element and three copies of a second 32 base pair element. The 33 base pair element differs from the 32 base pair element in the sequence found at the 3′ end. An additional version of the core repeat element also was found within the AT-rich region, Region II (Figure 9, repeat 10, see also Figure 8).

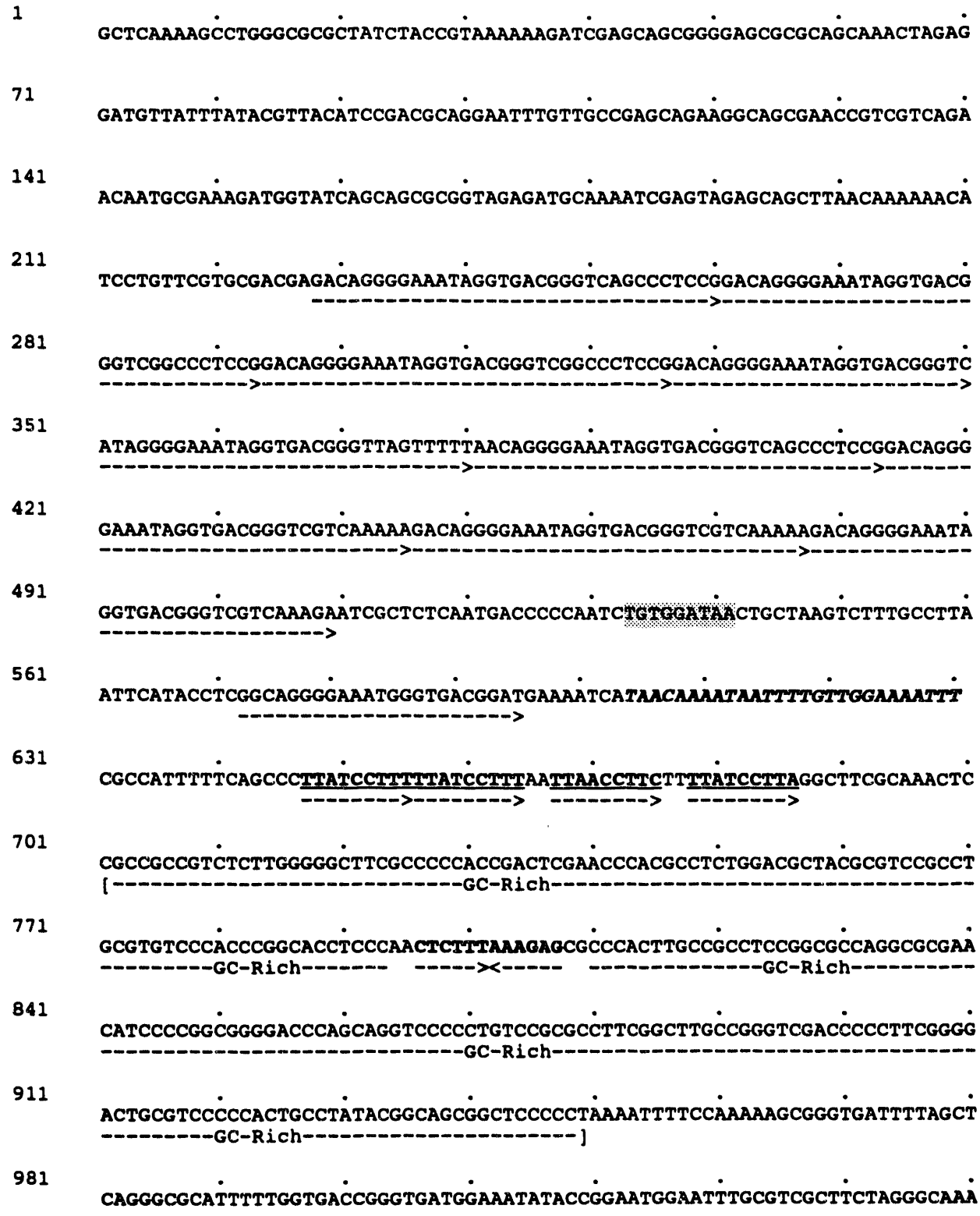
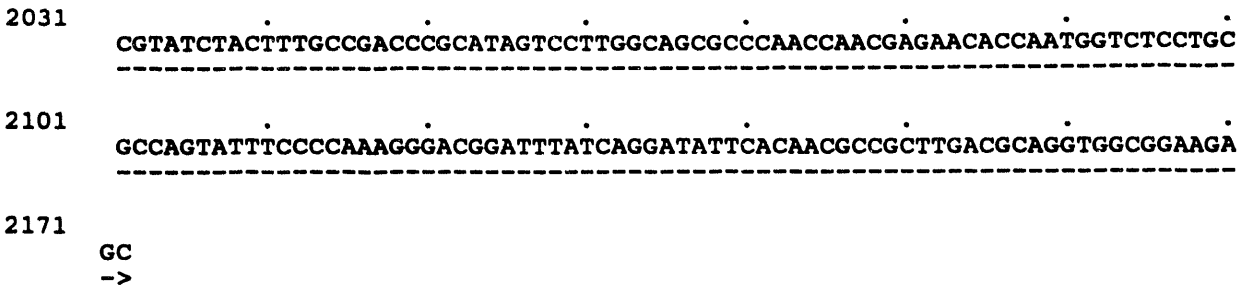


Figure 7. The nucleotide sequence of the 2.2 kb SacI fragment of pTFI91

```
1051
TACGGGCTATTTACGGCGGGGTTCTGCATCAGGCTTCTGGTCGGTTGTGCTGCGAATGATGCTTGGGACC

1121
GGCCTTGCTTGCGTTCGTTACCGTCACCGTGCGCTGCTGACCGGCTTTCTTCGAGTGCAAAGTGAGCGGT

1191
TTCTACTGGGGGGGCGCGCTCTAGAATGAAGTGCAACATCTCCGGTCGAAGGATGGCGGAGGTGAGTGAA
                                                          SD

1261
TGGGGACGAGATATCAGCAGTTGCAGTCGGAGCAGAGGAACCAGATTCAGCGAGGGCTGAACGAGGGGTT
->-----------------------------ORF-----------------------------------

1331
GAGTATGCGGGCCGTGGCCAAGCAGATAGGGAGGAGTCCCAGCACGGTCAGCCGGGAGGTACGCCGGGGT
----------------------------------------------------------------------

1401
TTGGTGGGAGAAACCTACGATGCGATACAAGGCCGGGAGGAGGCGCAGAGGCGCCTTCGTAAGGGGGTTA
----------------------------------------------------------------------

1471
GAAAGCTGGTGGGAGGCGCGCCCTTAACCAACGCGGTGACACACGCTATCCTGCAAAGGAAATGGTCACC
----------------------------------------------------------------------

1541
AGAGCAGGTGGCCGGGAGGTTGCGGATGGACTATCCCGAGGACAAGCAGTGGCGTGTCTCCCATGAGACC
----------------------------------------------------------------------

1611
ATTTATCAGTTCATCTATGCCCACCCGGCCGGTGAGCTGCGTAAGGCGCTGATAGCGGCGCTACGCCAGG
----------------------------------------------------------------------

1681
GACACGCAAAGCGCAAGCCGCGCACACGCGGGAAGGACCGGCGCGGACAACTGCGGAACATGCGTTCCAT
----------------------------------------------------------------------

1751
CGGGGAGCGTCCCTTGGAGGCCCAAGACCGCGAGATACCCGGCCACTGGGAAGGAGACTTCATCAAAGGG
----------------------------------------------------------------------

1821
GCTTTCAACGGCAGCGCCATTGGTACTCTGGTGGAGCGCAGCAGCCGTTTCGTGCTTCTGGTCAGGATGG
----------------------------------------------------------------------

1891
AAGGCACCGATGCCGACGCGGCCCTGGAGGGGTTCACCAGGCGCATGCGCACCTTGCCCAAGTCCATCCT
----------------------------------------------------------------------

1961
GCGGACCCTCACCTATGACCAGGGCAAGGAGATGGCACGGCACGAGGAGCTGGAGCGCAAGGTGGGCATC
----------------------------------------------------------------------
```

Figure 7. The nucleotide sequence of the 2.2kb SacI fragment of pTFI91 (Continued)


2031
CGTATCTACTTTGCCGACCCGCATAGTCCTTGGCAGCGCCCAACCAACGAGAACACCAATGGTCTCCTGC
2101
GCCAGTATTTCCCCAAAGGGACGGATTTATCAGGATATTCACAACGCCGCTTGACGCAGGTGGCGGAAGA
2171
GC
->
Figure 7. The nucleotide sequence of the 2.2kb SacI fragment of pTFI91 (Continued)
The complete nucleotide sequence of the 2.2kb SacI fragment of pTFI91 is shown. The tandem direct repeats extending from base pair 228 to 509 as well as the repeat present in the AT-rich region at base pairs 573 to 595 are underlined with dashed arrows. The sequence with homology to the DnaA binding site at position 533 to 541 is boxed. The sequence with homology to the consensus IHF binding site at positions 604 to 630 is shown in bold and italics. The nine base pair AT-rich direct repeats at positions 647 to 686 are shown in bold and underlined, and also are underlined with dashed arrows. The GC-rich region is underlined with a dashed arrow. The 12 base pair AT-rich palindrome within the GC-rich region at position 796 to 807 is indicated in bold and underlined with inverted arrows. A potential SD sequence at position 1248 to 1252, in front of the long open reading frame, is underlined. Finally, the extended open reading frame with homology to the IS30 transposase is underlined with a dashed line.

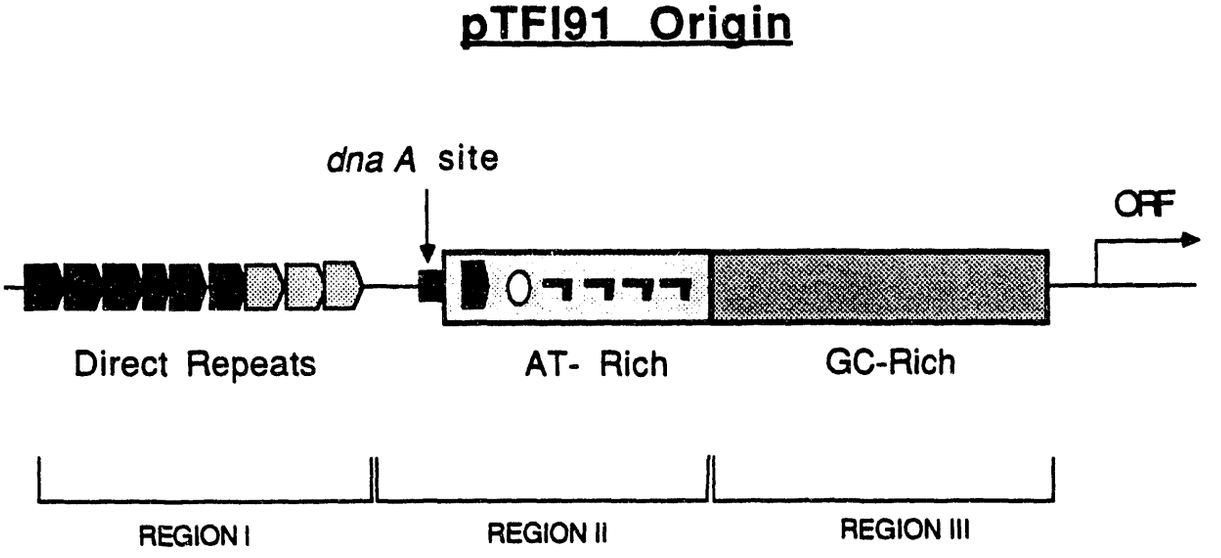


Figure 8: Structural features of the putative pTFI91 origin of replication

The boxes on the left indicate the positions of nine tandem direct repeats. Direct repeat sequences indicated by black boxes share 33 base pairs of homology. The dark gray boxes share 24 base pairs of homology and the light gray boxes share 32 base pairs of homology. All of the repeats share at least 24 base pairs of homology. An additional 23 base pairs of homology to the direct repeat element is located within the AT-rich region.

A sequence with nine base pairs of perfect homology to the *E. coli DnaA* binding site is located between the tandem direct repeats and an AT-rich sequence. The AT-rich region contains four tandem direct repeat sequences nine base pairs long. A seventeen base pair sequence with homology to the general consensus binding site for IHF is located between the 23 base pair version of the upstream direct repeats and the tandem nine base pair AT-rich repeats. The AT-rich sequence is followed by a GC-rich sequence. Finally, the location of a DNA sequence which appears to contain a *Thiobacillus* insertion sequence is indicated by the ORF.

1) GACAGGGGAAATAGGTGACGGGTCAGCCCTCCG
2) GACAGGGGAAATAGGTGACGGGTCGGCCCTCCG
3) GACAGGGGAAATAGGTGACGGGTCGGCCCTCCG
4) GACAGGGGAAATAGGTGACGGGTC
5) ATAGGGGAAATAGGTGACGGGTTAGTTTTT
6) AACAGGGGAAATAGGTGACGGGTCAGCCCTCCG
7) GACAGGGGAAATAGGTGACGGGTCGTCAAAAA
8) GACAGGGGAAATAGGTGACGGGTCGTCAAAAA
9) GACAGGGGAAATAGGTGACGGGTCGTCAAAGA
10) GGCAGGGGAAATGGGTGACGGAT

Figure 9: Direct repeats in the putative origin of replication of pTFI91

This figure shows the homology, between the direct repeat elements found at the putative origin of replication of pTFI91. Repeats numbered 1 to 9 occur in tandem between base pairs 228 and 509 (see Figure 1). Repeat number 10 is located in the AT-rich region at base pairs 573 to 595.

The organization of the repeats in Region I with respect to Regions II and III, is similar to the organization of several well characterized plasmid origins of replication. This similarity, which is unlikely to occur by chance, is consistent with the interpretation that the pTFI SacI restriction fragment contains the pTFI91 plasmid origin of replication. This interpretation is further supported by the identification of additional sequence elements located in Region II (see below).

Comparison of the DNA sequence of the pTFI91 repeats in Region I with other known plasmid sequences in the Genbank database failed to identify any known DNA sequence with significant homology to this portion of pTFI91. This result suggests that the pTFI91 plasmid origin represents a novel type of plasmid replicon.

3.3.3 Region II: the AT-Rich Region

The nine tandem direct repeats of Region I are immediately adjacent to a 160 base pair AT-rich region of complex structure. This region contains; a sequence element with perfect homology to the *E. coli* DnaA protein binding site, a version of the 23 base pair direct repeat core element (see Figure 9, repeat 10), a potential integration host factor (IHF) binding site, and a series of four tandem nine base pair AT-rich direct repeats (see Figures 7 and 8 for general organization). Each of these sequence features is described in greater detail below.

3.3.3.1 DnaA Binding Site

The bacterial DnaA protein has been shown to perform a key role in the initiation of DNA replication at origins of replication. (Fuller, et al., 1984, Funnell, and Kornberg, 1984; Bramhill, and Kornberg, 1988a). Sequence analysis of various plasmid origin of replication regions indicates that one or more potential binding sites for this protein are commonly located within essential cis-acting replication sites. The pTFI plasmid was found to contain a nine

base pair sequence (Figure 7, base pairs 533-541) with perfect homology to the consensus DnaA binding site (see Figure 10).

3.3.3.2 AT-Rich Direct Repeats

The DnaA protein has been shown to function at origin of replication regions by interacting with adjacent AT-rich repeat sequence elements, catalyzing the ATP dependant melting of these regions, and AT-rich repeats are commonly found in plasmid origins of replication. Inspection of the pTFI91 origin sequence (Figure 7, base pairs 647-686) reveals a series of four tandem direct repeats with imperfect sequence homology (Figure 10). It is interesting to note that the repeated AT-rich elements share partial homology to the consensus DnaA binding site (see Figure 10).

3.3.3.3 Integration Host Factor Binding Site

In addition to DnaA protein binding sites and AT-rich direct repeats, plasmid origins of replications sometimes contain binding sites for integration host factor (IHF). This multi-functional protein appears to contribute to the initiation of DNA replication by inducing a bent DNA structure around the site of pre-priming complex formation. A potential IHF binding site was identified in the pTFI origin region between base pairs 604 and 630 (Figure 7). The pTFI sequence has 88% (15/17 base pairs) homology to the general IHF consensus binding site (Tsai, et al. 1990). (Figure 10). Binding of IHF at this position could contribute to the formation of a pre-priming complex by bending the DNA of the origin region in a way which positions DnaA protein, bound at base pairs 533-541, in close proximity to the AT-rich repeats at positions 647-686. This event would be followed by a DnaA catalyzed unwinding of the AT-rich sequence and formation of a pre-priming complex.

a. DnaA Binding Site

Consensus *DnaA* Site	5′ TGTGGATAA 3′
pTFI91 2.2 kb SacI, bp 533-541	5′ TGTGGATAA 3′

b. Potential Integration Host Factor Binding Site

Consensus IHF Binding Site	5′	YAANNWNYTGWWWWNNNNNNNWWWWWW 3′
		*** * * **** ******
pTFI Sequence 604-630	5′	TAACAAAATAATTTTGTTGGAAAATTT 3′

c. A-T Rich Direct Repeats (bp 647-686)

5′ TTTTTCAGCCC**TTATCCTTT** **TTATCCTTT**AA**TTAACCTTC**TT**TTATCCTTA**GGCTTC 3′

1) TTATCCTTT
2) TTATCCTTT
3) TTAACCTTC
4) TTATCCTTA

Figure 10: Structural features within the AT-rich region of the Putative pTFI91 Origin of replication

a. Homology to the consensus DnaA binding site is shown. Note that the DnaA consensus sequence indicated is the complement of the sequence as usually written, 5′ TTATCCACA 3′, since the pTFI homology is positioned in inverted orientation from this common presentation.

b. Homology to the general IHF consensus binding site is shown. Asterisks indicate bases in pTFI91 which match the consensus. Note that the symbol "Y" means T or C and "W" means A or T.

c. The complete sequence of the AT-rich direct repeat region and homology between the repeated elements is shown.

In summary, analysis of the DNA sequence of Region II in the putative pTFI91 origin of replication indicates that this region contains a variety of sequence elements commonly found in plasmid origins of replication. This region appears likely to contain the site at which pTFI pre-priming complex formation occurs.

3.3.4 Region III: the GC-rich Region

Analysis of the putative pTFI origin of replication sequence indicates that, in addition to the direct repeats and AT-rich regions, pTFI also contains a third type of DNA sequence commonly found at plasmid origins of replication, a GC-rich region. In the sequence presented in Figure 7, the GC-rich region spans approximately 250 base pairs, between positions 700 and 950.

In contrast to the broad host range plasmid RSF1010 and certain other plasmid replicons, the pTFI91 GC-rich region does not appear to contain significant inverted repeat elements (Scholz, et al., 1989; Dorrington and Rawlings, 1990). However, there is a 12 base pair, AT-rich palindrome centered at base pair 801.

The function of the GC-rich regions in plasmid origins of replication is less well understood, compared with the direct repeats and AT-rich sequences. It has been suggested that a GC-rich sequence immediately adjacent to the AT-rich tandem direct repeats could function to limit the local unwinding of DNA at the AT-rich repeats, allowing precise positioning of the pre-priming complex. Furthermore, in cases such as the origin of replication of the IncQ plasmid RSF1010, inverted repeats within the GC-rich region may form cruciform structures which define the site of initiation of DNA replication. In the case of pTFI91, only the former hypothesis appears to be relevant.

3.3.5 Summary of Sequence Analysis

DNA sequence analysis of a portion of the pTFI91 plasmid has revealed a region, located within the 2.2 kb SacI restriction fragment, which contains structural features characteristic of plasmid origins of replication. These features include: a series of nine tandem direct repeats; an AT-rich region with a potential DnaA protein binding site, a copy of the direct repeat elements, a series of four directly repeated AT-rich elements, and a potential IHF binding site; and a 250 base pair GC-rich sequence element. The general organization of these sequence features is highly unlikely to have occurred by chance, supporting the interpretation that the reported sequence is the pTFI91 plasmid origin of replication.

While the general organization of the pTFI91 sequence elements is quite similar to that found at well characterized plasmid origins of replication, the specific DNA sequence determined lacks significant sequence similarity to known plasmid origin regions. This is true in particular to other sequenced *T. ferrooxidans* plasmids pTF-FC2 and pTF1, which share sequence homology to broad host range plasmids from the IncQ and IncP groups. Thus, it appears that pTFI91 represents a novel type of plasmid replicon.

The structural features identified in pTFI91 strongly suggest that this plasmid contains one or more genes encoding replication initiator proteins which would bind to the direct repeat sequence elements. The most likely location for these genes appears to be the region adjacent to the origin of replication, between the SacI site at pTFI position 1300 and the unique BamHI site. This region appears to be conserved in the related *T. ferrooxidans* plasmids pTFI35 and pDSM583, while the remainder of the plasmid sequence appears to be more variable.

The DNA sequence analysis of pTFI91, together with the Southern hybridization analyses of the related *T. ferrooxidans* plasmids pTFI35 and pDSM583 have identified conserved and variable regions in these plasmids. DNA sequence analysis of a portion of the most conserved sequence indicates that this region

contains the plasmid origin of replication. This information should facilitate the construction of useful plasmid vectors for genetic manipulation of *T. ferrooxidans* strains.

3.4 Identification of a *Thiobacillus* insertion sequence

The 2.2 KB SacI restriction fragment containing the pTFI91 plasmid origin also appears to contain an insertion sequence related to the IS30 family of insertion sequences. This putative *Thiobacillus* insertion sequence will be referred to here as IS3091.

3.4.1 Sequence Homology between IS30 and IS3091

IS element homology can be demonstrated by an extended similarity between a computer translated IS3091 ORF and the transposase coding region of IS30 as well as a high degree of DNA similarity between the pTFI91 DNA sequence and the IS30 element. The homology between these elements is further supported by similarity in the conserved genetic features. Similarity has also been detected between this element and the putative transposase protein of the *Bacteroides* clindamycin resistance element Tn4551, and a sequence from *Spiroides citri* virus, spvv1, indicating that IS30 and IS3091 are members of an extended family of IS sequences. The two other known *Thiobacillus* insertion sequences are not strongly related to the novel IS3091 sequence.

As is further described below, the present sequence of the IS3091 is incomplete and the element is expected to be extended approximately 100 nucleotides beyond the sequenced SacI fragment.

3.4.1.1. Protein Homology

Figure 11, shows the extensive homology between the proteins of IS30 and IS3091 and the computer analysis indicates that overall 124 out of 305 amino acids of IS3091 match IS30. The region of greatest similarity is located at the 3′ end

```
IS30            MRRTITAEEKASVFELWKNGTGFSEITNILGSKPGTIFTMLRDTGGIKPHERKRAVAHLT
IS3091          M--------------------------------GTRYQQLQSE----------------
                *                                ** .  *...

IS30             LSEREEIRAGLSAKMSIRAIATALNRSPSTISREVQRNRGRRYYKAVDANNRANRMAKRP
IS3091           --QRNQIQRGLNEGLSMRAVAKQIGRSPSTVSREVRRGLVGETYDAIQGREEAQRRLRKG
                   .*..*. **.. .*.**.*....*****.****.*.     *.*..... *.*. ..

IS30             KPCLLDQNLPLRKLVLEK-LEMKWSPEQISGWLRRTKPRQKTLRISPETIYKTLYFRSRE
IS3091           VRKLVGGA-PLTNAVTHAILQRKWSPEQVAGRLRMDYPEDKQWRVSHETIYQFIYAHPAG
                  . *.. . ** . * .  *..******..*.**.. * .*  *.*.****. .* .. .

IS30             ALHHLNIQHLRRSHSLRHGRRHTRKGERGTINIVNGTPIHERSRNIDNRRSLGHWEGDLV
IS3091           ELRKALIAALRQGHAKRKPRTRG-KDRRGQLRNMRS--IGERPLEAQDREIPGHWEGDFI
                 .*..  *. **..*. *. * .. *. ** .. ...  * **. . ..*   ******..

IS30             SGT-KNSHIATLVDRKSRYTIILRLRGKDSVSVNQALTDKFLSLPSELRKSLTWDRGMEL
IS3091           KGAFNGSAIGTLVERSSRFVLLVRMEGTDADAALEGFTRRMRTLPKSILRTLTYDQGKEM
                 .*. ..* *.***.*.**.....*. *.*. .. ...* .. .**... ..**.*.*.*.

IS30             ARHLEFTVSTGVKVYFCDPQSPWQRGTNENTNGLIRQYFPKKTCLAQYTQHELDLVAAQL
IS3091           ARHEELERKVGIRIYFADPHSPWQRPTNENTNGLLRQYFPKGTDLSGYSQRRLTQVAED>
                 *** *.. ..*...** **.***** ********.****** * *. *.*. *. **..

IS30             NNRPRKTLKFKTPKEIIERGVALTD
IS3091           -------------------------
```

* match across all sequences
. conservative substitutions

FIGURE 11: CLUSTAL program alignment of IS30 and IS3091 open reading frames

The CLUSTAL (Higgens and Sharp, 1989) program was used to align amino acids with the program parameters set for gap penalties (fixed and variable) of 10. The final alignment was adjusted by hand to rationalize the fit of the last amino acid aspartic acid) in IS3091.

```
 *  match across all sequences
 .  conservative substitutions

Scvirus          M----------------------------------------YSHLSFMDKVKLEQLL
Tn4551           MSKHITEE-------------------------------------------------
IS30             MRRTITAEEKASVFELWKNGTGFSEITNILGSKPGTIFTMLRDTGGIKPHERKRAVAHLT
IS3091           M-----------------------------------------------GTRYQQLQ
                 *

Scvirus          LSKIFLKKNGEQ---NISAIAKYLNRHRSTILREIKRFKTIDEYSAYKSDKMYYEKRKKN
Tn4551           --QRYAISMMLQIPMSKKAIAEAIGVDKSTVYREIKRNCDARSGSYSMELAQRKADRRKQ
IS30             LSEREEIRAGLSAKMSIRAIATALNRSPSTISREVQRN---RGRRYYKAVDANNRANRMA
IS3091           SEQRNQIQRGLNEGLSMRAVAKQIGRSPSTVSREVRRG---LVGETYDAIQGREEAQRRL
                   .            . .*.*. ..   **. **..*          .       ...

Scvirus          NKRFK-FTEEQLNFIHLRFNVYHD---SPSELIYRYFLKF--KVKFPVCVKTLYKWIRLG
Tn4551           QKHRKEVLTPAMRKRII---KLLKKGFSPEQIVGRSRLEGIA----MVSHETIYRWIWED
IS30             KRPKPCLLDQNLPLRKLVLEK-LEMKWSPEQISGWLRRTKPRQKTLRISPETIYKTLYF-
IS3091           RKGVRKLVGGA-PLTNAVTHAILQRKWSPEQVAGRLRMDYPEDKQWRVSHETIYQFIYA-
                 ..      . .             .   **...  .   .        .. .*.*. .

Scvirus          ------FYGFLKQNLRH-HGKKY-KRKGKSDNRGKLTDFKSI------WDIDNKVSNVGW
Tn4551            KRRGGKLHK----YLRR-QGRRYAKRGSKNAGRGFIPGRVD--IDERPEIVELKERFGDL
IS30            -RSREALHHLNIQHLRRSHSLRHGRRHTRKGERGTINIVNGTPIHERSRNIDNRRSLGHW
IS3091           -HPAGELRKALIAALRQGHAKRKPRTRG-KDRRGQLRNMRS--IGERPLEAQDREIPGHW
                       .       **. .. .  .  . .. ** .    .          . .

Scvirus          LEMDTVVGKDHKSAILVLVEQLSKKYFAIKLENHTAREVEKKFKDIIIKNNLIGKIKGII
Tn4551           -EIDTIIGKNHKGAILTINDRATSRVWIRKLSGKEAIPVAKIAVWALRKV--KNLIHTIT
IS30             -EGDLVSGT-KNSHIATLVDRKSRYTIILRLRGKDSVSVNQALTDKFLSLP-SELRKSLT
IS3091           -EGDFIKGAFNGSAIGTLVERSSRFVLLVRMEGTDADAALEGFTRRMRTLP-KSILRTLT
                  * * . *  . . * .. .. ..    .. . ..  . .    . .   .  ....

Scvirus          TDRRKEFSKWREMEIFAETQVYFCDAGSPQQKPLIEYMNSELRHWFPKGTDFNKVSQKQI
Tn4551           ADNGKEFAKHEEIAQKLEIKFYFCKPYHSWERGANENTNGLIRQYIPKGKDFSEVTNKQI
IS30             WDRGMELARHLEFTVSTGVKVYFCDPQSPWQRGTNENTNGLIRQYFPKKTCLAQYTQHEL
IS3091           YDQGKEMARHEELERKVGIRIYFADPHSPWQRPTNENTNGLLRQYFPKGTDLSGYSQRRL
                  *. .*...  *..   ... ** ..  . ..   *  *. .*...** . ..  ... .

Scvirus          DWVVNVINDKLRPCLNWISSKEMFLQNI---------
Tn4551           KWIENKLNNRPRKRLGYLTPNEKFKQIINQNSVAFAS
IS30             DLVAAQLNNRPRKTLKFKTPKE----IIE RGVALTD
IS3091           TQVAED
                 . . .
```

FIGURE 12: CLUSTAL program alignment of IS transposase open reading frames

The putative transposase open reading frame from IS3091 was compared with transposase genes from IS30 and Tn4551. Sequence similarity was also detected between this family of IS elements and an open reading frame found in a virus from *Spiroplasma citri*. The alignments were generated by computer as shown, except for a few minor modifications performed by hand to correct alignment of the N and C-termini of the proteins. Sequence data is from Genbank, accession numbers X51344, M17808, X00792 for sequences *S. c.* virus, Tn4551 and IS30.

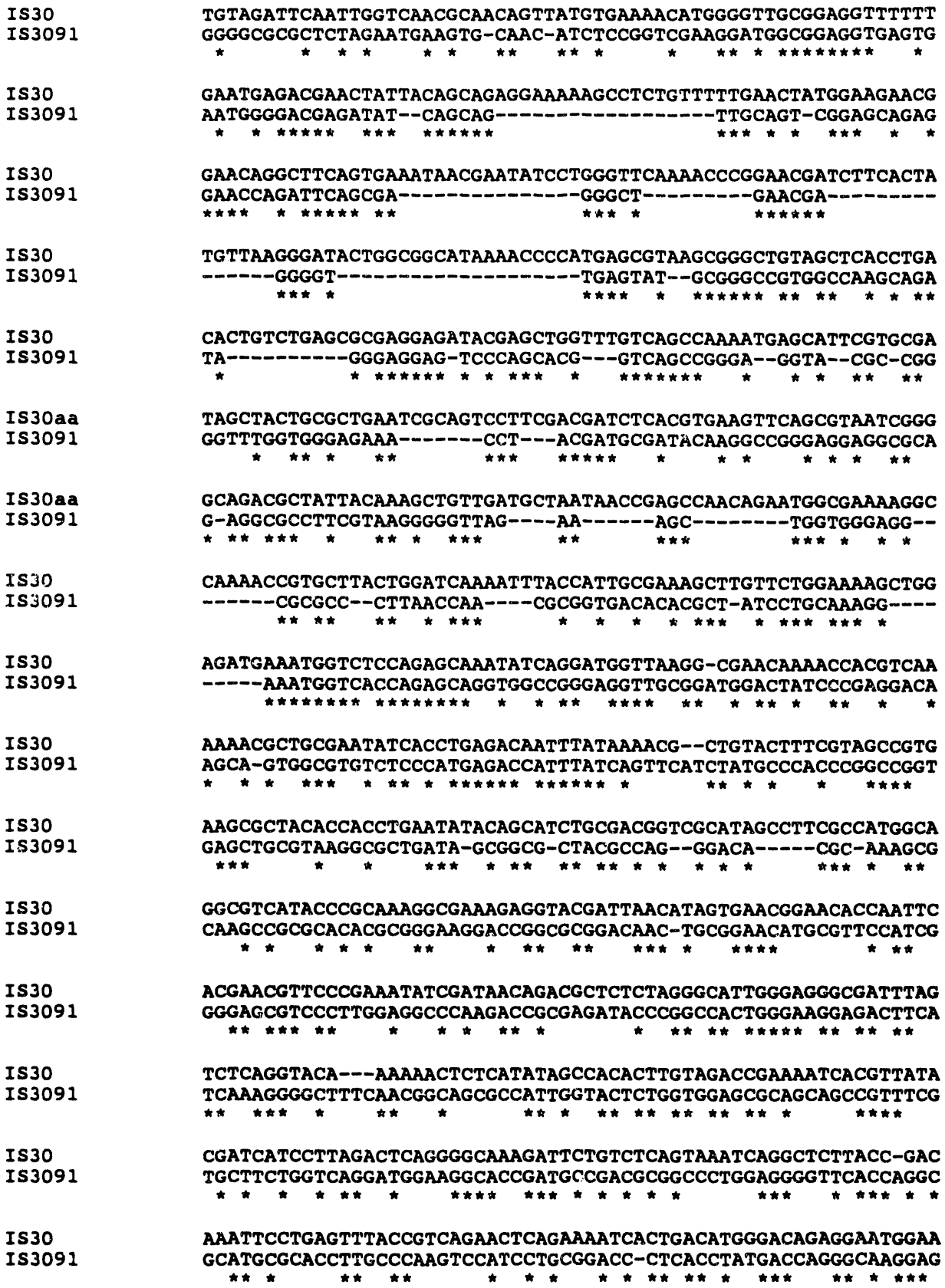


Figure 13: Alignment of IS30 and IS3091 DNA sequence

```
IS30            CTGGCCAGACATCTAGAATTTACTGTCAGCACCGGCGTTAAAGTTTACTTCTGCGATCCT
IS3091          ATGGCACGGCACGAGGAGCTGGAGCGCAAGGTGGGCATCCGTATCTACTTTGCCGACCCG
                 ****  * **    **  *      **     *** *     * *****   *** **

IS30            CAGAGTCCTTGGCAGCGGGGAACAAATGAGAACACAAATGGGCTAATTCGGCAGTACTTT
IS3091          CATAGTCCTTGGCAGCGCCCAACCAACGAGAACACCAATGGTCTCCTGCGCCAGTATTTC
                ** **************   *** ** ******** ***** **  * ** ***** **

IS30            CCTAAAAAAGACATGTCTTGCCCAATATACTCAACATGAACTAGATCTGGTTGCTGCTCAG
IS3091          CCCAAAGGGACGGATTTATCAGGATATTCACAACGCCG-CTTGACGCAGGTG--GCGGA-
                ** ***  ***   * *  *   **** * ****     ** **    * **  **  *

IS30            CTAAACAACAGACCGAGAAAGACACTGAAGTTCAAAACACCGAAAGAGATAATTGAAAGG
IS3091          --AG>>>
                  *

IS30            GGTGTTGCATTGACAGATTGAATCTACA
```

Figure 13: Alignment of IS30 and IS3091 DNA sequence (Continued)

Alignments were performed with the CLUSTAL program, gap weights set at 10 fixed and 10 variable, with minor adjustments in the alignment performed by hand at the 3′ end of the sequences. >>> denotes the end of the IS3091 SacI fragment which was sequenced.

of the ORF, near the probable 3′ end of the transposase protein (24 out of 35 amino acids are identical). One possibly significant difference between the two putative transposases is seen at the 5′ end of their ORF's. The IS30 transposase gene is almost thirty amino acids longer than the IS3091 transposase. The apparent lack of this extensive amino terminal region in IS3091 is probably not the result of a sequencing error in either protein. However, both elements have extensive similarity in the insertion element termini and both ORFs have extensive similarity in their translation start signals. The IS3091 protein N-terminus may resemble the N-terminus of the insertion sequence of Tn4551 (as seen in Figure 12). In addition, visual inspection of alternative reading frames upstream of the putative 3′ end of the IS3091 ORF do not reveal any obvious matches to the N-terminal amino acid sequences of the IS30 ORF. Further sequence similarity searches reveal that IS3091 is part of a more extended family of insertion sequences as shown in Figure 12. To our knowledge, the putative transposase gene in the *Spiroplasma citri* virus was previously unidentified. The sequence similarity between the four putative transposases supports the notion that the 3′ end of IS3091 is within approximately 100 base pairs of the end of the sequenced SacI fragment.

3.4.1.2 DNA Homology

The homology of IS3091 and IS30 can also be detected by comparison of their DNA sequences (Figure 13). Several of this homology can be demonstrated. The general similarity of the sequence is most extensive near the 3′ end of the ORF, where the protein sequence is also highly conserved, and the sequence has significant similarity upstream of the ORF, leading to the location of the IS 3091 terminus and translation initiation signals.

From the alignment it is obvious that the sequence of IS3091 is incomplete. A region of strong similarity is truncated by the end of the SacI fragment. This observation is consistent with the fact that our sequence does not contain the

typical IS elements. The conclusion is that only the upstream sequence (identified below), is included in the Sac I fragment.

It is tempting to speculate that some of the DNA sequence conservation found in the middle regions of the IS30 element is due to some important function of these nucleotides. Genetic elements that function at the level of nucleotide sequence include protein binding sites, regions involved in DNA bending or conserved RNA secondary structures. Function of these hypothetical genetic elements would contribute to the evolutionary pressure to conserve the DNA sequence in addition to the pressure to conserve the derived transposase protein sequence.

3.4.2 Shine Dalgarno sequences of IS30 and IS3091

The putative translation signals of IS30 and IS3091 bear a strong resemblance to each other and to an internal translation start present in Tn4551, leading to the conclusion that the transposase proteins actually initiate at these locations (Figure 14). All three of the putative initiation signals include strong Shine-Dalgarno sequences that might be sequestered in strong RNA secondary structure (data not shown). This type of sequestration is common to many IS elements (Kleckner et al. 1989) and is often associated with protection of the element from activation by outside promoters.

The IS30 and IS3091 ORFs are located very close to the ends of the elements, so only a very short stretch of sequence is available for encoding a transposase promoter. In the case of other IS elements such as IS10, an IS encoded promoter is located such that its transcription initiation site is within the RNA hairpin palindrome, and the resulting mRNA does not form the secondary structure which sequesters the translation initiation site. The result is that transcripts initiated under the control of the element's own genetic machinery determine the frequency and timing of transposase expression and the resulting transposition event.

A: Termini

```
Tn4551 CTTGAGTTCAACTTATAAATGCAACT TTTTGGGTGCGGATAATAAGCAATAA
IS30   TGTAGATTCAATTG     GT CAACAGTTATGAAAACATGGGGTTGCGGAGGT
IS3091 TCTAGAATGAA        GTGCAACA TCTCCGGTCGA AGGATGGCGGAGGT
```

B: Translation control

```
Tn4551 AGCAGATTGTCGGCAGGCAGGAGCCGCTT  GGAGGGAATTGCGATGGTATCTC
IS30   TCACGCAACAGTTATGAAAACATGGGGTTGCGGAGGTTTTTTGAATGA GACGAA
IS3091 TGAAGTGCAACATCTCCGGTCGAAGGATGGCGGAGGTG AGTGAATGGGGACGAG
```

Figure 14: Alignment of the transposon termini and the putative translation control regions of Tn4551, IS30 and IS3091

Alignments were performed by hand, important similarities are shaded. A: comparison of putative IS termini. The putative terminus of IS3091 begins at bp 1210 of the SacI fragment sequence. B: putative Shine Delgarno sequences (italicized in both A and B) begin at bp number 501, 43, and 1248 of the sequences for Tn4551, IS30 and IS3091 (SacI fragment), respectively. Sequences possibly involved in RNA secondary structure are underlined.

3.4.3 Preliminary Identification of the 5′ IS3091 Terminus

The IS3091 upstream insertion sequence terminus has been tentatively identified based on similarity to IS30 and Tn4551 (See Figure 14). The IS30 element utilizes the insertion sequence termini when transposing to a new DNA location, generating a 2 base pair repeat in the process. Unlike most transposons, the IS30 repeats generated by transposition do not appear to be part of the original sequence at the location of the element's new insertion. A more detailed discussion of the termini must await completion of the element's sequence to confirm the model.

3.4.4 Discussion of 1S3091

A novel *Thiobacillus* insertion element which is related by sequence and structural homology to the IS30 family of insertion sequences has been identified. Previous reports have described IS30 as a replicative transposable element which lacks a resolvase function. IS30 transposition apparently relies on host cell recombination functions to resolve cointegrate structures formed by the replicative transposition event (Caspers, et al., 1984). In this report, the IS30 family of insertion sequence has been briefly described. The data from the alignments of the protein, DNA, translation control signal and transposon termini sequences provides a rich source of information to direct future research on these interesting and potentially useful genetic elements.

The existence of an insertion sequence in the Sac I fragment clearly provides a boundary to the plasmid DNA replication origin (described previously). Parts of the plasmid's minimal replicon (for example, protein encoding open reading frames) may exist downstream of the insertion sequence. It seems likely that the large cis-acting DNA sequence normally

associated with plasmid DNA replication initiation sites would not tolerate the insertion of such a massive fragment of DNA sequence.

3.5 Molecular Genetics of Mercury Resistance in *T. ferrooxidans*

Development of a useful genetic engineering system for *T. ferrooxidans* requires a functional marker gene to allow the isolation of recombinant clones. Mercury resistance was chosen as a potential marker for three reasons: (1) cloned mercury resistance genes have been used as selectable markers in other bacteria; (2) naturally resistant *T. ferrooxidans* strains exist, demonstrating that the mercury resistance trait can function in *Thiobacillus* species growing at low pH (Shiratori et al., 1989); and (3) mercury is naturally present in some of the environments where recombinant *Thiobacillus* may be used, making manipulation of the mercury resistance trait a potential way to improve the utility of engineered *Thiobacillus* strains.

During this program a series of experiments was performed to analyze the mercury resistance determinants of *T. ferrooxidans* strains and to clone the mercury resistance genes. These included DNA hybridization studies to identify DNA sequences encoding mercury resistance functions in three *T. ferrooxidans* strains, analysis of induction of the mercury resistance trait, and cloning of the mercury resistance genes. Results of this work is summarized in the following section of the report.

3.5.1 Southern Blot Hybridization of *Thiobacillus* Chromosomal DNA Using a Tn501 *mer* Gene Probe

As described in the section on *T. ferrooxidans* strain purification and characterization, testing for sensitivity to heavy metals indicated that the strain DSM583 contained genes encoding resistance to mercury. The DSM583 *mer* genes were identified by their homology to the *mer* genes carried on Tn501. A radioactively labeled probe was made from a DNA fragment

encoding most of the Tn501 *mer* operon. This fragment was isolated from the plasmid pRSF1010::Tn501. There was strong hybridization of this probe to control DNA samples and to genomic DNA from the DSM583 strain, but reduced hybridization to negative control samples.

Chromosomal mercury resistance genes also were detected in the *Thiobacillus* strains TFI92, and TFI70 by Southern hybridization analysis of total genomic DNA (see 10th Quarterly Report; December 13, 1991). The use of the mercury resistance trait as a selectable marker will depend on the identification of a suitable mercury sensitive host which could be accomplished by screening potential host strains by hybridization analysis using a *mer* gene probe.

In addition to Southern hybridization analysis of chromosomal DNA fragments, hybridization conditions which reproducibly detected the presence of *mer* genes in colonies on nitrocellulose filters also were determined. This method was used to screen a recombinant library of DSM583 genomic DNA to identify *mer* gene clones.

3.5.2 Induction of Mercury Resistance in *Thiobacillus* Strains DSM583, TFI70 and TFI92

Three of the *T. ferrooxidans* strains used in this program were found to contain mercury resistance genes whose expression was induced by the presence of mercury in the culture medium. In experiments where 10^7 cells of *T. ferrooxidans* strains TFI70, TFI92, or DSM 583 were plated onto TSM plates containing 1.0-2.0 μg per ml $HgCl_2$, mercury resistant colonies were routinely recovered at a frequency of approximately 10^{-5}. *Thiobacillus* cells from these colonies were resistant to 4 μg/ml $HgCl_2$, indicating that these cells were adapted to grow in the presence of Hg^{++}. These mercury resistant colonies account for the background growth observed in transformation experiments described below. Subsequent experiments indicated that these three strains can be induced to growth on Hg^{++} by first culturing them on a sub-lethal dose of Hg^{++}. In addition, Southern

hybridization analysis of chromosomal DNA from these three strains, described above, demonstrated the presence of sequences with homology to the mercury resistance operon of Tn501.

3.5.3 Cloning of the *Thiobacillus mer* Operon of DSM583

The mercury resistance genes of *T. ferrooxidans* strain DSM583 were cloned by constructing a genomic library from this strain followed by screening with a Tn501 *mer* operon probe. Results of this work is summarized in this section.

3.5.3.1 Construction of the DSM583 Library

In order to clone the mercury resistance genes from DSM583 a genomic library was constructed and screened to isolate the DNA fragment encoding the mercury resistance gene cluster. Genomic DSM583 DNA was partially digested with PstI and restriction fragments of approximately 4 to 12 kb were isolated. These fragments were cloned into the pBluescript™ II SK vector and transformed into an *E. coli* host (DH10B). Over 3000 clones were transferred to microtiter dishes for screening and also to allow storage at -80°C. This library was replicated onto nitrocellulose filters for hybridization screening of *mer* gene clones. In addition, a plasmid preparation representative of over 10,000 recombinant clones was prepared and stored. Finally, a second library was constructed using a Sau3AI partial digest of *DSM583* DNA and 1500 clones from this library were isolated and stored.

The DSM583 library was screened directly for mercury resistance by replicating the *E. coli* transformants from the DSM583 PstI library onto selective media containing 10, 15 and 20 μg per ml of $HgCl_2$. Approximately 40 colonies exhibited some growth in the initial screening, but none were found to be reproducibly resistant to these mercury concentrations. Dot blot analysis of these putative mercury resistant clones failed to demonstrate homology to the Tn501 *mer* gene. Clones with significant homology to the Tn501 *mer* operon were identified in

subsequent hybridization experiments. These clones were found to confer resistance to lower levels of mercury after pre-induction by exposure to sub-lethal doses of mercury. These observations explain why *mer* gene clones were not detected by the preliminary plating experiments.

3.5.3.2 Identification of Recombinant Clones Containing the *mer* Operon of DSM583

Filter replicas of the DSM583 chromosomal DNA library were screened by hybridization using the Tn501 mer gene probe. Eighteen possible clones were identified in the primary screen, and plasmid DNA was prepared from each. Restriction analysis and Southern blotting identified four clones with good homology to the *mer* gene probe. Figure 15 shows an example of an autoradiogram identifying the *Thiobacillus mer* gene fragment from three of the clones. A more complete restriction analysis of these clones was performed and the restriction map of the cloned PstI fragment is shown in Figure 16.

Plasmids from two of these clones were re-transformed into *E. coli*, and the recombinant clones were tested for mercury resistance. Results indicated that both of these clones confer a moderate degree of resistance to mercury (2.5 to 5.0 μg/ml).

The *T. ferrooxidans* strain DSM583 *mer* genes are apparently inducible in *E. coli*; liquid (LB) cultures of the *E. coli* strain DH10B transformed with clone number 65-27 was grown with 0.1 μg/ml $HgCl_2$. These pre-induced cultures could then grow at higher concentrations of mercury (up to 4 to 5 μg/ml). This result may indicate that the clones contain a large portion of the *mer* gene cluster, including mercury responsive regulatory elements (Inoue et al., 1991).

The published sequence of a *Thiobacillus mer* gene predicts a restriction map which is different than the map of our cloned *mer* genes (Inoue et al., 1989), indicating that substantial and perhaps important differences exist between the two cloned versions.

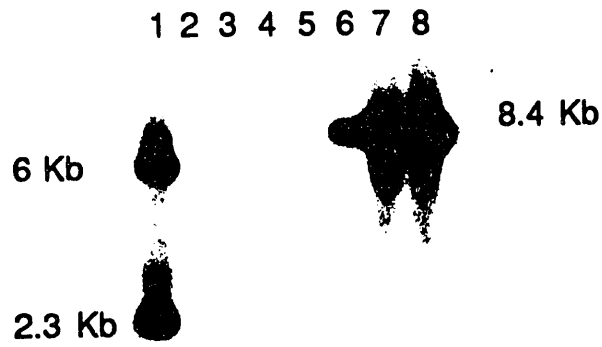


Figure 15: Southern hybridization analysis of DSM583 *mer* gene clones

This figure shows the result of Southern hybridization analysis of potential DSM583 *mer* gene clones in the pBluescript™ II SK vector (4 hr exposure). The blot had been probed with a radiolabeled probe isolated from Tn501. The samples are: lane 1, RSF1010::Tn501 + EcoRI; lane 2, DSM583 genomic DNA + PstI; lane 3, complete DSM583 genomic library in pBluescript™ II SK + PstI; lane 4, pBluescript™ II SK + PstI; lane 5, DSM clone #5-47 + PstI; lane 6, DSM583 clone #44-28 + PstI; lane 7, DSM583 clone #65-27 + PstI; lane 8, DSM583 clone #17-8 + PstI. Hybridization of the probe with 8.4 kb PstI fragments present in clones 44-28, 65-27, and 17-8 is clearly visible. The probe also was found to hybridize to a similar size band present in PstI digests of DSM583 genomic DNA (lane 2) and the complete DSM583 genomic DNA in the pBluescript™ II SK vector (lane 3) when the same blot was exposed for 16 hrs.

Restriction map of the 8.4 Kb PstI fragment
containing the *Thiobacillus* ***mer*** gene from DSM583

Figure 16: Restriction map of the *T. ferrooxidans* DSM583 *mer* Gene clone

This figure shows the restriction map of the 8.4 kb PstI fragment from DSM583 present in clone no. 65-27. The restriction site positions indicated are measured from the PstI site on the left of the diagram. The 8.4 kb PstI fragment was not cut by the enzymes HindIII, SacI, or XbaI.

3.6 Electroporation of *Thiobacillus ferrooxidans*

Electroporation is likely to be the best method for transforming *T. ferrooxidans* with plasmid vectors. Experiments were performed using various electroporation conditions, DNA concentrations and plating media in an attempt to demonstrate transformation of *Thiobacillus*. Plasmids used included the putative chloramphenicol resistant shuttle vectors described previously and a version of the RSF1010 plasmid with genes encoding mercury resistance. At the time this work was initiated, *T. ferrooxidans* had not been transformed by electroporation or any other reported technique. The general approach detailed below was to establish growth and harvesting conditions for preparing *T. ferrooxidans* for electroporation, establish the proper range of voltage and pulse length for electroporating the cells, and to identify plating conditions for selection of transformants. The work concluded with attempts to transform specific plasmids into the strains. Progress was made in determining the electroporation conditions, plating techniques and in techniques to analyze putative transformants, but demonstration of stable transformation awaits the development of a suitable host strain/plasmid combination.

3.6.1 Growing and Harvesting Cells for Electroporation

Two crucial variables in the electroporation of cells are the conditions under which the cells are grown and the conditions under which they are harvested. During these experiments the effectiveness of the cell harvest and preparation methods were determined. Iron III is precipitated in *Thiobacillus* culture media (TSM) as the bacteria oxidize the $FeSO_4$. This precipitate must be separated from the cells during harvesting since it would interfere with the electrophoresis. Also relatively high cell densities are desired for electroporation studies. Different ways to harvest TFI70 were investigated. It was found that much of the iron precipitate could be removed by simply allowing it to settle undisturbed for

5 minutes. The culture was then carefully decanted and harvested as before, with only a minimal number of dilute acid washing steps.

The results indicate that greater than 90% of the starting cells could be recovered using this technique. Overall, the preparation of cells for electroporation has no significant effect on cell survivability. Cell suspensions used for electroporation (in 300 mM sucrose) were at a concentration of approximately 5 X 10^{10} cells/ml.

Efficient electroporation of bacteria is usually achieved with cells harvested at the mid-log growth phase. Our initial experiments over-estimated the generation times of *T. ferrooxidans,* so that overly mature cultures were used. Growth conditions for the culture should be further developed to improve the yield of mid-log phase cells for electroporation. One of the conclusions of this study is that the media and growth conditions tested need to be improved so that high yields of mid-log phase cells can be obtained. It is likely that such cells would be susceptible to electroporation at more moderate voltages and pulse lengths.

3.6.2 General Electroporation Conditions

A series of experiments was performed to define the conditions (DNA concentration, voltage, and pulse length) for the electroporation of TFI70 and TFI91.

In the absence of a proven shuttle vector, the utility of the electroporation parameters was assessed by measuring cell viability. Because the electroporation conditions which allow DNA to enter cells also correlate with the loss of viability of a portion of the cell population, cell survival is a useful test of the potential induction of pores in the cell membranes. The electroporation of DNA into a cell is generally associated with a reduction in cell survival of approximately 50%.

As a second method for monitoring electroporation, FITC dextran (a fluorescent compound) was tested as a model for plasmid DNA. The FITC dextran used was about the same size as a

small plasmid. Cells were electroporated in the presence of FITC dextran and examined under fluorescent microscopy to determine if uptake of the fluorescent compound occurred.

Early experiments demonstrated that the *Thiobacillus* cells being used were highly resistant to electroporation. Cells were electroporated in the presence and absence of DNA under conditions in which control *E.coli* cells were efficiently transformed with common broad host range vectors such as RSF1010 and with the pBTL series of *Thiobacillus/E coli* putative shuttle vectors. Under these conditions, no *Thiobacillus* transformants were detected or selected, and the *Thiobacillus* cells showed undiminished viability. The lack of significant killing at long pulse lengths has been shown to indicate that the field strength was too low for efficient electroporation.

The absence of any loss of cell viability suggested that a more powerful electroporation device was needed for effective transformation of *T. ferrooxidans*. Using a second electroporator, field strengths of 14 to 24 kv/cm were tested. While the cultures exposed to 22 and 24 kv/cm appeared to be growing less well 5 hours after electroporation, no difference in survival was apparent after longer incubation. Additional experiments were performed in which the electric field strength was further increased to up to 32 kv. In this higher range, cell viability was significantly reduced, with over 90% of the cells killed at voltages of 30 to 32k v. Between 24 and 30 kv, the degree of cell killing also depended on the pulse length, and conditions could be found in which approximately 50% of the cells were killed. In addition, fluorescence studies indicated that at high field strengths, the *Thiobacillus* cells were able to take up high molecular weight FITC dextran molecules. As mentioned in the previous section, optimal electroporation conditions will vary depending on the method of cell preparation.

3.6.3 Conditions for Transformation Selection

To select for transformed cells by either antibiotic or mercury resistance, the sensitivity of the potential host strain and the level of expected resistance provided by the cloned marker gene must be known. The sensitivity of strains TFI70, TF192 and DSM583 to mercury and chloramphenicol were assessed. In general, the sensitivity of *Thiobacillus* to these agents depends on whether the cells are cultured on plates or liquid media, on the number of cells and for inoculation and on the strain being tested.

The sensitivity of *Thiobacillus* strains DSM583, TFI92 and TFI70 was tested, and the results were used to define the range of mercury or chloramphenicol concentrations used for selection of transformed *Thiobacillus*. Potentially transformed cells were plated at a series of cell densities onto TSM solid medium containing a range of antibiotic concentrations. Cell densities ranged from 10^4 to 10^7 bacteria per plate. The results are summarized in the following sections.

3.6.4 Electroporation of *Thiobacillus* with RSF1010/Tn501

The possibility that the broad host range plasmid RSF1010 was capable of replicating in *Thiobacillus* was tested. An RSF1010 plasmid bearing the mercury resistance genes of transposon Tn501 was electroporated into *Thiobacillus* strain TFI70. Mercury resistance colonies were obtained using the TSM medium containing 0.5, 1 and 2 μg per ml $HgCl_2$. Individual colonies were cored from the selective plates and cultured in liquid media containing mercury. Potential transformants were tested for the presence of recombinant plasmid by restriction analysis, by Southern hybridization and by transformation of competent *E. coli* with *Thiobacillus* DNA preparations from the selected colonies. The results of this experiment were inconclusive due to the presence of inducible *mer* genes in the host cells.

As previously discussed, the strain DSM583 initially appeared to have a higher resistance to mercury than the *Thiobacillus* strains TFI92 and TFI70, with maximum tolerances reaching up to 4 μg per ml $HgCl_2$ for the DSM583 but less than 1 μg per ml $HgCl_2$ for the other two strains. As work on *Thiobacillus* mercury resistance progressed, it was determined that all of the strains being used had inducible chromosomal *mer* genes, limiting utility of the mercury resistance trait as a selectable marker in these particular *Thiobacillus* strains.

3.6.5 Electroporation with the BTL series of vectors

Four potential *Thiobacillus/E. coli* shuttle vectors (pBTL107, pBTL201, pBTL305, and pBTL501) were electroporated into the host vector TFI92. The plasmids tested represent those which insert the Co1E1 plasmid origin and CM selectable marker into each of four locations in the *Thiobacillus* plasmid with little or no deletion of the *Thiobacillus* sequences.

Solid media containing 150, 200 and 250 μg per ml of CM was used as a selection for *Thiobacillus* transformants resistant to this antibiotic. Severe growth inhibition of *Thiobacillus* colonies was seen when up to 10^7 electroporated cells were plated. Small *Thiobacillus* colonies present on CM plates were used to inoculate liquid media containing 50 to 100 μg per ml of CM.

As many as 100 slow growing colonies per plate were observed in experiments where the shuttle vectors pBTL107, pBTL201 and pBTL305 were electroporated into cells, but 5 to 10 colonies were seen when plasmid pBTL501 or the minus DNA control electroporations were plated.

DNA samples from these colonies were prepared and analyzed as above. No stable transformants were observed, possibly because the difference in chloramphenicol resistance between transformed and untransformed cells is too low for efficient selection under these conditions.

The electroporation experiment was repeated with pBTL plasmid DNA samples prepared from *dam, dcm E. coli* cells to

reduce the possibility of *Thiobacillus* host restriction of methylated DNA (by analogy to the *E. coli mcr* restriction system).

From the analysis of the putative transformants, it was found that TFI92 had a plasmid that was not detected in the initial plasmid screening effort. This plasmid is virtually identical to the plasmid from TFI91, the plasmid used in the construction of the BTL series of shuttle vectors.

The presence of a plasmid endogenous to TFI92 complicates the transformation experiment in two ways: first, detection of transformed shuttle vector DNA is difficult in strains bearing a plasmid with a somewhat similar restriction pattern. More importantly however, establishment of a plasmid transformed into a strain with a resident, incompatible plasmid is expected to be more difficult than transformation into a strain not bearing such a plasmid. One possible solution is to use a laboratory strain of *Thiobacillus* which does not have the plasmid, or to use a broad host range plasmid which would not be incompatible with the resident plasmid.

Because the relatively high resistance of the *T. ferrooxidans* strains used to chloramphenicol, it was concluded that either a different genetic marker is needed to select for transformants or different, milder, growth conditions are needed.

3.6.6 Evaluation of *Thiobacillus cuprinus* as a Potential Host for Electroporation

The *T. cuprinus* strain was found to be sensitive to both chloramphenicol and ampicillin at concentrations above 5 μg per ml. Resistance to either marker may be useful to select for cells electrotransformed with drug resistance plasmids. Southern hybridization analysis of *T. cuprinus* DNA preparations did not reveal the presence of any endogenous plasmids.

T. cuprinus has been successfully grown to densities exceeding 1 x 10^8 cells per ml in SY 64 media with elemental sulfur as the primary energy source. The composition of the medium is 0.8 g $(NH_4)_2SO_4$, 0.2 g KH_2PO_4, 0.08 g $MgSO_4$ $7H_2O$, and

0.01 g yeast extract (Difco) per liter. The pH of the media is initially adjusted to 5.5 with sulfuric acid, but growth of the bacteria lowers the pH to about 2.5 at saturation.

T. cuprinus appears to be an ideal host for transformation experiments, since it is closely related to *ferrooxidans,* is easy to grow in liquid culture, and has no endogenous plasmids or interfering drug resistances. T *Cuprinus,* however, does not grow well on agar media and this limited the scope of the electroporation experiments. A logical next step would be to improve plating techniques for this potentially useful strain.

4.0 CONCLUSIONS

This report has summarized the results from the program entitled Molecular Biological Enhancement of Coal Biodesulfurization. The most substantial progress made in this program was in the area of *Thiobacillus* molecular biology, specifically in the area of vector development. At the initiation of this program, the basic elements needed for performing genetic engineering in *T. ferrooxidans* were either not yet developed or unavailable to us. Improved techniques have been described which will make it easier to construct and analyze the genetic structure and metabolism of recombinant *T. ferrooxidans*. Major accomplishments of the program are summarized below:

The metabolism of the model organic sulfur compound DBT by certain heterotrophic bacteria was confirmed and characterized. Techniques were developed to analyze the metabolites of DBT, so that individual "4S" pathway metabolites could be distinguished. These techniques are expected to be valuable when engineering organic sulfur metabolism in *Thiobacillus*.

Strain isolation techniques were used to develop pure cultures of *T. ferrooxidans* for further study. Seven of these strains of *T. ferrooxidans* were assessed as potential recombinant hosts. The mixotrophic strain *T. cuprinus* was also characterized

for potential use as an electroporation host.

Several fundamental contributions were made toward the establishment of genetic engineering for *T. ferrooxidans*. A family of related *Thiobacillus* plasmids was discovered in the seven strains of *T. ferrooxidans* mentioned above. One of these plasmids, pTFI91, was cloned into a pUC-based plasmid vector, allowing it to propagate in *E. coli*. A key portion of the cloned plasmid was sequenced. This segment, which is conserved in all of the related plasmids characterized, contains the vegetative origin of DNA replication, and fortuitously, a novel insertion sequence, designated IS3091. The sequence of the DNA origin revealed that these *Thiobacillus* plasmids represent a unique class of replicons not previously described. The potentially useful insertion sequence IS3091 was identified as a new member of a previously undefined family of insertion sequences which include the *E. coli* element IS30.

The experiments performed to develop the technique of electroporation for *T. ferrooxidans* set defined conditions needed for electroporation of these recalcitrant bacteria by identifying the criteria for a suitable host/vector/selection combination. Finally the cloned *mer* genes from *T. ferrooxidans* strain DSM583 should be valuable not only as a potential genetic marker for *T. ferrooxidans* transformation, but also as a source of regulatory elements for expression of foreign genes in this species.

5.0 REFERENCES

Bramhill, D. and A. Kornberg. 1988a. "Duplex opening by DnaA protein at novel sequences in the initiation of replication at the origin of the *E. coli* chromosome". Cell 52:743-755.

Bramhill, D. and A. Kornberg. 1988b. "A model for initiation at origins of DNA replication". Cell 54:915-918.

Caspers, P., B. Dalrymple, S. Iida, and W. Arber. 1984. "IS30 a new insertion sequence of *Escherichia coli* K12". Mol. Gen. Genet. 196:68-73.

Dorrington R. A. and D. E. Rawlings. 1990. "Characterization of the minimum replicon of the broad-host-range plasmid pTF-FC2 and similarity between pTF-FC2 and the IncQ plasmids". J. Bacteriol. 172:5697-5705.

Drolet, M., P. Zanga, and P.C.K. Lau. 1990. "The mobilization and origin of transfer regions of a *Thiobacillus ferrooxidans* plasmid: relatedness to RSF1010 and pSC101. "Mol. Microbiol. 4(8):1381-1391.

Dutt, D., P. Saadi, S. Ciatti and S. Krawiec. 1988. "Isolation of bacteria with the ability to desulfurize dibenzothiophene to o,o'-Biphenol. Abstr 88th Gen Mtg. Am. Soc. Microbial., K112, p. 225.

Fry, I.J. R.E. Wyza, T.J. Zupancic, D.T. Palmer, J.H. Litchfield and O.H. Tuovinen 1990. "Molecular Genetic Analysis of Desulfurizing Bacteria". Sixth Annual Coal Preparation, Utilization, and Environmental Control Contractors Conference Proceedings. Pittsburgh, PA. pp. 217-225.

Fry, I.J. and E.G. Garcia. 1989. "Isolation and cloning of an endogenous plasmid from Thiobacillu ferrooxidans". Abst. 89th Gen. Mtg. Am. Soc. Microbiol., J18, p. 244.

Fry, I.J. O. Delgado, D. Palmer, R.E. Wyza, T.J. Zupancic, J.H. Litchfield and O.H. Tuovinen 1991a. "Toxicity of Antibiotics. metal ions and dibenzothiophene. Metabolites in Thiobacillus ferrooxidans" Abstr. 91st Gen. Mtg. Am. Soc. Microbiol. Q-20-p. 279.

Fry, I.J. C.G. Traunero., D.T. Palmer, T.J. Zupancic, R.E. Wyza, J.H. Litchfield, L. Chakravarty, O. Delgado and O.H. Tuovinen, 1991b". Molecular genetic toxicological analysis of pyritic sulfur oxidizing bacteria". Seventh Annual Coal Prepration Utilization, and Enviornmental Control Contractors Conference Proceedings. Pittsburgh, PA. pp. 361-367

Fuller, R.S., B.E. Funnell, and A. Kornberg. 1984. "The DnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites". Cell 38:889-900.

Gibbs, H.D. 1927. "Phenol tests. III. The indophenol test", J. Biol. Chem., 72:649-664.

Hanna, M.L. and Taylor R.T. 1989. "Radioactive Electrophoretic assay of adenosine-5-triphosphate sulfatase activity in crude extracts with sulfate or selenate as a substrate". Anal. Biochem. 176:3294-302.

Higgins, D.G. and Sharp, P.M. (1989). "Fast and sensitive multiple sequence alignments on a microcomputer". CABIOS, 5 (2):51-153.

Holmes, D.S., J.H. Lobos, L.H. Bopp, and G.C. Welch. 1984. "Cloning of a *Thiobacillus ferrooxidans* plasmid in *Escherichia coli*". J. Bacteriol. 157:324-326.

Hou, C.T. and A.I. Laskin, 1976. "Microbial Conversion of Dibenzothiophene". Develop. Ind. Microbiol., 17:351-362.

Huber, H., and K.O. Stetter. 1990. "*Thiobacillus cuprinus* sp. nov., a novel facultative organotrophic metal-mobilizing bacterium". Appl. Environ. Microbiol. 56:315-322.

Inoue, C., K. Sugawara, and T. Kusano. 1991. "The *merR* Regulatory gene in *Thiobacillus ferrooxidans* is spaced apart from the *mer* structural genes". Mol. Microbiol. 5:2707-2718.

Inoue, C., K. Sugawara, T. Shiratori, T. Kusano, and Y. Kitagawa. 1989. "Nucleotide sequence of the *Thiobacillus ferrooxidans* chromosomal gene encoding mercuric reductase". Gene. 84:47-54.

Isbister, J.D and R.C. Doyle 1985. "A novel mutant microorganism and its use in removing organic sulfur from coal". European Patent Appl. 85112450.302.10.85, Publ. No. 0218734.

Jin, S.M., W.M. Yan, and Z.N. Wang. 1992. "Transfer of IncP plasmids to extremely acidophilic *Thiobacillus thiooxidans*". Appl. Environ. Microbiol. 58:429-430.

Kilbane, J.J. 1989. "Desulfurization of coal: the microbial solution". TIBTECH, 97-101.

Kilbane, J.J and B.A. Bielaga 1989. "Molecular biological enhancement of coal biodesulfurization". Abstr. Fifth Annual Coal Preparation Utilization and Environmental Control Contractors Conference. Pittsburgh, PA July 31 - August 3, 1989.

Kleckner, N. 1989. In: Mobile Genetic Elements, ed. Shapiro, J. Academic Press, Orlando, FL.

Kodama, K. 1977. "Induction of dibenzothiophene oxidation by Pseudomonas jianii". Agric. Biol. Chem. 41:1193-1197.

Kodoma, K.K. Umehara, K. Shimizu, S. Nakatani, Y. Minoda and K. Yamada. 1973. "Identification of microbial products from dibenzothiophene and its proposed oxidation pathway". Agric. Biol. Chem. 37:45-50.

Kodama, K., S. Nakatani, K. Umehara, K. Shimizu, Y. Minoda and K. Yamada. 1970. "Microbial conversion of petro-sulfur compounds Part iii. Isolation and identification of products from dibenzothiophene". Agric. Biol. Chem 34:1320-1324.

Krawiec, S. 1990. "Bacterial desulfurization of thiophenes: screening techniques and some speculations regarding the biochemical and genetic bases". Develop. Ind. Microbiol. 31:103-114.

Kues, U., and U. Stahl. 1989. "Replication of plasmids in Gram-negative bacteria". Microbiol. Rev. 53:491-516.

Larsson, L. 1989. Licentiate thesis, Department of Biotechnology, University of Lund, Sweden.

Martin, P.A., P.R. Dugan and O.H. Tuovinen, 1981. "Plasmid DNA in Acidophilic, Chemolithotrophic Thiobacilli". Can. J. Microbiol. 27:850-853.

Martin, P.A.W., P.R. Dugan and O.H. Tuovinen. 1981. "Plasmid DNA in acidophilic, chemolithotrophic thiobacilli". Can. J. Microbiol. 27:850-853.

Monticello, D.J. D. Bakker and W.R. Finnerty. 1985. "Plasmid-mediated degradation of dibenzothiophene by Pseudomonas species". Appl. Environ. Microbiol. 49:756-760.

Mormile, M.R. and Atlas, R.M. 1988. "Mineralization of the dibenzothiophene biodegradation products 3-hydroxy-2-formyl benzothiophene and dibenzothiophene sulfone". Appl. Environ. Microbiol., 54:3183-3184.

O'Connor, M. and M.H. Malamy. 1983. "A new insertion sequence, IS121, is found on the Mu dI1(Ap lac) bacteriophage and the *Escherichia coli* K-12 chromosome". J. Bacteriol. 156:669-679.

Sanchez, H., E. Hevia, B. Caceres, and A. Venegas. 1986. "Studies on native strains of *Thiobacillus ferrooxidans*. Biotechnol. Appl. Biochem. 8:300-308.

Scholz, P., V. Haring, B. Wittmann-Liebold, K. Ashman, M. Bagdasarian, and E. Scherzinger. 1989. "Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010". Gene 75:271-288.

Shiratori, T., C. Inoue, K. Sugawara, T. Kusano, and Y. Kitagawa. 1989. Cloning and expression of *Thiobacillus ferrooxidans* mercury ion resistance genes in *Escherichia coli*. J. Bacteriol. 171: 3458-3464.

Silverman, M.P. and D.G. Lundgren, 1959. "Studies on the chemoautotrophic iron bacterium *Ferrobacillus ferrooxidans*. I. An improved medium and harvesting procedures for securing high cell yields. " J. Bacteriol 77:642-647.

Smith, A.W. and B.H. Iglewski. 1989. "Transformation of Pseudomonas aeruginosa by electroporation". Nucleic Acids Res. 17:10509.

Smith, C.J. 1987. "Nucleotide sequence analysis of Tn4551: use of ermFS operon fusions to detect promoter activity in *Bacteroides fragilis*". J Bacteriol. 169:4589-4596.

Stanier, R.Y., N.J. Palleroni and M. Doudoroff. 1966. The aerobic pseudomonads: a taxonomic study". J. Gen. Microbiol 43:159-271.

Stoner, D.L., J.E. Wey, K.B. Barett, J.G. Jolley, R.B. Wright and P.R. Dugan. 1990. "Modification of water-soluble coal-derived products by dibenzothiophene-degrading microorganisms". Appl. Environ. Microbiol. 56:2667-2676.

Tsai, M., Y.H. Florence Fu, and R.C. Deonier. 1990. "Intrinsic bends and integration host factor binding at F plasmid *oriT*". J. Bacteriol. 172:4603-4609.

Visca P., E. Bianchi, M. Polidoro, V. Buonfiglio, P. Valenti, and N. Orsi. et. al., 1989. "A new solid medium for isolating and enumerating *Thiobacillus ferrooxidans*". J. Gen. Appl. Microbiol. 35:71-81.

Wirth, R.A. Friesenegger and S. Fiedler, 1989. "Transformtion of various species of gram-negative bacteria belonging to 11 different genera by electroporation". Mol. Gen. Genet. 216:175-177.

Wyza, R.E. and J.D. Isbister 1989. "Enhancement of coal desulfurization". Abstr. Fifth Annual Coal Preparation, Utilization and Environmental Control Contractors Conference. Pittsburgh, PA. July 31 - August 3, 1989.

Yates, J.R. and D.S. Holmes. 1987. "Two families of repeated DNA sequences in *Thiobacillus ferrooxidans*". J. Bacteriol. 169: 1861-1870.

Yates, J.R., R.P. Cunningham, and D.S. Holmes. 1988. "IST2: An insertion sequence from *Thiobacillus ferrooxidans*". Proc. Natl. Acad. Sci. USA. 85:7284-7287.

Zupancic T.J., L. Chakravarty, B. Baker, J.D. Kittle, I. Fry and O.H. Tuovinen 1992. "Molecular analysis of plasmids from *Thiobacillus ferrooxidans*." Abstr. 92nd Gen. Mtg. Am. Soc. Microbiol. H-121, p 203.

END

DATE FILMED

2 / 19 / 93