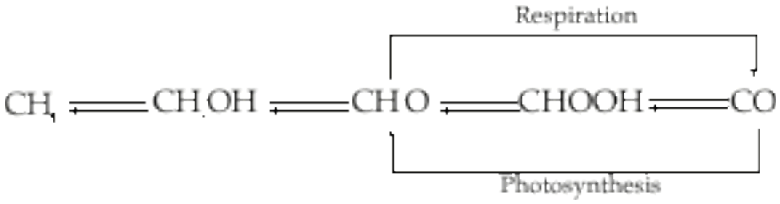
Chapter Two

Heterologous Expression of the *Methylococcus capsulatus* (Bath) Soluble Methane Monooxygenase Hydroxylase

Introduction

Carbon atoms, along with other elements necessary for the chemistry of life, circulate in the biosphere in a characteristic pattern. This pattern, termed the carbon cycle, can be represented in a redox fashion as shown in eq. 1 below:



(1)

This equation shows the circulation of carbon between high oxidation states (carbon dioxide in the atmosphere and carbonates in the hydro/geosphere) and low oxidation states (methane in the atmosphere and petrochemicals in the geosphere). The two most well known processes in the carbon cycle are labeled in eq. 1. Photosynthesis is the process of storing solar energy in carbon by reducing carbon dioxide to the formaldehyde oxidation state, generally in the form of saccharides. Respiration is the release of that energy through the oxidation of saccharides to carbon dioxide.

There are biological participants in the carbon cycle at all oxidation levels. The metabolic oxidation of methane to carbon dioxide occurs in methylotrophic bacteria (*1-5*). About half of the total organic matter degraded by anaerobic microbes is converted to methane, but only 0.5% of this methane reaches the

atmosphere (2). Most is consumed by methanotrophs. These organisms, some of which are capable of subsisting on methane as a sole source of carbon and energy, form a small but crucial ecological niche for recycling carbon from decaying biomass. Figure 2-1 portrays a greatly simplified schematic of the role methanotrophs play in the carbon cycle. Complex carbon molecules that are not easily broken down settle in soil layers that are very low in oxygen. Heterotrophs and methanogens anaerobically break down these complex molecules into smaller and more reduced pieces. Methane gas is the ultimate product of this process, for it can be neither disassociated into smaller pieces nor further reduced. It diffuses upward into soil containing relatively high oxygen levels where it is oxidized by methanotrophs to form either biomass or carbon dioxide. Since this methane would otherwise reach the atmosphere where methane can act as a greenhouse gas, methanotrophs occupy an important and interesting niche in the carbon cycle.

Methanotrophs are capable of subsisting on nothing but methane and oxygen for all of their carbon and energy. Methane metabolism in these organisms is summarized in Figure 2-2. The first two steps, catalyzed by methane monooxygenase (MMO) and methanol dehydrogenase, are energy neutral, since the MMO reaction consumes NADH and the second step produces a NADH equivalent. The organism then makes a choice between carbon fixation or energy production in the form of NADH. This metabolism mimics the larger scale carbon cycle, in that energy is stored as CH_2O units and released by burning those units to form CO_2.

The methanotroph metabolic reactions are interesting from both a chemical and economic standpoint. The first step in the process, the controlled oxidation of methane to methanol by dioxygen, is particularly interesting due to its kinetic difficulty. Current industrial processes to produce methanol from

methane require two steps (*6*). The first is the conversion of methane to "synthesis gas" (CO + H_2), which takes place over Ni, Al or Si catalysts, and the second is conversion of the "synthesis gas" to methanol, which is done by using Zn/Cu or Zn/Cr catalysts. Both reactions require pressures of 50-200 bar and reaction temperatures of 250-900 °C. The MMO system performs this reaction at 25 °C and 1 atmosphere utilizing water as a solvent. It would clearly be cheaper and safer to use catalysts that mimic the chemistry of MMO in industrial methanol production processes. To further that goal, examination of the reaction of methane and dioxygen in the MMO system is a first step in the design of new catalysts.

The enzymes that catalyze the MMO reaction come in two forms, a particulate MMO (pMMO) that contains copper and is membrane bound, and a soluble MMO (sMMO) that contains iron and is found in the cytoplasm (*7-10*). The sMMO enzymes are relatively simple to purify and are quite robust when compared to the pMMO, and they have therefore been studied in more detail. The stoichiometry of the MMO reaction is shown below in eq. 2.

$$CH_4 + NADH + H^+ + O_2 \xrightarrow{MMO} CH_3OH + NAD^+ + H_2O \qquad (2)$$

The sMMO hydroxylase protein (sMMOH) is the site of methane activation. It is a heterodimer consisting of two copies each of three subunits, MMOHα (60646 Dal), MMOHβ (45132 Dal) and MMOHγ (19847 Dal), resulting in a holoenzyme with a molecular weight of 251 kDal. It contains a non-heme, carboxylate-bridged diiron center that is the site of methane oxidation. This diiron unit is similar to centers found in a wide variety of enzymes, including ribonucleotide reductase (RR) (*11, 12*), stearoyl acyl carrier protein Δ-9 desaturase (*13*), phenol hydroxylase (*14*) and toluene-4-monooxygenase (*15*). The sMMO

system also contains a reductase and a coupling protein. The structure of the hydroxylase has recently been determined under a variety of conditions (*16*, *17*), and those data supplied important information about the mechanism of the enzyme.

Three proposals relate the structure of the hydroxylase to its function. Specifically, the MMOHα residues cysteine 151, threonine 213 and leucine 110 are hypothesized to have functions suggested by their positions in the crystal structure. It was first postulated by Dalton and Nordlund that C151 in the *M. capsulatus* (Bath) MMOH would occupy the same position in the active site as tyrosine 122 from ribonucleotide reductase subunit R2, based on sequence comparisons (*18*). The R2 iron center is responsible for converting Tyr122 to a radical. Subsequently, it was proposed that a thiyl radical at C151 might be a part of the mechanism of methane hydroxylation (*19*). The crystal structure confirmed that C151 occupies a position similar to the RR R2 tyrosine residue, although the sulfur atom of the cysteine is farther away from the iron center than the tyrosine oxygen atom due to the difference in size between tyrosine and cysteine side chains. A thiyl radical intermediate was recently observed by EPR spectroscopy in a coenzyme B12-dependent RR that is responsible for abstraction of a hydrogen atom from a ribose ring (*20*). A thiyl radical is possibly involved in the mechanism of hydrocarbon hydroxylation.

Threonine 213 also has a corresponding residue in RR R2 subunit. It occupies a position similar to that of serine 211 in the R2 protein, which is responsible for shuttling protons from solvent into the active site (*21*). In the structures of the reduced and oxidized frozen crystals, T213 is hydrogen bonded to a solvent molecule and an acetate ion, respectively (*17*, *22*). Both the water molecule and acetate are further linked by hydrogen bonds to the diiron center

via solvent ligands. The threonine residue is, therefore, part of an extensive hydrogen bonding network that may function to feed protons into the active site.

Another possible interpretation of the roles of the Cys151 and Thr213 residues is based on the observation that they are among the few protonated residues in the largely hydrophobic active site cavity. As such, they may serve to communicate structural changes via hydrogen bonding in the hydroxylase caused by binding of protein B or the reducatase to the diiron center, tuning its activity to reflect the state of the protein.

The third residue mentioned above is leucine 110, which crystallographic methods reveal to adopt two different orientations in different crystal forms (*23*). This residue may act as a gate between two hydrophobic pockets, termed cavities 1 and 2, which are proposed to be involved in substrate access to the active site iron center.

The ideal method to approach the testing of these proposals is site-directed mutagenesis, combined with analysis methods previously established in our laboratory, including X-ray crystallography and stopped-flow kinetics. Interesting mutants at position 151 would include C151A and C151S. MMOH C151S preserves the proton, but changes the electronic character of the residue. The net effect would be that the hydrogen bonding role of C151 will very likely be preserved, but that any radical that may be formed at that position in the native enzyme may not be formed in the mutant C151S. The C151A mutant will be useful as a control since neither hydrogen bonding nor radical formation is possible. The mutants at T213 that would be most informative include T213A, T213S, and T213C. Both T213S and T213C preserve the hydrogen bonding capability, but remove a methyl group. The hydrogen bonding strengths of serine and cysteine differ, and that difference may be observable kinetically if T213 were responsible for delivering protons to the active site during turnover.

The T213A mutant is, again, a control that removes the hydroxyl proton completely. L110 mutants would include a number of smaller residues to observe the effects of removing the proposed gate.

Generating these site-directed mutants depends on having the hydroxylase expressed in a system that allows for easy changes in the hydroxylase subunit gene sequence, and that allows the production of significant amounts of protein for study. This topic is the subject of the present thesis chapter.

A number of proteins homologous to the sMMOH have been expressed in *E. coli*, including ruberythrin (*24*). It was produced in an insoluble, iron-deficient form, which could be reconstituted by denaturing the protein with guanidinium chloride, followed by dilution in the presence of Fe(II). The Δ-9 desaturase has also been expressed in *E. coli* (*25*). High yields of soluble desaturase were prepared through the use of an optimized lactose fed-batch fermentation methodology, where lactose was used instead of IPTG as an inducer (*26*). The toluene-4-monooxygenase (T4M) system is quite similar to the sMMO system (*15, 27*). It has a hydroxylase with three subunits, a reductase and a protein B homologue. It has been expressed in *E. coli* by inducing the expression systems at very high cell densities, at low temperatures, at low IPTG concentrations and with Fe(II) media supplements. All three of these proteins were produced in an insoluble form under standard expression conditions, requiring specialized conditions for production of soluble material.

All of the proteins mentioned above, with the exception of T4M, are single polypeptides. Expression of proteins with more than one subunit is often a particular challenge, due to the need for subunit assembly in a foreign environment. Few multisubunit proteins have been expressed in *E. coli* (*28*). The most common strategy for expressing those kinds of systems is to prepare the

individual subunits separately, and recombine them at a later stage to form the holoprotein. Examples of this type of expression include recombinant actin and myosin subunits (*29, 30*), the two subunits of the human class I histocompatability antigen (HLA)-A2 (*31*), RNA polymerase I initiation factor SL1 (*32*), and two subunits from the corrinoid/iron-sulfur protein from *C. thermoaceticum* (*33*). A second strategy for expression is to produce an "operon" on a plasmid, either by replication of the wild-type operon or by construction of an artificial operon. The T4M system described above was expressed by using the latter methodology.

Several groups have reported progress in expressing the sMMO genes. The *M. capsulatus* (Bath) sMMOB and sMMOR proteins have been expressed in *E. coli* (*34-36*). sMMOH$^-$ mutants have also been generated in *M. trichosporium*, an important step for producing homologous mutants (*37*). Heterologous expression of the *M. trichosporium* sMMO system in *P. putida* F1 has been accomplished by using the pMMB broad host range (*38*), but no subsequent purification of the protein products was reported.

Expression of heterologous proteins in *E. coli* often results in the formation of protein aggregates, termed inclusion bodies (*28, 39-42*). These inclusion bodies contain the protein of interest in a non-native and non-active form.

Inclusion bodies are the result of the absence of some condition required for protein folding in a heterologous system. This problem has been most evident for eukaryotic proteins that have post-translational modification requirements, in proteins that have more than one subunit, in proteins that may require specific chaperonins for correct folding, and in proteins requiring the assistance of other proteins for insertion of a cofactor (*43*).

There are two broad methods for solving the problem of inclusion bodies, *in vivo* and *in vitro*. *In vivo* methods prevent the formation of inclusions by

changing the conditions inside the cell during protein induction (*28*). Examples include growing cultures at a lower temperature, using lower IPTG concentrations, using higher media salt concentrations (to increase the hydrostatic pressure inside the cell), and changing expression systems or organisms. *In vivo* methods often fail, since there is very little that one can do to change the physical environment inside the cell. *In vitro* solutions involve collecting the inclusions and subjecting them to a procedure to refold the protein (*42, 44*). There are two common examples of such procedures. One can solubilize the inclusion bodies by using a denaturant such as urea or guanidinium hydrochloride and then refold them by dialysis to remove the denaturant. Alternatively one can solubilize the inclusions by using a detergent such as sarcosyl, and then renature the protein by detergent removal. *In vitro* methods are the most common solutions reported in the literature, and if they work well, the formation of inclusion bodies can be an advantage since they are easily isolated in high purity from soluble protein impurities.

This chapter reports on progress toward the expression of the soluble MMOH in *E. coli.* Construction of a series of vectors for both "operon" and individual subunit expression is described, together with the use of those systems to produce substantial amounts of the three sMMOH subunits as inclusion bodies. Both refolding and expression conditions were systematically changed in attempts to produce active sMMOH. Finally, based on observations and deductions from on the data reported, a future course of action to express the hydroxylase is outlined.

Materials and Methods

All chemicals were from Sigma-Aldrich, except IPTG (Boeringer-Mannheim) and when noted otherwise. All enzymes were from New England Biolabs unless stated explicitly. Manufacturer's instructions were followed exactly in their use unless stated. All restriction digests were carried out at 37 °C over 6 h. Double digests were performed first with low-salt optimal enzyme (3 h) followed by addition of high-salt buffer and the second enzyme (3 h). Ligations were carried out overnight at 14 °C. *E. coli* strains INVαF' or XL1-Blue were used for all DNA manipulations. *E. coli* strain JM105 was used as expression host at 37 °C unless mentioned otherwise.

Low Pressure Chromatography sMMOH Purification Procedure. The sMMOH and a crude mixture of sMMOB and sMMOR (termed B/C mix) were prepared as described previously (*45, 46*). Average iron content of the hydroxylase was 3.4-3.6 Fe/protein and the average activity (in units as described in the literature) was 220-250 mU/mg. An SDS-PAGE gel showing the quality of the protein obtained in this manner is presented in Figure 2-3.

Production of MMOH Polyclonal Antibodies. Two female New Zealand white rabbits were obtained. Blood samples (15 mL) were taken from each animal before any injections were made. The blood was incubated at room temperature for 4 h to clot, then left at 4 °C for 12 h to retract the clot. The serum was decanted and spun in a table-top centrifuge for 10 min at 200 rpm to remove additional blood particulates. The serum was frozen in liquid nitrogen and stored at –20 °C.

Protein antigen injections were prepared by first exchanging and concentrating a native sMMOH sample into PBS at a concentration of 3 mg/mL. A 2 mL sample of the exchanged protein solution was mixed with 2 mL of

Freund's Complete Adjuvant for the first injection or 2 mL of Freund's Incomplete Adjuvant for all subsequent injections. This solution was either vortexed vigorously at 4 °C for 30-60 min or exchanged rapidly between two syringes connected by a narrow stopcock until a thick emulsion formed. This emulsion was divided into two 2 mL syringes and given to MIT Animal Care personnel to inject each rabbit at multiple subcutaneous sites. Two to three weeks after each injection, the animals were bled (15 mL), and serum containing antibodies to the hydroxylase was isolated as described above.

This cycle was repeated 6 times. The final bleed was done by fatal cardiac puncture in which most of the blood from the rabbit was collected, resulting in approximately 50 mL of material. Total serum collected was approximately 120 mL. The serum was used in subsequent experiments without further purification. Protein concentration was approximately 25 mg/mL in the sera samples.

ELISA Assays. A standard sMMOH solution of known purity and quality was used as the standard for all ELISA assays. ELISA 96-well polystyrene plates were purchased from Corning.

In each ELISA assay, one column was reserved as a blank, no-protein control, and another was reserved for a sMMOH standard. Each row was a 1:3 dilution of the row above it, resulting in a typical range of concentrations in the sMMOH column from 160 μg/mL to 73 ng/mL. All dilutions were done in PBSN buffer (140 mM NaCl, 10 mM KCl, 25 mM $NaPO_4$ at pH 7.4). The protein solution to be tested was diluted and treated in an identical manner to the sMMOH standard.

Protein aliquots of 50 μL were placed in each well. The plate was wrapped in plastic and either baked at 37 °C for 2 h or left to incubate at 25 °C overnight. The protein solution was washed out of the wells with 3 rinses of deionized

water. The remaining sites for protein binding in the wells were blocked by filling each well to the brim with approximately 500 μL ELISA blocking buffer (170 mM H_3BO_4 at pH 8.5, 120 mM NaCl, 0.5% Tween 20, 1 mM EDTA, 0.25% BSA, 0.05% NaN_3) and soaking for 30 min. The blocking buffer was removed by using 3 rinses of deionized water. A solution of 625 ng/mL antiserum was made by dilution of the anti-sMMOH stocks described above in blocking buffer. Aliquots of 50 μL anti-sMMOH antibody were added to each well and allowed to soak at room temperature for 2 h. The antibody was washed out by rinsing 3 times with deionized water, soaking in blocking buffer for 10 min, and rinsing again 3 times in deionized water. A solution of 500 ng/mL Protein A-alkaline phosphatase conjugate was made by dilution of a stock solution in blocking buffer. Aliquots of 50 μL volume of the conjugate were added to each well and allowed to soak in for 2 h. The conjugate was washed out by rinsing 3 times with deionized water, soaking in blocking buffer for 10 min, and rinsing again 3 times in deionized water. A solution of 1.0 mg/mL *p*-nitrophenol phosphate (pNPP) was prepared by dissolving a pre-weighed tablet in an appropriate volume of 200 mM Tris buffer at pH 8.0. Aliquots of 200 μL volume of the pNPP solution were added to each well. The color of the pNP product of the alkaline phosphatase reaction was allowed to develop over 2 to 16 h, depending on the strength of the color.

The plates were read on an EFLab Titertek Multiskan model 110V multiwell plate reader utilizing a lamp filter set to 405 nm. The absorbance values for the sMMOH standard were used to construct a standard curve, and the protein concentration of sMMOH in the unknown samples was determined from that curve.

Western Blots. An SDS-PAGE gel containing one lane of sMMOH pure standard and other lanes with unknown samples were prepared and blotted onto

nitrocellulose. After blotting, the nitrocellulose membrane was sealed in a chamber containing 100 mL TTBS (0.1% Tween 20, 100 mM Tris pH 7.5, 0.9% NaCl) and allowed to soak for 1 h. The TTBS was removed and replaced with 40 mL of a 12.5 μg/mL solution of anti-sMMOH serum, which was prepared by dilution of the stock antisera described above in TTBS. After a 1 h soak, the membrane was removed and washed four times with 50 mL TTBS. A stock solution of goat anti-rabbit IgG(H+L)-horseradish peroxidase (HRPO) conjugate (obtained from Biorad) was diluted 1:3000 in 20 mL of TTBS. The membrane was soaked in that solution for 1 h, and washed four times with 50 mL TTBS. A 20 mL solution of 0.015% H_2O_2, 3 mM 4-chloro-1-naphthol, 0.9% NaCl and 100 mM Tris pH 7.5 was prepared and the membrane was soaked in that solution until adequate color development was observed. The HRPO reaction was stopped by a water wash, and the membrane was allowed to dry in air. A sample western blot prepared by using that protocol is shown in Figure 2-4.

Genomic DNA Sample Preparation. All procedures were carried out at 25 °C unless noted. *M. capsulatus* (Bath) cell paste (1 gram) was thawed and resuspended in 9.8 mL of TE (pH 8). SDS (200 μL of a 25% solution) and 1.5 mg proteinase K were added and mixed gently, followed by incubation at 37 °C for 2 h and 60 °C for 20 min. SDS (600 μL of a 25% solution) was added and incubated at 60 °C for 10 min. Sodium perchlorate (2.5 mL of a 5 M solution) was added, mixed gently, then the solution was again incubated at 60 °C for 15 min. The resulting solution was extracted with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol by shaking gently for 30 min and centrifuged at 20,000 g for 20 min. The viscous aqueous layer containing the genomic DNA was removed by using a wide end pipet tip. This extraction procedure was repeated until no white precipitate formed at the phase boundary (typically 2-3 times). The DNA/RNA solution was precipitated by slow addition of 95%

ethanol at –80 °C to form two layers. The layers were mixed slowly by gentle agitation with a thin glass rod, curved at the tip. The DNA/RNA pellet accumulated on the glass, and was subsequently dried and resuspended in 7 mL of 150 mM NaCl/150 mM sodium citrate (pH 7). DNase-free RNase was added to the solution to 50 mg/mL, and the mixture was incubated at 37 °C for 30 min. The solution was phenol extracted and precipitated as above, except that the DNA pellet was washed in 70% ethanol before speed-vac drying. The pellet was resuspended in 5 mL TE (pH 7.4).

Purification of PCR Primers for pDEC002. The following primers (5' to 3') were ordered from MIT Biopolymers and stored as deprotected, lyophilized powders at –20 °C: XN1.1: GGCCATGGAGGAGGTAAGTAATGGCACTTAG-CACCGCA, XC1.1: CTCGAATTCTCAATTGAATGCCTTCACCG-GGTT, YN1.1: GGGAATTCAGGAGGTTCGATATGAGCATGTTAGGACAA, YC1.1: CTCAC-GCGTTCAATCCTGCCAGAACCGCTTTAA, ZN1.1: GGACGCGTAGGAGGTA-TGACATGGCGAAACTGGGTATA, ZC1.1: CTCGAGCTCTCAGTGCGGGCA-CTGCAGATGCAC. Each primer sample was resuspended in 100 μL of water, a 40 μl aliquot of which was mixed with 10 μL of USB Sequenase stop solution. These six samples were loaded onto 7 M urea 12% polyacrylamide gels, which were run for 1.75 h at 250 mV at 50 °C. The bands corresponding to the desired primers were visualized by using UV shadowing and were excised from the gel by using fresh, clean razor blades. The gel slices were crushed and the DNA was eluted by soaking overnight at 37 °C in 1 mL of 3 M sodium acetate. The samples were extracted by addition of an equal amount of 25:24:1 phenol:chloroform: isoamyl alcohol, vortexing for 15 s, spinning down in a microcentrifuge for 30 s, and pipetting off the top aqueous layer. The water layer was added to 2.5X volumes of 95% ethanol, vortexed and kept at –80 °C for 30 min. The tubes were spun at 4,000 rpm for 30 min to pellet the primers. The supernatant was decanted

and the pellets were washed in 1 mL of 70% ethanol. The samples were spun down at 4,000 K for 30 min and the supernatant was decanted. The pellet was dried under vacuum and resuspended in 40 μL TE (pH 7.4).

PCR Conditions for pDEC002. The following were added to a total volume of 100 μL for each of the three subunits: 1000 ng of genomic DNA, 100 pmol of the N-terminal primer (_N1.1), 100 pmol of the C-terminal primer (_C1.1), 1 μL of 20 mM dNTPs, 10 μL of 10X *Taq* polymerase reaction buffer and 1 μL *Taq* polymerase (5 units). These samples were placed in a thermocycler that was programmed to with the following cycle: 1 x (7 min at 94 °C, 2 min at 60 °C, 2 min at 72 °C), 38 x (2 min at 94 °C, 2 min at 60 °C, 2 min at 72 °C), 1 x (2 min at 94 °C, 2 min at 60 °C, 10 min at 72 °C). Each of the three samples was purified from reaction buffer salts, primers and dNTPs by using the Qiagen PCR spin quantitation kit. An agarose gel showing the results of the PCR reactions is given in Figure 2-5.

Construction of pDEC002. Direct 4-piece ligations of *mmoX*, *mmoY*, *mmoZ* and the chosen expression vector pTrc99A did not perform adequately, so a more complex serial cloning procedure was adopted, as diagrammed in Figure 2-6. A double digest of 1 μg of pSL301 (Novagen) and 1 μg of X1.1 PCR product was performed by using EcoR I and Nco I. The cut pSL301 was treated with CIP (Boeringer-Mannheim) and both restriction digests were purified by agarose gel electrophoresis followed by treatment with the USBioClean kit. A ligation was performed with 2 μL Boeringer-Mannheim T4 ligase, 2 μL 10X T4 ligation buffer, 2 μL 10 mM ATP, 0.1 μg digested, phosphatased pSL301, and 0.1 μg digested X1.1 in a total volume of 20 μL. The reaction was carried out at 4 °C overnight. The entire ligation reaction was transformed into *E. coli* strain JM105 made competent by using the RbCl Hanahan protocol (*47*). The transformed cells were plated onto LB plates (75 μg/mL ampicillin). Colonies were screened by

restriction mapping with EcoR I and Nco I. The positives were labeled pSL301-*mmoX*. Approximately 1 μg of pSL301-*mmoX* and 1 μg of Y1.1 PCR product were double digested with EcoR I and Mlu I. The cut pSL301-*mmoX* was treated with CIP and both digests were purified by agarose gel electrophoresis followed by treatment with the USBioClean kit. A ligation was done with 2 μL Boeringer-Mannheim ligase, 2 μL 10X ligation buffer, 2 μL 10 mM ATP, 0.1 μg digested, phosphatased pSL301-*mmoX*, and 0.1 μg digested Y1.1 in a total volume of 20 μL. The reaction was carried out at 4 °C overnight. The entire ligation reaction was transformed into competent *E. coli* strain JM105. The transformed cells were plated onto LB plates (75 μg/mL ampicillin). Colonies were screened by restriction mapping with EcoR I and Mlu I. The positives were labeled pSL301-*mmoXY**. Approximately 1 μg of pSL301-*mmoXY** and 1 μg of Z1.1 PCR product were double digested with Mlu I and Sac I. The cut pSL301-*mmoXY** was treated with CIP, and both restriction digests were purified by agarose gel electrophoresis followed by treatment with the USBioClean kit. A ligation was carried out with 2 μL ligase, 2 μL 10X ligation buffer, 2 μL 10 mM ATP, 0.1 μg digested, phosphatased pSL301-*mmoXY**, 0.1 μg digested Z1.1 in a total volume of 20 μL. The reaction was performed at 4 °C overnight. The entire ligation reaction was transformed into competent *E. coli* strain JM105. The transformed cells were plated onto LB plates (75 μg/mL ampicillin). Colonies were screened by restriction mapping with Mlu I and Sac I. The positives were labeled pSL301-*mmoXY*Z*.

A mistake in the published sequence of the *mmoY* gene was discovered midway through the cloning procedure described above (see chapter one for a description of the error). The C-terminal cloning primer for the *mmoY* gene was, therefore, designed erroneously, and the sequence of the pDEC002 gene was not as expected. A new PCR primer, YC20 (5'-TAGTTACGCGTTTATTTCAGTC-

CTGCCAGAAC-3') was designed to correct the error. YC20 was synthesized on a Cruachem DNA synthesizer and PAGE-purified according to manufacturer's instructions. A PCR reaction was carried out by adding 20 pmol of primer YC20, 20 pmol primer YC11, 0.1 μg of pCH4, 2 μL of 10 μM dNTPs, 10 μL 10X *Taq* polymerase reaction buffer and 1 μL of Gibco BRL *Taq* polymerase (licensed for PCR by Hoffman-La Roche) in a total volume of 100 μL. The samples were overlaid with 75 μL mineral oil and placed in a thermocycler set to run the following cycle: 1X (5 min at 94 °C), 30 X (1 min at 94 °C, 1 min at 60 °C, 1 min at 72 °C), 1 X (8 min at 72 °C). The PCR reactions were cloned into the pCRII vector (purchased from Invitrogen) directly by adding 1 μL of the PCR reaction without purification to a 10 μL ligation reaction containing 1 μL pCRII vector, 1 μL ligase (from Boeringer-Mannheim) and 1μL ligase buffer. The ligation reactions were transformed into INVαF' cells and selection of positive *mmoY* insert clones was done by using a combination of α-complementation and restriction mapping. The *mmoY* insert of interest was isolated by cutting selected pCRII-*mmoY* clones with Mlu I and EcoR I. A sample of plasmid pSL301-*mmoXY*Z* was also cut by Mlu I and EcoR I to excise the incorrect *mmoY* insert. The appropriate DNA fragment from each restriction digest was purified by use of Millipore MP filter units for purification of DNA from agarose gel slices. The *mmoY* correction insert and the pSL301-*mmoXY*Z* plasmid shell were ligated together to form pSL301-*mmoXYZ*. The corrected sMMOH artificial operon was cut with Nco I and Hind III to excise the operon fragment, which was subsequently ligated into the multiple cloning site of pTrc99A to form pDEC002. The structure of this plasmid was verified by restriction mapping and is shown in Figure 2-7.

Sequencing of pDEC002. The entirety of the pDEC002 plasmid was sequenced by using the protocols and primers described in chapter one. The *mmoX*, *mmoY* and *mmoZ* genes were all sequenced completely on both sense and

antisense strands, along with approximately 100 bp on either side of the cloned genes.

Expression of MMOH by using the pDEC002 system. The pDEC002 plasmid was transfected into JM105. The standard MMOH expression experiment was as follows. A 5 mL culture grown in LB media supplemented with 100 μg/mL of ampicillin was grown to saturation overnight. A 100 μL sample of this culture was diluted into 100 mL of LB media containing 100 μg/mL ampicillin. The culture was allowed to grow until the cells reached on O.D. of 0.6 to 0.8. IPTG was added to a final concentration of 0.4 mM and the cells were allowed to grow for another 2 h. The culture was centrifuged at 4,000 rpm in a Sorvall centrifuge for 10 min, and the media was discarded. The cells were resuspended in 20 mL of cracking buffer (25 mM MOPS at pH 7, 5 mM $MgCl_2$, 5 mM sodium thioglycolate, 1 mM 6-aminocaproic acid, 1 mM 3,5-diaminobenzoic acid, and approximately 1 mg/mL DNase), and passed through a French Pressure Cell (SLM-Aminco) at 15,000 psi twice. The crude cell lysate was ultracentrifuged at 40,000 rpm in a Beckman ultracentrifuge for 90 min. The supernatant was saved as the soluble fraction, and the pellet was solublized in cracking buffer with the addition of 6 M urea. An SDS-PAGE gel showing the results of a sample pDEC002 expression experiment is shown in Figure 2-8.

MMOH Inclusion Purification. A 10 mL sample of pDEC002 system inclusions (approximately 1 mg/ml) solubilized in standard cracking buffer plus 6 M urea was loaded on a 30 mL DEAE Biogel (Pharmacia) column equilibrated in 25 mM MOPS (pH 7)/6 M urea. The column was washed with 32 mL of buffer at 1 mL/min, and was washed in a NaCl gradient from 0 to 300 mM over 192 min at 1 mL/min. The column was washed with 300 mM NaCl buffer for 76 min (at 1 mL/min).

Post-Expression Refolding Experiments

Dialysis Reconstitution of Native MMO Hydroxylase Isolated from M. capsulatus *(Bath).* A 10 mL sample of active hydroxylase (1 mg/mL) was prepared. Powdered urea was added slowly with slow stirring to bring the sample to 6M. Stirring continued for 1 h to insure complete denaturation of protein. This sample was placed in a dialysis bag and dialyzed against 1 L of cracking buffer plus 0.5 mM ferrous ammonium sulfate for 8 h. The buffer was exchanged and dialysis was continued for another 8 h. During this time, a flocculent precipitate formed in the bag. After dialysis was complete, the contents of the bag were ultracentrifuged at 40,000 rpm for 90 min. The supernatant was drop frozen at -80 °C and the pellet was solublized in 5 mL of cracking buffer plus 6M urea. Similar experiments were performed by using 4 M guanidinium hydrochloride in place of the 6 M urea, with identical results.

Dialysis Reconstitution of Insoluble Expressed Hydroxylase Isolated from E. coli. A 3 mL sample of the insoluble sMMOH pellet described above was diluted in 27 mL of cracking buffer plus 4 M guanidine hydrochloride (approximately 1 mg/mL). A 10 mL sample of the diluted protein was placed in each of three dialysis bags (10 kDa MW cutoff). These three bags were each placed in 1 liter of cracking buffer, 1 liter of cracking buffer plus 0.5 mM ferric EDTA, and 1 liter of cracking buffer plus 0.5 mM ferrous ammonium sulfate respectively. Stirring continued for 8 h, after which time the outer buffer was exchanged, and stirring was continued for another 8 h. During this dialysis, a white flocculent precipitate formed in the bags, usually after 1 or 2 h. After dialysis was complete, the contents of the bags were ultracentrifuged at 40,000 rpm for 90 min. The supernatant was drop frozen at –80 °C and the pellet was solublized in cracking

buffer plus 4M guanidine hydrochloride. The supernatant from each of the three dialyses was analyzed on SDS-PAGE gels, shown in Figure 2-9.

Anaerobic Fe(II) Dialysis Reconstitution of Insoluble Expressed Hydroxylase Isolated from E. coli. In analogy with the *Desulfovibrio vulgaris* rubrerythrin refolding procedure (*24*), an anaerobic denaturant dialysis was attempted to refold the MMOH inclusions. A 5 mL volume of a 50 μM MMOH solution in 6 M urea/cracking buffer solution was thawed in a 500 mL RB flask with a 70 μL aliquot of β–mercaptoethanol. The flask was sealed with a septum and degassed by 20 cycles of evacuation followed by back filling with argon gas. This protein sample along with 250 mL of degassed 25 mM MOPS (pH 7) buffer and a vial of 0.1 g Fe(II)$(NH_4)_2(SO_4)_2$ and 0.01 g ascorbic acid were cycled into an anaerobic box, where the iron/ascorbate powders were added to the 250 mL of buffer. The buffer was added slowly to the protein solution over two h, diluting the urea to 120 mM. The protein solution was ultracentrifuged to remove the precipitated protein and concentrated to a volume of 10 mL in an Amicon concentrator.

This anaerobic Fe(II) refolding experiment was also carried out by using denatured MMO hydroxylase as isolated from *M. capsulatus* (Bath). All conditions were identical, except that the starting concentration of MMOH in the 5 mL sample was 5 μM.

Sarcosyl Lysis Growth. The standard growth protocol was followed up to the resuspension step. The cells were resuspended in 10 mL of cracking buffer. A 200 μL aliquot of 10% w/v N-lauroyl sarcosine (sarcosyl detergent) was added, and the sample was passed through the French press (2 passes, 1000-1200 psi). A 1 mL sample of 20% 1-O-octyl-β-D-glucopyranoside (octylglucoside detergent) was added and the mixture was stirred for 20 min to allow the sarcosyl to diffuse into the octylglucoside micelles. The sample was ultracentrifuged at 40,000 rpm for 90 min. The supernatant was drop frozen at –80 °C and the pellet was

solublized in cracking buffer plus 6 M urea. Both the supernatant and pellet were examined by SDS-PAGE gels, shown in Figure 2-10. All of the hydroxylase remained in the insoluble form.

Detergent Lysis Screening. Four buffer solutions containing various detergents were made according to the following procedure. All buffers contained 25 mM MOPS (pH 7), 1 mM aprotinin, 1 mM pepstatin, 1 mM PMSF and 5 mM sodium thioglycolate. Buffer NI contained 0.5 mM Nonidet P40. Buffer TW contained 0.1 mM Tween 20. Buffer DM contained 1.6 mM dodecylmaltoside. Buffer OG contained 25 mM octylglucoside.

For each of these buffers, the standard MMOH inclusion preparation was followed, excepting that the cracking buffer was replaced by detergent buffer. An SDS-PAGE gel of those samples is shown in Figure 2-11.

Sarcosyl Pellet Solubilization. The standard growth protocol for the pDEC002/JM195 system was followed to the ultracentrifugation step. After ultracentrifugation, the supernatant was drop frozen, and the pellet was resuspended in 3 mL of SPS buffer (25 mM MOPS at pH 6.95, 5 mM sodium thioglycolate {added immediately before use}, 1 mM PMSF, 20 μg/mL aprotinin, 5 μg/mL leupeptin, 2.5 μg/mL pepstatin, 1.5% sarcosyl detergent). This suspension was homogenized in a glass homogenizer. The sample was ultracentrifuged at 40,000 rpm for 20 min and the supernatant was mixed gently into 24 mL OG buffer (25 mM MOPS at pH 6.95, 5 mM sodium thioglycolate {added immediately before use}, 1 mM PMSF, 20 μg/mL aprotinin, 5 μg/mL leupeptin, 2.5 μg/mL pepstatin, 2.5% octylglucoside detergent, 1 M NaCl). This sample was ultracentrifuged for 30 min at 40,000 rpm. The supernatant was drop frozen at -80 °C and the pellet was solublized in 5 mL cracking buffer plus 6 M urea. The first and second supernatants and the pellet were analyzed by SDS-PAGE as shown in Figure 2-12. A 10 mL sample of the second supernatant was

placed in a 10 mL dialysis bag and dialyzed against 1 liter cracking buffer for 8 h to remove the octylglucoside/sarcosyl micelles. During this initial period, no protein precipitated. The outer buffer was exchanged and dialysis continued for 8 h. During this period, a white precipitate formed in the bag. After dialysis, the sample was ultracentrifuged for 30 min at 40,000 rpm. The supernatant was concentrated to 2 mL by using Centricon-10 concentrators, and the pellet was solubilized in 2 mL cracking buffer plus 6 M urea.

Alteration of Expression Conditions

Low Temperature Induction Growths. The standard growth protocol was followed, except that the 2 h induction period after addition of IPTG was carried out at 30 °C and 25 °C for two separate 500 mL growths. An additional trial was done under standard conditions, excepting that the IPTG was added at O.D. 0.2. Results as monitored by SDS-PAGE gels were identical to the 37 °C induction growths.

E. coli *Strain Screening.* The pDEC002 vector was transfected into the following *E. coli* strains: BL21(DE3), HMS174, HMS174(DE3), C600, HB101, K12, XL1-Blue, and JM105 (as a control). These strains were grown by following the pDEC002 protocol. All strains performed identically to JM105.

Variation of Media Content. Variants on the protein expression procedure were attempted to induce production of soluble protein. The medium was changed to M-9 minimal media, and to TYPGN (super-rich). Results were identical to those obtained under standard conditions.

Expression System Screening (pDEC010). The fragment of pDEC002 containing the artificial operon (*mmoXYZ*) was excised from the plasmid by using Nco I and HinD III and ligated into the pET11d vector to form pDEC010.

This plasmid was transfected into the BL21(DE3) strain and grown by using the pDEC002 protocol. An SDS-PAGE gel showing the results of that system is shown in Figure 2-13.

Individual Subunit Expression

Construction of pDEC05x. Linker oligonucleotides linkX1 (5'-AATTTGC-GGATCCGCTCGA-3'), linkX2 (5'-AGCTTCGAGCGGATCCGCA-3'), linkY1 (5'-CATGGAATGAGGATCCGCGT-3') and linkY2 (5'-AATTACGCGGGATCC-TCATTC-3') were synthesized on a Cruachem DNA synthesizer and were purified by PAGE according to Cruachem directions.

Plasmid pDEC002 was cut with BamH I and EcoR I and was purified from an agarose gel slice by using a Millipore MP filter unit. Approximately 0.1 μg of the pDEC002 fragment was ligated with 100 pmol of each linker oligonucleotide linkX1 and linkX2. The resulting plasmid, with the EcoR I site removed and a BamH I site added via linker design, was restriction mapped and labeled pDEC050.

Plasmid pDEC002 was cut with Nco I and EcoR I and was purified from an agarose gel slice by using a Millipore MP filter unit. Approximately 0.1 μg of the pDEC002 fragment was ligated with 100 pmol of each linker oligonucleotide linkY1 and linkY2. The resulting plasmid, with the EcoR I site removed and a BamH I site added via linker design, was restriction mapped and labeled pDEC055. The construction and structure of plasmids pDEC050 and pDEC055 are shown in Figures 2-14 and 2-15, respectively.

Construction of pDEC060. This construction is diagrammed in Figure 2-16. Plasmids pDEC050 and pET20b were digested with Nco I and HinD III. The agarose gel bands corresponding to the *mmoX* cassette and the pET20b fragment

were purified by using Millipore MP filter units. Approximately 0.1 μg of each DNA fragment were ligated together. The resulting pDEC060 plasmid was restriction mapped to confirm its structure as shown in Figure 2-17.

Construction of pDEC070. Primer oligonucleotides XN30 (5'-GCGCCATGGCACTTAGCACCGCAACC-3') and XC30 (5'-CTCAAGCTTTCA-ATTGAATGCCTTCACCGG-3') were synthesized on a Cruachem DNA synthesizer and purified by PAGE according to manufacturer's specifications. The PCR reactions were carried out by adding 2.60 μg pCH4 template DNA, 20 pmol of each primer, 2 μL of 10 μM dNTPs, 10 μL 10X *Taq* polymerase reaction buffer and 1 μL of Gibco BRL *Taq* polymerase in a total volume of 100 μL. They were placed in a thermocycler set to cycle for 1X (2 min at 94 °C), 10 X (2 min at 94 °C, 2 min at 60 °C, 6 min at 72 °C), 1 X (10 min at 72 °C). The reaction products were cloned into the pCRII vector (purchased from Invitrogen) directly by adding 1 μL of the PCR reaction directly to a 20 μL ligation reaction containing 1 μL pCRII vector, 1 μL ligase (from Boeringer-Mannheim) and 2 μL ligase buffer. The ligation reactions were transformed into INVαF' cells. Selection of insert clones was done by using α-complementation followed by EcoR I digests to identify inserts of the correct size.

The *mmoX* insert was excised by using Nco I and HinD III and ligated into the multiple cloning site of pET32a. The plasmid was labeled pDEC070 and restriction mapped. Its construction is shown in Figure 2-16, and its structure is shown in Figure 2-17.

Simple MMOH subunit expression - pDEC05x. Both pDEC050 and pDEC055 were transformed into JM105. A 5 mL culture of each strain was grown in LB amp media (100 μg/mL ampicillin) to saturation overnight. A 1 mL sample of this culture was diluted into 1000 mL of LB amp media. The culture was allowed to grow until the cells reached on O.D. of 0.6. IPTG was added to a final

concentration of 0.4 mM and the cells were allowed to grow for another 2 h. The culture was centrifuged at 4,000 rpm in a Sorvall centrifuge for 10 min, and the LB supernatant was discarded. The cells were resuspended in 40 mL of cracking buffer (25 mM MOPS at pH 7, 5 mM $MgCl_2$, 5 mM sodium thioglycolate, 1 mM 6-aminocaproic acid, 1 mM 3,5-diaminobenzoic acid, and approximately 1 mg/mL DNase), and passed through a French Pressure Cell (SLM-Aminco) at 15,000 psi twice. The crude cell lysate was ultracentrifuged at 40,000 rpm in a Beckman ultracentrifuge for 90 min. The supernatant was saved as the soluble fraction, and the pellet was solublized in cracking buffer with the addition of 6 M urea. The results of this expression are shown in Figure 2-18.

Periplasmic Expression of sMMOHα - pDEC060. The pDEC060 plasmid was transformed into BL21(DE3) cells. A cell culture procedure identical to the one described above was carried out by using that strain. Results are shown in Figure 2-19.

Thioredoxin-sMMOHα subunit expression. A 10 mL culture of an AD494(DE3)/pDEC070 strain containing the sMMOHα-thioredoxin fusion gene was grown in LB amp media (100 μg/mL ampicillin) to saturation overnight. A 6 mL culture sample was diluted into 6 L of LB amp media. The culture was cooled to 20 °C after the cells reached on O.D. of 0.6. IPTG (Boeringer-Mannheim) was added to a final concentration of 0.4 mM and the cells were allowed to grow for another 2 h at 20 °C. The culture was centrifuged at 4,000 rpm in a Sorvall centrifuge for 10 min, and the LB supernatant was discarded. The cells were resuspended in 40 mL of cracking buffer (25 mM MOPS at pH 7, 5 mM $MgCl_2$, 5 mM sodium thioglycolate, 1 mM Pefabloc SC, and approximately 1 mg/mL DNase), and passed through a French Pressure Cell (SLM-Aminco) at 15,000 psi twice. The crude cell lysate was ultracentrifuged at 40,000 rpm in a Beckman

Ultracentrifuge for 90 min. The supernatant contained the soluble portion of the sMMOH alpha-thioredoxin fusion, and was saved for purification.

Purification of sMMOHα. The soluble cell lysate was applied to a 25 mL Ni^{2+}-NTA resin column at 1 mL/min. Following loading, the column was washed with buffer containing 20 mM Tris at pH 7.9, 500 mM NaCl and 5 mM imidazole at 1 mL/min. The bound proteins were eluted by using a gradient of 5 mM imidazole to 150 mM imidazole over 1 h at 1 mL/min, followed by washing of the column with 150 mM imidazole over 1 h at 1 mL/min. The alpha-thioredoxin fusion elutes at approximately 100 mM in a very broad peak.

The crude alpha-thioredoxin fraction was exchanged into buffer containing 20 mM Tris pH 7.9, 150 mM NaCl and 2.5 mM $CaCl_2$ and concentrated to a volume of approximately 10 mL. 3 units of thrombin (Novagen) were added to the crude mixture, and the cleavage reaction was allowed to proceed for 16 h at 4 °C.

The crude protein solution was applied to a 25 mL Ni^{2+}-NTA resin column at 1 mL/min. The flow-through fractions containing pure alpha subunit were collected.

Protein samples were analyzed by using SDS-PAGE gels and Western blotting. Protein concentrations were determined by using ELISAs. Iron concentration was determined by using the ferrozine assay. Results are shown in Figure 2-21.

EPR Mixed-Valent Sample Preparation. Protein samples, either native hydroxylase or expressed sMMOHα, of approximate concentration 50 μM and approximate volume of 250 μL were prepared by concentration in Centriprep 10 concentrator units. Stoichiometric amounts of the mediators potassium indigo tetrasulfonate, methylene blue and phenazine methosulfate were added to the sample. The sample was degassed by 15 cycles of vacuum evacuation followed

by argon backfilling. A degassed solution of 5 mM sodium dithionite in 50 mM MOPS (pH 8.6) buffer was made and used to titrate the protein solution until the methylene blue changed from blue to yellow. The sample was anaerobically transferred to a quartz EPR tube and frozen in liquid nitrogen. Results are shown in Figure 2-22.

EPR Reduced Sample Preparation. Expressed sMMOHα samples of approximate concentration 100 μM and approximate volume of 250 μL were prepared by concentration in Centriprep 10 concentrator units. A stoichiometric amount of the mediator methyl viologen was added to the sample. The sample was degassed by 15 cycles of vacuum evacuation followed by argon backfilling. A degassed solution of 5 mM sodium dithionite in 50 mM MOPS (pH 8.6) buffer was made and used to titrate the protein solution until the methyl viologen changed from yellow to blue. The sample was anaerobically transferred to a quartz EPR tube and frozen in liquid nitrogen.

Results

The template for most of the PCR cloning discussed above was the pCH4 plasmid, a gift from Prof. J. Colin Murrell. This plasmid is a cDNA copy of the entire sMMO operon cloned into pBR325 (*48*). The sequence and structure of this operon can be found in chapter one. The initial pDEC002 plasmid cassettes (*mmoX*, *mmoY** and *mmoZ*) were amplified from a genomic DNA sample directly prepared from *M. capsulatus* (Bath) cell paste.

The plasmids in this chapter were designed by using the commonly used expression-cassette strategy (*49*). This methodology allows for very precise control over the design of the expression plasmid. By designing PCR primers with mismatches at the 5′ end, one can insert any desired sequence at the ends of

the PCR cassettes, including spacers, restriction sites and ribosome binding sites. The disadvantage is that PCR based cloning often results in sequences with mistakes since *Taq* polymerase does not have the full array of error checking machinery that many other DNA polymerases do.

The design for the first useful plasmid constructed, pDEC002, shown in Figure 2-7, was intended to create a simplified sMMO operon that would only express the three subunits of the hydroxylase. Upon induction of expression, the *E. coli* cell culture color changed from an ordinary yellowish color to white. Observation of the cells under the microscope revealed that the cells had large white organelles taking up approximately 10-25% of the cell volume. These white 'bodies' were collected by cell lysis followed by centrifugation. The peptide component of the aggregate was solublized with denaturants such as urea and guanidinium hydrochloride, which showed the white material to be entirely protein. SDS-PAGE analysis revealed that the protein was approximately 50% hydroxylase peptides that had been misfolded and packaged as inclusion bodies.

Several lines of research were followed in an attempt to obtain active hydroxylase from the pDEC002 plasmid. A system that produces inclusion bodies can actually be an advantage if one is able to induce the collected protein to refold into an active form. The packaging of the misfolded, expressed protein allows it to be easily collected and purified away from other *E. coli* proteins. Therefore, among the first experiments tried was to refold the hydroxylase via a denaturation/renaturation cycle.

Use of hydrogen-bond disrupting salts such as urea and guanidinium hydrochloride as a denaturant and then removing those salts by dialysis or dilution did not afford any detectable soluble hydroxylase. This basic theme was varied by changing the rate of dialysis/dilution, the presence of 'helper' molecules such as thiols, Fe(III) or Fe(II), and the nature of the denaturant. A few

detergents were screened, including N-lauroyl sarcosine, which is reported in the literature as an excellent additive for refolding. The only method that exhibited encouraging results was a technique labeled in the text above as sarcosyl pellet solublization. The detergent N-lauroyl partially solublized the sMMOH aggregate, but removal of the detergent induced the protein to precipitate out of solution.

Since the sMMOH expresses and folds correctly in its native host, *M. capsulatus* (Bath), it seems reasonable to conclude that there is either some factor missing in the *E. coli* cytoplasm that is necessary for correct folding, or some factor present in *E. coli* that is causing misfolding. It follows that an effort to make the *E. coli* cytoplasm more like the *M. capsulatus* (Bath) cytoplasm might cause the sMMOH to fold properly upon expression

One of the most basic differences between the sMMOH native expression system and the designed system is the speed of expression. The native system produces peptides slowly and steadily. The artificial system produces peptides very quickly upon induction. One consequence of this difference is that there is a large number of folding intermediates present in the cell after induction. One might imagine that these folding intermediates might interact with one another, forming long, polymeric, insoluble chains that would aggregate. This model for inclusion body formation is the most prevalent one in the literature (*43*).

To reduce this folding intermediate problem, there are a number of methods one can employ. Reducing the temperature, the IPTG concentration, or both, can greatly reduce the amount of folding intermediates. This method often works in the literature, but was unsuccessful in the sMMOH case. Changing media content has also been successful in some cases, but was not useful for the pDEC002 system.

Changing the expression system has been helpful in producing soluble protein from some inclusion body systems. In particular, with multisubunit proteins, expression of the subunits individually can be of use. Accordingly, a series of plasmids with the three subunits separated was constructed. Emphasis was placed on expression of the MMOHα subunit, as there was the possibility of it being useful on its own, even if expression of MMOHβ and MMOHγ were unsuccessful, or if reconstitution of the three subunits was not possible.

The first vector constructed, pDEC050, was designed to express the native MMOHα. Not unexpectedly, the subunit was expressed as inclusion bodies. The pDEC055 plasmid, constructed as a counterpart to pDEC050, was intended to serve as a source of β and γ subunit. The system failed to produce any proteins, either soluble or insoluble. It seems most likely that the long space between the pTrc promoter and the ribosome binding site, which was enforced by the linker design, caused expression to fail. The plasmid pDEC060 was designed to attach the *pelB* leader sequence onto the α subunit. The *pelB* sequence directs the nascent polypeptide to be secreted into the periplasmic space. The periplasm is a very different environment from the cytoplasm. The protein concentration is lower, and the reduction potential is much more oxidizing. Unfortunately, no soluble protein was afforded by this system.

The next vector made, pDEC070, was much more promising. The construction of the vector expressed a fusion protein of the MMOHα with the protein thioredoxin. Thioredoxin fusions are often quite soluble, and are popular for use in solving inclusion body problems. In addition, the vector contained a His6 site for simple purification and an S-Tag site for simple detection via Western blot. Although approximately 90% of the expressed MMOHα-thioredoxin was insoluble, the rest was soluble. It was quite unstable to protease digestion, however, requiring the presence of the protease inhibitor Pefabloc SC

(Boeringer-Mannheim). Purification on two successive nickel resin columns, with a thrombin cleavage step between the two to separate the thioredoxin from the MMOHα, afforded milligram amounts of pure α with approximately 2 irons per subunit, as expected.

The expressed soluble MMOHα has no measurable hydroxylase activity, however. The recombinant material also exhibited no mixed-valent or reduced signal in the EPR, which are hallmarks of the native hydroxylase. It seems, therefore, that the expressed material is folded only into a soluble form, not a native one. It is not unusual to see small amounts of soluble, non-native, misfolded material produced along with insoluble inclusion bodies. Samples of MMOHα with no measurable properties are of limited utility.

Discussion and Conclusions

The work described above has taught us about the folding pathway of the hydroxylase. The MMOHγ subunit can refold into a soluble unit spontaneously, as shown by refolding experiments. This result is not surprising, given knowledge of the crystal structure. The MMOHγ subunit caps the ends of the flat hydroxylase molecule. A large proportion of its surface area is exposed to solvent. This property contrasts with that of the MMOHα and MMOHβ subunits, which have extensive contacts with one another. It seems likely that the α and β subunits might have to fold together to reach a native conformation.

However, the folding of the MMOHα subunit must proceed normally to a point. The material obtained from the inclusion bodies contained iron at an approximate level of two irons per MMOHα. There are two domains in the α subunit, domain 1 and domain 2 (*16*). It is possible that domain 1, which houses

the iron site and that has the same fold as the soluble R2 subunit, folds well, and that domain 2 does not.

Nevertheless, the hydroxylase does not fold into its native form in the *E. coli* cytoplasm or under *in vitro* conditions. It seems that there is some specific, unknown condition for correct folding that has not been met in the experiments described above. There are several possibilities for the identity of this condition. It may be some cytoplasmic physical condition, such as pH, salt concentration or reduction potential, that was not reproduced in the conditions described above. It may be that some specific *M. capsulatus* (Bath) "foldase" protein is necessary for refolding, such as a chaperonin, or a disulfide bridge reducing protein. It might also be that it is important for the sMMO polypeptides to be produced with specific timing enforced by the *M. capsulatus* (Bath) ribosomal machinery. This might be controlled by the spacing between the ribosomal binding sites and the beginnings and ends of the sMMO genes, or by pauses induced by *M. capsulatus* (Bath) codon preferences.

These possibilities suggest several possible methods for expression of the hydroxylase. The possibilities for use of the *E. coli* organism as a host have not been exhausted. Most of the simple "*in vivo*" experiments have been done. It does not seem that the *E. coli* cytoplasm is a hospitable place for sMMOH refolding. Refolding the collected hydroxylase inclusion bodies outside the cell is still a viable possibility. The refolding work described above does not exhaust all possibilities. A methodical search through possible refolding conditions, varying pH, temperature, salt concentration, thiol concentration and reduction potential, might result in native enzyme.

A second possibility is to use the existing system for producing soluble α subunit and combine it with a system for producing β and γ subunit. Since the existing samples of the α subunit have 2 irons per peptide, as expected, this

subunit is most likely partially assembled. It is possible that combining the expressed α subunit with expressed or native β and γ subunit would produce active hydroxylase indistinguishable from the native material. It is not clear how to recombine the subunits to form active material, because the native hydroxylase cannot be disassembled except under highly denaturing conditions. Nevertheless, this method is a possible avenue for hydroxylase expression.

A third possibility for expression of the sMMO hydroxylase is to extend efforts into other organisms that may have cytoplasmic conditions more like *M. capsulatus* (Bath). The psuedomonad *Psuedomonas pudita* would be a prime choice, since the *M. trichosporium* sMMOH has been expressed successfully in it already (*38*). The methylotrophic yeast *Pichia pastoris* is a possibility. It is a valuable alternative in expression of proteins that have been problematic in *E. coli*, and it may share C1 metabolic enzymes with *M. capsulatus* (Bath). A final option is to express in *M. capsulatus* (Bath) itself. This is, by far, the most challenging option, but it is the only organism that the sMMOH is known to fold correctly in. It is challenging because little is known of its molecular biology, other than it is very difficult to culture on solid media. In addition, there is the problem of *M. capsulatus* (Bath) expressing large amounts of the native sMMOH in addition to the mutant sMMOH. This problem could be solved by expressing the pMMO, instead of the sMMO, by use of culture media containing copper. Alternatively, one could produce *M. capsulatus* (Bath) hydroxylase knock-out strains by using procedures described in the literature (*37*). If one assumes that the technical problems associated with *M. capsulatus* (Bath) could be solved, then use of a wide host range plasmid that contains both the genes necessary for its own replication and the sMMO genes under an inducible promoter would be the favored route for expression.

References

1. Higgins, I. J., Best, D. J., Hammond, R. C., and Scott, D. (1981) *Microbiol. Rev. 45*, 556-590.

2. Higgins, I. J., Best, D. J., and Hammond, R. C. (1980) *Nature 286*, 561-564.

3. Dalton, H. (1992) in *Methane and Methanol Utilizers* (Murrell, J. C., and Dalton, H., Eds.) pp 85-114, Plenum Press, New York.

4. Lidstrom, M. E., and Stirling, D. I. (1990) *Annu. Rev. Microbiol. 44*, 27-58.

5. Best, D. J., and Higgins, I. J. (1983) in *Topics in Enzyme and Fermentation Biotechnology I* pp 38-75.

6. Sheldon, R. A. (1983) *Chemicals from Synthesis Gas*, D. Reidel Publishing Co., Dordrecht, Holland.

7. Valentine, A. M., and Lippard, S. J. (1997) *J. Chem. Soc., Dalton Trans. 21*, 3925-3931.

8. Wallar, B. J., and Lipscomb, J. D. (1996) *Chem. Rev. 96*, 2625-2657.

9. Liu, K. E., and Lippard, S. J. (1995) *Adv. Inorg. Chem. 42*, 263-289.

10. Lipscomb, J. D. (1994) *Annu. Rev. Microbiol. 48*, 371-399.

11. Nordlund, P., and Eklund, H. (1993) *J. Mol. Biol. 232*, 123-164.

12. Stubbe, J. (1990) *J. Biol. Chem 265*, 5329-5332.

13. Fox, B. G., Shanklin, J., Ai, J., Loehr, T. M., and Sanders-Loehr, J. (1994) *Biochemistry 33*, 12776-12786.

14. Powlowski, J., and Shingler, V. (1990) *J. Bact.*, 6834-6840.

15. Yen, K.-M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., Lu, H. S., Harcourt, A. A., and Chen, K. K. (1991) *J. Bact. 173*, 5315-5327.

16. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) *Nature 366*, 537-543.

17. Rosenzweig, A. C., Nordlund, P., Takahara, P. M., Frederick, C. A., and Lippard, S. J. (1995) *Chemistry & Biology 2*, 409-418.

18. Nordlund, P., Dalton, H., and Eklund, H. (1992) *FEBS Lett. 307*, 257-262.

19. Feig, A. L., and Lippard, S. J. (1994) *Chem. Rev. 94*, 759-805.

20. Licht, S., Gerfen, G. J., and Stubbe, J. (1996) *Science 271*, 477-481.

21. Regnström, K., Åberg, A., Ormö, M., Sahlin, M., and Sjöberg, B.-M. (1994) *J. Biol. Chem. 269*, 6355-6361.

22. Rosenzweig, A. C., Frederick, C. A., and Lippard, S. J. (1996) in *Microbial Growth on C1 Compounds* (Lidstrom, M. E., and Tabita, F. R., Eds.) pp 141-149, Kluwer Academic Publishers, Dordrecht.

23. Rosenzweig, A. C., Brandstetter, H., Whittington, D. A., Nordlund, P., Lippard, S. J., and Frederick, C. A. (1997) *Proteins 29*, 141-152.

24. Gupta, N., Bonomi, F., Kurtz, D. M., Ravi, N., Wang, D. L., and Huynh, B. H. (1995) *Biochemistry 34*, 3310-3318.

25. Fox, B. G., Shanklin, J., Somerville, C., and Münck, E. (1993) *Proc. Natl. Acad. Sci. USA 90*, 2486-2490.

26. Broadwater, J. A., and Fox, B. G. (1996) *iNovations 4a*, 8-9.

27. Yen, K.-M., and Karl, M. R. (1992) *J. Bact. 174*, 7253-7261.

28. Teschke, C. M., and King, J. (1992) *Curr. Op. Biotech. 3*, 468-473.

29. McNally, E., Sohn, R., Frankel, S., and Leinwand, L. (1991) *Meth. Enz. 196*, 368-389.

30. Frankel, S., Sohn, R., and Leinwand, L. (1991) *Proc. Natl. Acad. Sci. USA 88*, 1192-1196.

31. Garboczi, D. N., Hung, D. T., and Wiley, D. C. (1992) *Proc. Natl. Acad. Sci. USA 89*, 3429-3433.

32. Zomerdijk, J. C. B. M., Beckmann, H., Comai, L., and Tijan, R. (1994) *Science 266*, 2015-2018.

33. Lu, W.-P., Schiau, I., Cunningham, J. R., and Ragsdale, S. W. (1993) *J. Biol. Chem. 268*, 5605-5614.

34. Wu, W., Rosenzweig, A., and Lippard, S. J. (1989) *Unpublished Results*.

35. Blazyk, J., and Lippard, S. J. (1998) *Unpublished Results*.

36. West, C. A., Salmond, G. P. C., Dalton, H., and Murrell, J. C. (1992) *J. Gen. Microbiol. 138*, 1301-1307.

37. Martin, H., and Murrell, J. C. (1995) *FEMS Microbiol. Lttrs. 127*, 243-248.

38. Jahng, D., and Wood, T. K. (1994) *App. Env. Microbiol. 60*, 2473-2482.

39. Mitraki, A., and King, J. (1989) *Biotechnology 7*, 690-697.

40. Jaenicke, R., and Ruldolph, R. (1990) in *Protein structure: A practical approach* (Creighton, T., Ed.) pp 191-223, IRL Press, Oxford.

41. Marston, F. A. O. (1986) *Biochem. J. 240*, 1-12.

42. Fischer, B., Sumner, I., and Goodenough, P. (1993) *Biotech. Bioeng. 41*, 3-13.

43. Mitraki, A., Haase-Pettingell, C., and King, J. (1991) in *Protein Refolding* (Georgiou, G., and Bernardez-Clark, E. D., Eds.), ACS Symposium Series.

44. Lin, K.-H., and Cheng, S.-Y. (1991) *BioTechniques 11*, 748-751.

45. DeWitt, J. G., Bentsen, J. G., Rosenzweig, A. C., Hedman, B., Green, J., Pilkington, S., Papaefthymiou, G. C., Dalton, H., Hodgson, K. O., and Lippard, S. J. (1991) *J. Am. Chem. Soc. 113*, 9219-9235.

46. Liu, K. E., Johnson, C. C., Newcomb, M., and Lippard, S. J. (1992) *J. Am. Chem. Soc. 115*, 939-947.

47. Hanahan, D. (1985) in *DNA Cloning: A Practical Approach* (Glover, D. M., Ed.) pp 109-35, IRL press, Oxford.

48. Stainthorpe, A. C., Murrell, J. C., Salmond, G. P. C., Dalton, H., and Lees, V. (1989) *Arch. Microbiol. 152*, 154-159.

49. MacFerrin, K. D., Chen, L., Terranova, M. P., Schreiber, S. L., and Verdine, G. L. (1993) in *Meth. Enz.* pp 79-102.

Figure 2-1. Methylotrophs in the carbon cycle.

This schematic depicts the reclamation of small organic compounds into the carbon cycle. The process begins with heterotrophic catabolsis of complex hydrocarbons, typically biological in origin, to small molecules. The products of this process descend to lower layers of the soil where the dioxygen concentration is low. Organisms called methanogens reductively metabolize the small organic waste products to methane, generating energy for themselves. The methane gas produced diffuses towards the surface. Where the dioxygen concentration is high, the methylotrophs can subsist on this methane as a source of carbon and energy, releasing carbon dioxide as a waste product. Thus, carbon is returned to the large pool of CO_2 in the atmosphere.

This cartoon shows one location where methylotrophs can prosper. Any space where methane and oxygen coexist, such as marshlands or geothermal gas vents, could be imagined as a home for these unique organisms.

Figure 2-2. Metabolism of methane in *M. capsulatus* (Bath).

The first step, oxidation of methane to methanol, is catalyzed by MMO. One NADH molecule must be consumed for every turnover. The second step, oxidation of methanol to formaldehyde, is catalyzed by methanol dehydrogenase. A NADH equivalent, $PQQH_2$, is generated by this reaction, making the first two reactions energy neutral up to this point. At the formaldehyde step, the organism can choose between assimilation of the carbon or generation of energy by oxidizing the CH_2O to CO_2 via formaldehyde dehydrogenase and formate dehydrogenase.

A more detailed account of this metabolic pathway along with related reactions can be found at <http://www.expasy.ch/cgi-bin/show_image?D5>.

Figure 2-3. Native sMMOH PAGE.

An SDS-PAGE gel showing the purified native sMMOH from *M. capsulatus* (Bath). The samples vary in quality, and represent a range of the purest protein available at the time.

Figure 2-4. Sample Western Blot.

An SDS-PAGE gel with a matched Western blot showing typical results obtained. The sMMOHα variation shown in the Figure is a sMMOHα-thioredoxin fusion protein discussed later in the chapter. The additional proteins visible in the hydroxylase lane of the Western blot are due to strong interactions between the hydroxylase impurities and antibodies against those impurities in the sMMOH polyclonal antiserum.

Figure 2-5. PCR of the sMMOH genes.

Agarose gel showing the results of PCR amplifications of the *mmoX*, *mmoY* and *mmoZ* genes from *M. capsulatus* (Bath) genomic DNA.

Figure 2-6. Construction of pDEC002.

PCR cassettes for the *mmoX*, *mmoY* and *mmoZ* genes were inserted one at a time into the superlinker plasmid pSL301. The *mmoY** gene was generated from an incorrect PCR primer pair that was designed before it was discovered that the published gene sequence of the *mmoY* gene contained a frame shift mistake (see chapter one). A new primer pair based on the correct sequence was designed and the new *mmoY* gene was inserted into the pSL301-*mmoXY*Z* plasmid. The assembled artificial sMMOH operon was excised and ligated into the pTrc99A gene to make pDEC002.

Figure 2-7. Structure of pDEC002.

Expression of the *mmoX*, *mmoY* and *mmoZ* genes are controlled by the pTrc promoter, based on a combination of the *E. coli* lac and trp promoters. Immediately following the *mmoZ* gene is a small multiple cloning site and the rrnB transcription termination region, designed to prevent run-through transcription. Additional features of the plasmid include the β-lactamase gene for ampicillin resistance, the pBR322 origin of replication, and a copy of the lacIq gene for production of the lac repressor and successful expression in lac^- hosts.

Figure 2-8. Expression of sMMOH genes in *E. coli* results in inclusion bodies.

An SDS-PAGE gel showing the results of expression of the sMMOH genes in pDEC002 in *E. coli* JM105. The sMMOH subunits are clearly visible in only the insoluble lanes. The dark band near the MMOHβ subunit in the soluble fractions is a native *E. coli* protein present in controls.

Figure 2-9. Attempted refolding of expressed, insoluble sMMOH under a variety of conditions.

An SDS-PAGE gel shows the results of several refolding experiments with two different iron sources. Note that the γ subunit refolds.

Figure 2-10. Expression of sMMOH from pDEC002 by using a sarcosyl lysis procedure.

An SDS-PAGE gel shows the results of lysing the *E. coli* containing expressed sMMOH inclusions in the presence of sarcosyl detergent. The photo is a compilation of several lanes from a single gel.

Figure 2-11. Attempted reconstitution of the expressed, insoluble sMMOH by using detergents.

An SDS-PAGE gel shows the results of reconstituting the expressed sMMOH in a variety of detergents.

Figure 2-12. Attempted reconstitution of the expressed, insoluble sMMOH by using sarcosyl detergent.

An SDS-PAGE gel shows the results of reconstituting the expressed sMMOH with sarcosyl detergent. Note that the sarcosyl resolublized a portion of the sMMOH, but removal of the sarcosyl resulted in the reprecipitiation of the protein.

Figure 2-13. Expression by using the pDEC010 plasmid in *E. coli.*

An SDS-PAGE gel shows the result of expression of the sMMOH and B genes in *E. coli.* Note that *mmoZ* does not seem to be expressed in this system. The soluble protein near the α subunit is not MMOHα, but an *E. coli* protein.

Figure 2-14. Construction of pDEC050 and pDEC055.

pDEC050 and pDEC055 were constructed by using similar methods. A fragment (either *mmoYZ* or *mmoX*) was cut out of pDEC002. The remaining plasmid was spliced together by using a short linker. These linkers were designed to remove the single EcoR I site in pDEC002, and insert a BamH I site in its place. This allowed for easy screening of the resulting plasmids.

Figure 2-15. Structure of pDEC050 and pDEC055.

These plasmids are very similar to their parent pDEC002, including the pTrc promoter, the rrnB terminator, ampicillin resistance, and a copy of the lac repressor. pDEC050 was intended for expression of the *mmoX* gene, and pDEC055 was intended for expression of the *mmoY* and *mmoZ* genes.

Figure 2-16. Construction of pDEC060 and pDEC070.

The plasmid pDEC060 was constructed by excision of the *mmoX* gene from pDEC050 and insertion into the pET20b plasmid. Use of the Nco I site allowed fusion with the pelB tag. Due to the design of the pDEC050 plasmid, there are nine amino acids in between the pelB and *mmoX* genes that are part of neither sequence.

The plasmid pDEC070 was constructed by designing a new set of PCR primers tailored for the pET32a plasmid. These primers were used to amplify the *mmoX* gene from the pCH4 cDNA copy of the sMMO operon. The *mmoX* PCR product was directly cloned into pCRII. This procedure was done to solve to problem often seen in earlier cloning work of cutting restriction sites close to the end of PCR products. The *mmoX* gene was excised from pCRII-*mmoX* and ligated into the pET32a plasmid.

Figure 2-17. Structure of pDEC060 and pDEC070.

The pDEC060 plasmid's major feature is a fusion with the pelB leader sequence. This allows proteins that are attached to it to be exported to the periplasm. This can be advantageous in a variety of ways. The periplasm in different from the cytoplasm, in that the protein concentration is lower and the reduction potential is oxidizing. Periplasmic expression also allows for easy harvest of the expressed protein, as the cell wall can be broken via a simple freeze/thaw cycle.

The pDEC070 plasmid has several useful features. It fuses the *mmoX* gene with the thioredoxin gene. This feature is specifically designed to help solubilize difficult to fold proteins. In between the *mmoX* and trxA genes are a number of small useful sequences, including a His6 tag for easy purification, an S-Tag for easy Western blots, and thrombin and enterokinase cleavage sites to remove the thioredoxin fusion at will.

Figure 2-18. Expression of sMMOH genes from pDEC050 and pDEC055.

An SDS-PAGE gel shows the results of expression of the pDEC050 and pDEC055 plasmids in *E. coli.* The pDEC050 plasmid exhibits strong expression of the α subunit in an insoluble form. The pDEC055 plasmid does not show any expression of the β and γ subunits at all. This indicates that the space between the promoter and the RBS of the *mmoY* gene, which is enforced by the linker design, may be too long.

Figure 2-19. Periplasmic expression of the sMMOHα from pDEC060.

An SDS-PAGE gel shows the results of expression of the pDEC060 plasmid in *E. coli.* The pelB-tagged α subunit did not fold correctly, and was found in the insoluble fraction, despite its direction to the periplasm. No soluble protein was visible in the soluble fractions.

Figure 2-20. Expression of the thioredoxin-α subunit fusion protein from pDEC070.

An SDS-PAGE gel shows the results of the expression of the pDEC070 plasmid in *E. coli.* At 37 °C the thioredoxin fusion protein is expressed as insoluble inclusion bodies.

Figure 2-21. Purification of the thioredoxin-α subunit fusion protein.

An SDS-PAGE gel and Western blot show the purification of the thioredoxin fusion protein from the bulk of the soluble *E. coli* proteins by using a Ni^{2+}-NTA column.

Figure 2-22. EPR spectrum of putatively Fe(II) Fe(III) thioredoxin-α subunit fusion protein.

a. EPR spectrum of 150 μM thioredoxin fusion protein prepared under conditions expected to produce mixed-valent diiron center at 7.2 K. b. EPR spectrum of native sMMOH prepared under conditions expected to produce mixed-valent diiron center at 8.0 K. Both spectra: 100 kHz modulation frequency, 6.67 G modulation amplitude, 9.41 GHz microwave frequency, 1 mW microwave power. The signals at g~4.3 and g=2.00 are due to ferric iron impurities and protein-based radical induced by the chemical reduction, respectively. The g=1.93, 1.90, 1.85 signal is due to the antiferromagnetically coupled S=1/2 Fe(II) Fe(III) center in the hydroxylase (*45*).