JOURNAL OF BACTERIOLOGY, Apr. 1993, p. 2436–2444
0021-9193/93/082436-09$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 175, No. 8

# Sequence of the Gene Coding for Ammonia Monooxygenase in *Nitrosomonas europaea*

HUGH McTAVISH,[1,2] JAMES A. FUCHS,[3] AND ALAN B. HOOPER[1*]

*Department of Genetics and Cell Biology,[1] Graduate Program in Biochemistry,[2] and Department of Biochemistry,[3] University of Minnesota, St. Paul, Minnesota 55108*

Received 9 November 1992/Accepted 5 February 1993

***Nitrosomonas europaea*, a chemolithotrophic bacterium, was found to contain two copies of the gene coding for the presumed active site polypeptide of ammonia monooxygenase, the 32-kDa acetylene-binding polypeptide. One copy of this gene was cloned, and its complete nucleotide sequence is presented. Immediately downstream of this gene, in the same operon, is the gene for a 40-kDa polypeptide that copurifies with the ammonia monooxygenase acetylene-binding polypeptide. The sequence of the first 692 nucleotides of this structural gene, coding for about two-thirds of the protein, is presented. These sequences are the first sequences of protein-encoding genes from an ammonia-oxidizing autotrophic nitrifying bacterium. The two protein sequences are not homologous with the sequences of any other monooxygenase. From radioactive labelling of ammonia monooxygenase with [$^{14}$C]acetylene it was determined that there are 23 nmol of ammonia monooxygenase per g of cells. The $k_{cat}$ of ammonia monooxygenase for $NH_3$ in vivo was calculated to be 20 $s^{-1}$.**

*Nitrosomanas europaea* is an autotrophic gram-negative bacterium of soils and fresh water that obtains all of its energy for growth from the oxidation of ammonia to nitrite. This oxidation occurs in two steps. Ammonia is first oxidized to hydroxylamine by ammonia monooxygenase (AMO) (8):

$$NH_3 + O_2 + 2e^- + 2H^+ \rightarrow NH_2OH + H_2O$$

Hydroxylamine is then oxidized to nitrite in an energy-yielding dehydrogenase reaction by the periplasmic enzyme hydroxylamine oxidoreductase (1, 9). The acceptor for the four electrons generated in the latter reaction is the tetraheme cytochrome $c_{554}$, which is thought to carry the electrons in two-electron steps (2). Two of the four electrons are destined for an oxidative electron transfer chain and cytochrome $aa_3$ terminal oxidase (5); the other two must be used in the AMO reaction.

Little is known about AMO. It can be irreversibly inhibited by acetylene, a suicide substrate for many monooxygenases, and inhibition with [$^{14}$C]acetylene radioactively labels a 28-kDa polypeptide that is associated with the membrane fraction (16). AMO is also inhibited by metal chelators, such as thiourea and allylthiourea (10, 20). The fact that these chemicals bind several metals (28), with a preference for copper, has been considered suggestive evidence that AMO might contain copper (3, 30).

AMO activity has been demonstrated with a crude membrane fraction (35, 36), but no further purification has been accomplished. The electron donor for AMO in vivo is unknown, although with the crude membrane fraction it was found that the addition of cytochrome $c_{554}$ stimulated ammonia oxidation activity (35), indicating that cytochrome $c_{554}$ can serve as at least an indirect electron donor for AMO.

AMO is significant because of the key role that it plays in the nitrogen cycle and because it is able to degrade a wide range of hydrocarbons and halogenated hydrocarbons (14, 39, 40). The study of AMO is also important as a model for the related enzyme methane monooxygenase. Methane monooxygenase exists in a soluble form and a membrane form in the methanotrophic bacteria, with the membrane form being more widespread (3). It is thought that AMO and the membrane form of methane monooxygenase are probably homologous enzymes because, in addition to both being membrane enzymes, they have very similar substrate specificities (both enzymes can oxidize both ammonia and methane) and similar inhibitor profiles (3, 4). Also, the membrane methane monooxygenase may, like AMO, be a copper-containing enzyme (27, 31, 34).

In this paper we report the partial purification of the acetylene-binding polypeptide of AMO and the cloning and sequencing of the gene for this polypeptide. We also report the cloning and partial sequence of a second gene in this operon which codes for a polypeptide that copurifies with the acetylene-binding polypeptide.

## MATERIALS AND METHODS

**Bacterial growth.** *N. europaea* was grown in continuous culture (23) in a 55-liter fermentor on medium containing (per liter) 1.33 g of $Na_2HPO_4$, 0.11 g of $KH_2PO_4$, 4 g of $(NH_4)_2SO_4$, 3 ml of a solution containing 68 g of $MgCl_2 \cdot 6H_2O$ per liter and 3.08 g of $CaCl_2 \cdot 2H_2O$ per liter, and 1 ml of a solution containing 1.34 g of $FeSO_4 \cdot 7H_2O$ per liter, 1.0 g of $CuSO_4 \cdot 5H_2O$ per liter, and 1.6 g of $Na_2EDTA$ per liter. The pH was continually titrated to 7.87 with 50% (wt/vol) $K_2CO_3$. The flow rate of medium through the fermentor was 27 liters/24 h. After flowing out of the fermentor, cells were stored at 4°C until concentration by tangential flow across a Pellicon cassette filter unit (Millipore Corp., Bedford, Mass.) and collection by centrifugation.

**Electrophoresis, protein sequencing, and amino acid analysis.** Analytical sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was done on 0.75-mm-thick minigels as described by Porzio and Pearson (26), with the modification that a 5% stacking gel containing 125 mM Tris-HCl (pH 6.8) was added. Preparative SDS-PAGE to purify the AMO acetylene-binding polypeptide for protein sequencing was done by using a single lane across the top of a 1.5-mm-thick minigel; it was done in the same way as

* Corresponding author.

analytical SDS-PAGE except that Tricine replaced glycine in the buffers in order to maintain a lower pH during electrophoresis, which reduces N-terminal modification that blocks protein sequencing (25). After electrophoresis a slice of the gel was stained to determine the position of the 27-kDa polypeptide. This band was then cut out of the unstained gel and electroeluted at 150 V for 3 h in a Centrilutor apparatus (Amicon Corp., Danvers, Mass.) with 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)–NaOH buffer (pH 7.8). The sample was then concentrated with a Centricon 10 microconcentrator (Amicon), washed four times with water, and concentrated to a volume of 80 μl. It was then subjected to Edman degradation with a model 470 gas phase sequencer (Applied Biosystems, Inc.) with on-line phenylthiohydantoin (PTH) analysis.

For amino acid analysis, the acetylene-binding polypeptide was purified in the same way, except that it was eluted from the gel slice by diffusion in 5 mM Tris-HCl (pH 7.5)–0.1% SDS. The purified samples were sealed under a vacuum in 6 N HCl–1.5% phenol, hydrolyzed at 110°C for 24 h, and analyzed with a Beckman model 6300 amino acid analyzer equipped with a 12-cm column by using HiRez Na-E, -S, and -D buffers.

The 40-kDa polypeptide that copurifies with the AMO acetylene-binding polypeptide was sequenced by Edman degradation after SDS-PAGE and blotting of the gel onto an Immobilon 0.45-μm-pore-size polyvinylidene difluoride membrane filter (Millipore), using the procedure described by Matusdaira (24).

**Radioactive or fluorescent labelling of AMO.** [$^{14}$C]acetylene was generated from $Ba^{14}CO_3$ as described by Hyman and Arp (12). Fluorescein isothiocyanate was coupled to propargylamine as follows. Seven volumes of 142 mM fluorescein isothiocyanate in dimethyl sulfoxide was mixed with 3 volumes of 400 mM proparglyamine hydrochloride–250 mM $Na_2CO_3$ in water, and the preparation was allowed to react at room temperature for 2 h. To label AMO, the resulting fluorescein thiocarbamoylpropargylamine (FTCP) (100 μM) was incubated with a 2% (wet wt/vol) suspension of cells in 50 mM $NaPO_4$ (pH 7.5).

**DNA techniques.** *N. europaea* genomic DNA was prepared by the following procedure. Cells (1 g) were digested with lysozyme (15% [wt/vol] cells and 1.5 mg of lysozyme per ml in 50 mM Tris-HCl [pH 8.0]–10 mM EDTA–12.5% sucrose) for 30 min at 37°C and lysed with 0.4% SDS. The extract was treated with DNase-free RNase (9 μg/ml) for 10 min at 37°C and then with proteinase K (0.3 mg/ml) for 40 min and then extracted with phenol, phenol-chloroform, and chloroform at room temperature; the DNA was precipitated with 150 mM NaCl–65% ethanol and redissolved in TE (29).

Agarose gel electrophoresis of DNA was carried out in Tris-acetate-EDTA buffer as described by Sambrook et al. (29). The gels were blotted onto ICN Biotrans$^+$ nylon membranes by using the manufacturer's improved transfer protocol.

Oligonucleotides were obtained from Oligos Etc., Wilsonville, Oreg. For probing of Southern blots, 25 pmol of oligonucleotide was end-labelled with T4 polynucleotide kinase and 12.5 μl of 5'-[$\gamma$-$^{32}$P]ATP (7,500 Ci/mmol; 10 mCi/ml). Southern blot filters were prehybridized and hybridized in standard solutions (29), except that 10% (wt/vol) polyethylene glycol (molecular weight, 8,000) was included in the hybridization solution. After hybridization each membrane was rinsed twice in medium-stringency buffer (0.5× SSPE, 0.2% SDS; 1× SSPE is 150 mM NaCl, 1 mM EDTA, and 10 mM $NaPO_4$, [pH 7.4]) and then washed for 30 min in medium-stringency buffer or high-stringency buffer (0.1× SSPE, 0.2% SDS).

A genomic library was created from *N. europaea* DNA partially digested with *Sau*3AI to yield fragments that were approximately 10 to 20 kb long. This DNA was ligated into lambdaGEM-11 *Bam*HI arms (Promega Corp., Madison, Wis.), packaged into phage with Packagene obtained from Promega Corp., and transfected into *Escherichia coli* KW251.

A size-fractionated library of *Kpn*I fragments approximately 1.6 kb long was created with the pUC119 vector. DNA was size fractionated by agarose gel electrophoresis and then purified from each gel slice with a Prep-A-Gene kit (Bio-Rad Corp., Richmond, Calif.). Following digestion, vector DNA was treated with alkaline phosphatase before ligation to the insert. DNA was transformed into transformable *E. coli* DH5αMCR purchased from Bethesda Research Laboratories, Gaithersburg, Md.

Lambda and plasmid libraries were plated and screened as described by Sambrook et al. (29). Amplification and purification of lambda DNA and plasmids were done as described by Sambrook et al. (29).

Dideoxy sequencing of single- and double-stranded DNAs was carried out with a Sequenase kit obtained from U.S. Biochemicals, Cleveland, Ohio.

**Homology searches.** Protein sequences were used to search for homologous proteins in the GenBank, EMBL, and SwissProt data banks with the FastDB program in the Intelligenetics software package. The IFind program was used for alignments with specific proteins.

**Purification of AMO acetylene-binding polypeptide.** *N. europaea* cells (10 ml of a 20% [wt/vol] suspension in 50 mM $NaPO_4$ [pH 7.5]) were broken by three freeze-thaw cycles. A few grains of pancreatic DNase was added to reduce viscosity. The membranes were sedimented by centrifugation in a Sorvall GSA rotor at 20,000 × *g* for 15 min at 4°C. The pellet was washed three times in 50 mM $NaPO_4$ (pH 7.5) and resuspended in 70 ml of the same buffer. A 23-ml volume of prechilled 20% (wt/vol) Triton X-100 in water was added, and the solution was incubated at 4°C for 15 min and then centrifuged at 25,000 × *g* for 10 min at 4°C to sediment the insoluble material. The supernatant was loaded onto DEAE-Sepharose CL-6B column (2 by 10 cm) equilibrated with 37.5 mM $NaPO_4$ (pH 7.5)–0.5% Triton X-100. The eluate was collected. The AMO polypeptides were precipitated by adding to the eluate an equal volume of 40% (wt/vol) polyethylene glycol (molecular weight, 3,500), incubating the preparation on ice for 40 min, and centrifuging the preparation in a GSA rotor at 25,000 × *g* for 15 min.

**Nucleotide sequence accession number.** The DNA sequence reported here has been deposited in the GenBank data base under accession number L08050.

## RESULTS

**Purification of AMO acetylene-binding polypeptide.** The first step in purifying the presumed active site polypeptide of AMO was to label it with either a radioactively labelled or chromophorically labelled suicide substrate ([$^{14}$C]acetylene and FTCP, respectively). FTCP was synthesized by reacting propargylamine with fluorescein isothiocyanate, which leaves an acetylene functional group attached to a fluorescein label. Both of these suicide substrates were found to specifically inhibit oxidation of ammonia in whole cells without affecting oxidation of hydroxylamine (data not shown), which indicates that they are specific for AMO.

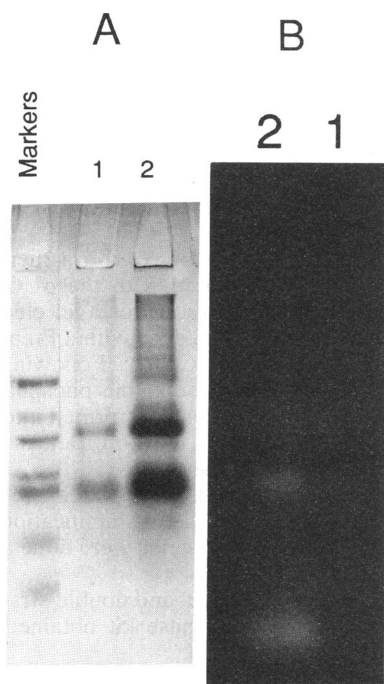


FIG. 1. (A) Coomassie-stained SDS-PAGE gel of DEAE-Sepharose eluate of Triton X-100-solubilized membrane supernatant from cells labelled with FTCP. Lane 1, 7 μg; lane 2, 28 μg. The molecular weights of the markers were (from the top to the bottom) 66,000, 45,000, 36,000, 29,000, 24,000, 20,100, and 14,200. (B) Photograph of the unstained gel shown in panel A when it was illuminated with long-wavelength UV light.

TABLE 1. Amounts of different amino acids found during each cycle of Edman degradation sequencing of the AMO acetylene-binding polypeptide[a]

| Cycle | Residue | Amt (pmol) |
|---|---|---|
| 1 | Ser | 115 |
| 2 | Ile | 92 |
| 3 | Phe | 162 |
| 4 | ? | |
| 5 | Thr | 114 |
| 6 | Glu | 164 |
| 7 | Glu | 204 |
| 8 | Ile | 157 |
| 9 | Leu | 293 |
| 10 | Lys | 229 |
| 11 | Ala | 308 |
| 12 | Ala | 343 |
| 13 | Lys | 238 |
| 14 | Met | 130 |
| 15 | Pro | 191 |
| 16 | Pro | 198 |
| 17 | Glu | 124 |
| 18 | Ala | 295 |
| 19 | Val | 76 |
| 20 | ? | |
| 21 | Met | 30 |

[a] Assuming one acetylene molecule per polypeptide molecule, the preparation was calculated, from its level of radioactivity, to have 461 pmol of acetylene-binding polypeptide prior to being concentrated by ultrafiltration and then sequenced.

After cells were labelled in this way, the label was followed during purification (see Materials and Methods) by optical absorbancy or radioactivity. DEAE-Sepharose chromatography of the Triton X-100-solubilized membranes yielded just two major polypeptides (Fig. 1). When cells had been treated with FTCP, the smaller of these two polypeptides, migrating at 27 kDa, was labelled with the fluorescent chromophore (Fig. 1). When the cells had been treated with [$^{14}$C]acetylene, this polypeptide was radioactive (data not shown).

**N-terminal amino acid sequence of acetylene-binding polypeptide.** The labelled AMO polypeptide was separated from the other major polypeptide in this preparation by SDS-PAGE and was electroeluted from the polyacrylamide gel slice, and the N-terminal sequence was obtained by Edman degradation. The amino acid sequence obtained in this way was SIFRTEEILKAAKMPPEAVCM. Quantitative data support the idea that the sequence determined in this way was indeed the sequence of the acetylene-binding polypeptide rather than a contaminant. The maximum number of moles of amino acid removed during the cycles of sequencing approximately matched the number of moles of acetylene bound to the polypeptide (determined by its level of radioactivity) (Table 1). This preparation was calculated to have 461 pmol of bound acetylene prior to undergoing four cycles of concentration and buffer exchange with water on a Centricon-10 centrifugal ultrafiltration membrane. During sequencing, the yields of some amino acid residues are greater than the yields of other amino acid residues because the former are more resistant to the harsh chemistry of the Edman degradation procedure. Thus, the minimum estimate for the amount of polypeptide is the amount obtained for the amino acid residue whose yield is greatest, in this case 343 pmol of alanine in cycle 12 (Table 1). Allowing for sequencing yield loss due to blockage and termination during the previous cycles of the Edman degradation procedure, and possibly allowing for some sample loss during the concentration and buffer exchange step, this value is in good agreement with the value of 461 pmol of polypeptide that we calculated if we assumed 1 bound acetylene molecule per polypeptide molecule.

**Cloning and sequencing of the AMO gene.** The chemically determined N-terminal amino acid sequence of the 27-kDa AMO polypeptide was used to design a degenerate oligonucleotide probe for the AMO gene matching the amino acid sequence TEEILKAAKMP. This oligonucleotide, AMO oligonucleotide 1, had the sequence 5′-ACI-GA(A,G)-GA(A,G)-ATI-(C,T)TI-AA(A,G)-GCI-GCI-AA(A,G)-ATG-CC-3′ (inosine was used at positions of three- or fourfold degeneracy in the genetic code). AMO oligonucleotide 1 was used to probe Southern blots of *N. europaea* genomic DNA (Fig. 2A). When DNA cut by *Eco*RI or *Sma*I was used, the AMO oligonucleotide was found to hybridize to two different bands. With increasing washing stringency, the probe washed off both bands simultaneously (data not shown), suggesting that the genome contains two copies of the AMO gene. Two bands were also seen after hybridization with Southern blots of DNA digested with *Xho*I, *Sal*I, *Bcl*I, and *Hin*dIII (data not shown). A single band was seen with DNA digested with *Bam*HI and *Kpn*I (Fig. 2A) and with *Pst*I (data not shown), which indicates that in these cases the flanking restriction sites are conserved in the two copies of the gene.

Next, the oligonucleotide probe was used to try to clone one of the copies of the AMO gene. However, screening of a lambda library of *N. europaea* DNA partially digested with *Sau*3AI yielded no complete clones but only clones containing a single 432-bp *Sau*3AI fragment of the gene (Fig. 3). The

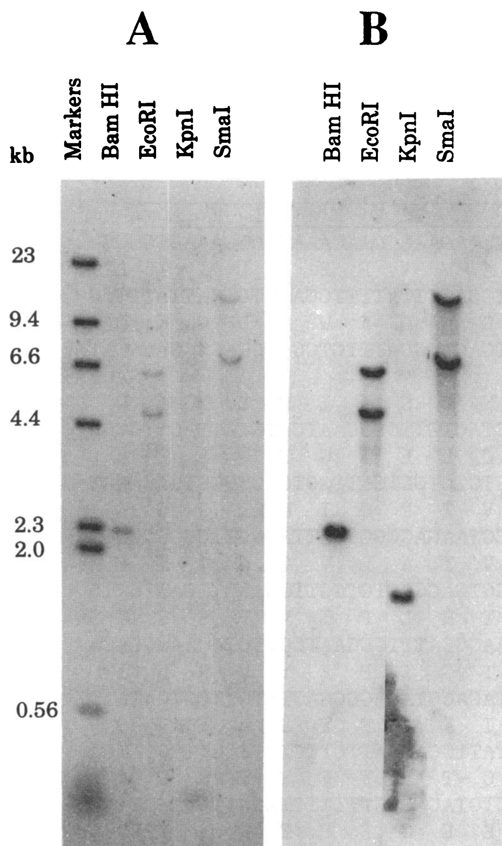


FIG. 2. (A) Autoradiograph of a Southern blot of *N. europaea* cut with four restriction enzymes and electrophoresed through a 0.6% agarose gel. The blot was hybridized with radiolabelled AMO oligonucleotide 1 and after hybridization was washed in high-stringency buffer at 42°C for 30 min. (B) Autoradiograph of the same Southern blot filter used in panel A, except that this filter was hybridized to radiolabelled AMO oligonucleotide 2. After hybridization the filter was washed in high-stringency buffer at 52°C for 30 min.

nucleotide sequence of this fragment matched exactly the chemically determined N-terminal amino acid sequence of the polypeptide.

This *Sau*3AI fragment included a *Kpn*I site, so an attempt was made to clone a *Kpn*I fragment that would contain the remainder of the gene. A 27-mer oligonucleotide, AMO oligonucleotide 2, corresponding to bases 380 to 406 of the sequence shown in Fig. 3, was synthesized, radiolabelled, and used to probe the same Southern blot filter shown in Fig. 2A (Fig. 2B). This oligonucleotide appeared to hybridize to the same one or two bands as AMO oligonucleotide 1 except in the *Kpn*I lane. The *Kpn*I result was expected since the *Kpn*I site at bases 366 to 371 (Fig. 3) lies between the binding sites of the two AMO oligonucleotides. This result confirms that there are two copies of the AMO gene and that the hybridization of AMO oligonucleotide 1 to two bands was not due to fortuitous hybridization of the oligonucleotide to a second, unrelated sequence.

AMO oligonucleotide 2 hybridized to a 1.6-kb *Kpn*I fragment (Fig. 2B), so a plasmid library of size-fractionated 1.6-kb *Kpn*I-digested genomic DNA was created and screened with AMO oligonucleotide 2. This experiment identified several clones, six of which were analyzed by agarose gel electrophoresis and Southern blotting. All six were found to have identically sized 1.6-kb *Kpn*I inserts that hybridized with AMO oligonucleotide 2 and that matched the size of the *Kpn*I fragment from genomic DNA that hybridized with oligonucleotide 2.

One of these clones was sequenced completely. The resulting sequence, combined with the sequence of the 432-bp *Sau*3AI fragment described above, is shown in Fig. 3. The *Sau*3AI and *Kpn*I clones overlapped at bases 366 to 432 (Fig. 3), and in this overlap region their sequences matched exactly. Collectively, the sequence of the two clones shown in Fig. 3 includes an open reading frame at positions 249 to 1079 that should code for a polypeptide having a molecular weight of 31,861; this is approximately the molecular weight of the AMO acetylene-binding polypeptide determined by SDS-PAGE, which was 27,000. We designate this gene *amoA*.

**Amino acid composition of the acetylene-binding polypeptide.** To confirm that the sequence shown in Fig. 3 was not the sequence of a contaminant and that *amoA* is indeed the gene for the acetylene-binding polypeptide, a purified preparation of the polypeptide was analyzed quantitatively for amino acid composition (Table 2). The radiolabelled acetylene-binding polypeptide was purified as described above. From its level of radioactivity and the specific activity of the [$^{14}$C]acetylene, the sample analyzed in the experiment shown in Table 2 was calculated to contain 200 pmol of acetylene and, assuming one bound acetylene molecule per polypeptide molecule, 200 pmol of polypeptide. From this information and from the sequence of *amoA*, we predicted that the sample contained 46.2 nmol of the amino acids analyzed. (Trp and Cys were excluded from the analysis because they were destroyed in the acid hydrolysis; Gly was excluded because glycine was in the electrophoresis buffer and was not completely removed from the sample prior to the amino acid analysis.) We observed 40.13 nmol of these amino acids, 87% of the predicted amount. Table 2 also shows that the relative mole percentages of the different amino acids agreed reasonably well with the mole percentages predicted from the sequence of *amoA*.

**Second gene in the *amo* operon.** Downstream from *amoA* a second open reading frame was found, which continued through the remainder of the 1,516-bp cloned *Kpn*I fragment (Fig. 3). The 40-kDa polypeptide that copurified with the acetylene-binding polypeptide had been sequenced by Edman degradation. Its N-terminal sequence, ?GE?SQEP-FL?M?TVQWYDI, exactly matched the sequence of the protein encoded by the second open reading frame in the AMO clone (Fig. 3). The fact that the gene for this copurifying polypeptide immediately follows the gene for the acetylene-binding polypeptide is evidence that this polypeptide is involved in AMO activity, so we designated the gene *amoB*. The partial sequence of *amoB* shown in Fig. 3 encodes 26.346 kDa (not including the hypothesized cleaved leader sequence) of the polypeptide that has an apparent molecular weight as determined by SDS-PAGE of 40,000.

**Analysis of *amoA* and *amoB* sequences.** The start codon for *amoA* must be the GTG beginning at base 249. GTG, rather than the usual ATG, is the start codon in 7% of the genes of *E. coli* (7). The evidence that this GTG is the start codon includes the fact that it immediately precedes the codon for the first amino acid of the mature protein and the fact that it is preceded by a sequence that matches reasonably well the consensus sequence for prokaryotic ribosome binding sites. Also, the nearest other possible in-frame start codon is the GTG beginning at base 182 (Fig. 3). However, between this

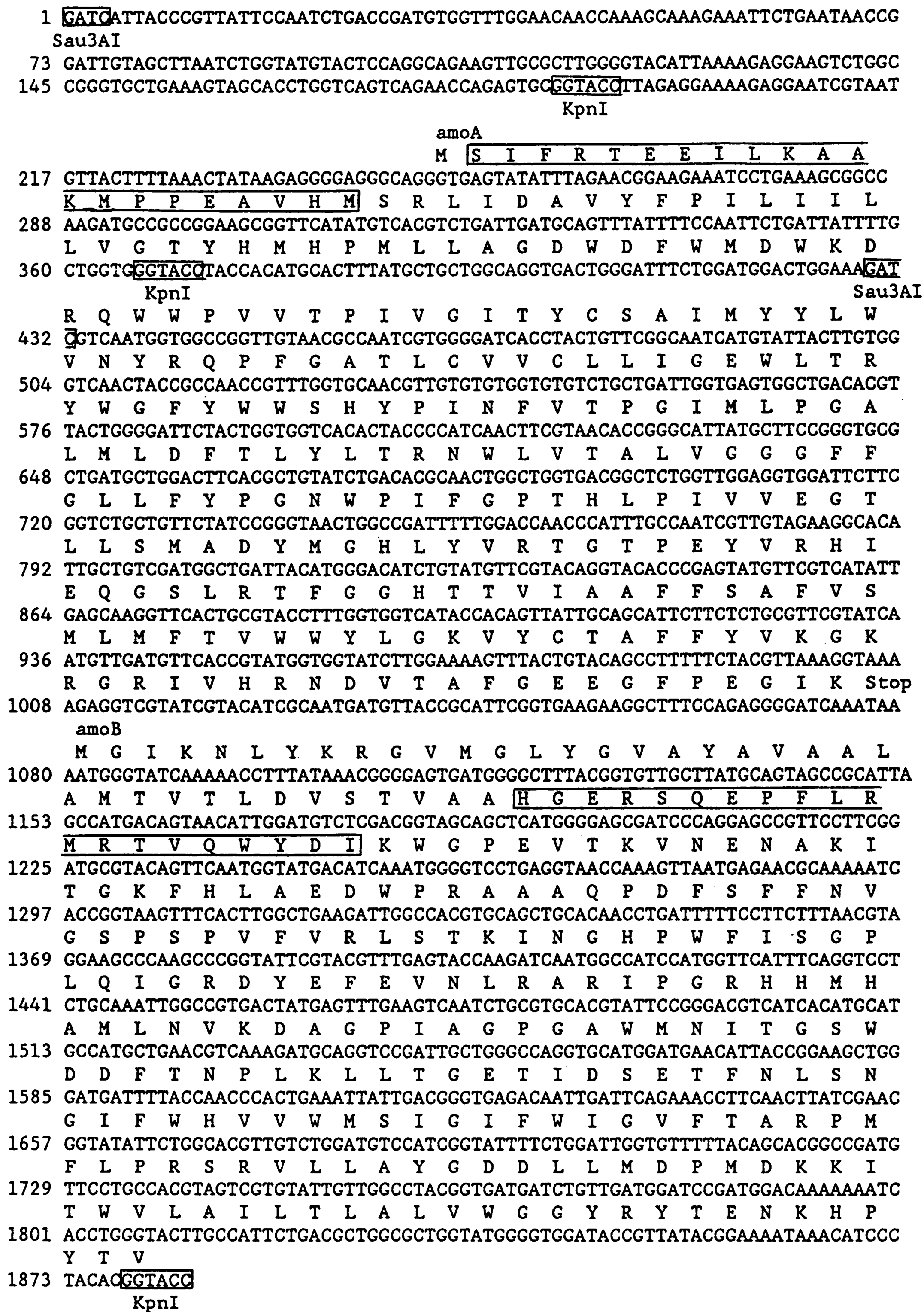


FIG. 3. Combined nucleotide sequence of cloned *Sau*3AI (bases 1 to 432) and *Kpn*I (bases 366 to 1883) fragments of *N. europaea* genomic DNA, identified by AMO oligonucleotides 1 and 2, respectively (see text). The sequence encodes two open reading frames, *amoA* and *amoB*. The amino acid sequences of the proteins encoded by *amoA* and *amoB* are shown above the nucleotide sequence. The portions of the amino acid sequences encoded by *amoA* and *amoB* enclosed in boxes are the chemically determined amino acid sequences of the two polypeptides shown in Fig. 1. *Kpn*I sites and two *Sau*3AI sites are enclosed in boxes, as are the most likely *amoA* and *amoB* start codons.

TABLE 2. Quantitative amino acid analysis of the [$^{14}$C]acetylene-binding polypeptide purified by SDS-PAGE and comparison with the amino acid composition predicted from the sequence of *amoA*

| Amino acid(s) | No. of residues predicted from gene sequence | nmol observed | nmol predicted | mol% observed | mol% predicted |
|---|---|---|---|---|---|
| Val | 23 | 3.8 | 4.6 | 9.4 | 10.0 |
| Ala | 15 | 3.0 | 3.0 | 7.5 | 6.5 |
| Leu | 28 | 4.7 | 5.6 | 11.8 | 12.1 |
| Ile | 18 | 2.8 | 3.6 | 7.1 | 7.8 |
| Pro | 15 | 2.2 | 3.0 | 5.4 | 6.5 |
| Phe | 21 | 3.5 | 4.2 | 8.8 | 9.1 |
| Met | 12 | 1.8 | 2.4 | 4.6 | 5.2 |
| Tyr | 17 | 2.7 | 3.4 | 6.8 | 7.5 |
| Ser | 8 | 2.0 | 1.6 | 5.1 | 3.5 |
| Thr | 20 | 2.6 | 4.0 | 6.4 | 8.7 |
| His | 9 | 1.3 | 1.8 | 3.3 | 3.9 |
| Lys | 7 | 1.6 | 1.4 | 4.1 | 3.0 |
| Arg | 12 | 2.0 | 2.4 | 5.0 | 5.2 |
| Glu | 10 | | | | |
| Gln | 3 | | | | |
| Glu + Gln | (13) | 2.8 | 2.6 | 7.1 | 5.6 |
| Asp | 8 | | | | |
| Asn | 5 | | | | |
| Asp + Asn | (13) | 3.1 | 2.6 | 7.7 | 5.6 |
| Gly | 26 | | | | |
| Trp | 14 | | | | |
| Cys | 4 | | | | |
| Total | | 40.1 | 46.2 | | |

codon and the codon for the first amino acid residue of the mature protein is an in-frame stop codon. Thus, there are no possible start codons other than the GTG at bases 249 to 251.

There are three in-frame ATGs upstream from *amoB*, beginning at bases 1081, 1114, and 1156, plus a GTG at base 1111. Of these, the most likely start codon is probably the one beginning at base 1081, since it is preceded by the best ribosome binding site of the candidates and it almost immediately follows the stop codon for *amoA*. If this is the start codon, the gene would encode a 37-amino-acid leader sequence, which might direct insertion of the polypeptide in the plasma membrane. The *amoA* gene product apparently has an internal signal sequence.

No significant level of overall amino acid identity was observed between the *amoA* or *amoB* gene products and any of the gene products in the protein and DNA sequence data bases. Alignments with particular gene products were also examined in order to detect possible regions of local homology. However, alignment of *amoA* and *amoB* with the proteins encoded by the soluble methane monooxygenase gene cluster (32, 33), the bacterial integral membrane monooxygenases alkane hydroxylase and xylene monooxygenase (17, 18, 37), and the copper-containing monooxygenases tyrosinase (11, 21) and dopamine B-hydroxylase (38) revealed no regions of significant local homology.

Hydropathy plots of the *amoA* and *amoB* gene products (Fig. 4) show that the *amoA* gene product, the acetylene-binding protein, is very hydrophobic and contains four potential membrane-spanning regions. The partial sequence of the *amoB* product is less hydrophobic but still contains at least one possible membrane-spanning region in addition to a quite hydrophobic hypothesized leader sequence.

The pI of the *amoA* gene product was calculated to be 7.9.

**Amount of AMO in *Nitrosomonas* cells and $k_{cat}$ of the enzyme.** Radioactive labelling of AMO in vivo followed by purification of the protein allowed the determination of the concentration of the enzyme in cells. Cells were labelled with [$^{14}$C]acetylene (specific activity, 0.86 mCi/mmol), and membranes were isolated and solubilized with Triton X-100. Insoluble material was then pelleted by centrifugation, and proteins from the soluble material were precipitated with 2.5% (wt/vol) trichloroacetic acid. The trichloroacetic acid pellet was washed twice with water and then counted for radioactivity. The trichloroacetic acid pellet from 1 g (wet weight) of cells was found to contain 42,800 dpm of $^{14}$C. As determined by SDS-PAGE and fluorography, the 27-kDa AMO protein was the only radioactive protein in this material. Assuming one bound acetylene molecule per AMO molecule, this works out to 23 nmol of AMO per g (wet weight) of cells.

This batch of cells consumed 42 nmol of $O_2$ per min per mg (weight weight) of cells with $NH_3$ as the substrate. Assuming 1.5 mol of $O_2$ consumed per mol of $NH_3$ consumed and using the value of 23 nmol of AMO per g of cells, the $k_{cat}$ in vivo for AMO was calculated to be approximately 20 $s^{-1}$.

## DISCUSSION

The use in this study of a fluorescein tag coupled to a suicide substrate functional group was a novel approach to labelling an enzyme active site. However, this technique should be adaptable to other suicide substrates and other enzymes. Any compound with an amino group or other nucleophilic group can be reacted with fluorescein isothiocyanate, which should allow many suicide substrates to be covalently attached to fluorescein and allow enzyme active sites to be fluorescently labelled.

The gene for the AMO acetylene-binding polypeptide and two-thirds of the gene for a 40-kDa polypeptide that copurified with the acetylene-binding polypeptide were cloned and sequenced in this study. Among the evidence that the first gene is the gene for the acetylene-binding polypeptide is the fact that its sequence exactly matches the chemically determined sequence of the N terminus of the protein and the fact that it codes for a polypeptide having a molecular weight of 31,861, which is approximately the molecular weight determined by SDS-PAGE. The finding that the actual molecular weight (31,861) is greater than the SDS-PAGE molecular weight estimates (28,000 [16] and 27,000) is common for membrane proteins (18, 41).

The evidence that the chemically determined N-terminal sequence is the sequence of the acetylene-binding polypeptide rather than the sequence of a contaminant protein is as follows. First, the number of moles of amino acid determined during each cycle of sequencing approximately equalled the number of moles of bound acetylene present. Second, the amino acid composition of a purified sample of the [$^{14}$C]acetylene-labelled polypeptide matches fairly closely the composition predicted from the *amoA* sequence. Furthermore, the quantity of amino acids closely matches the quantity predicted from the level of radioactivity in the sample and the sequence of *amoA*.

A cleaner proof that *amoA* is indeed the gene for the acetylene-binding polypeptide would be to express *amoA* and demonstrate ammonia-oxidizing activity in vitro or in recombinant cells. Unfortunately, however, no effective in vitro assay system has been developed for AMO, and furthermore, we were unable to clone a complete copy of *amoA*, which precludes synthesis of the polypeptide.

The first important finding from the DNA sequence re-

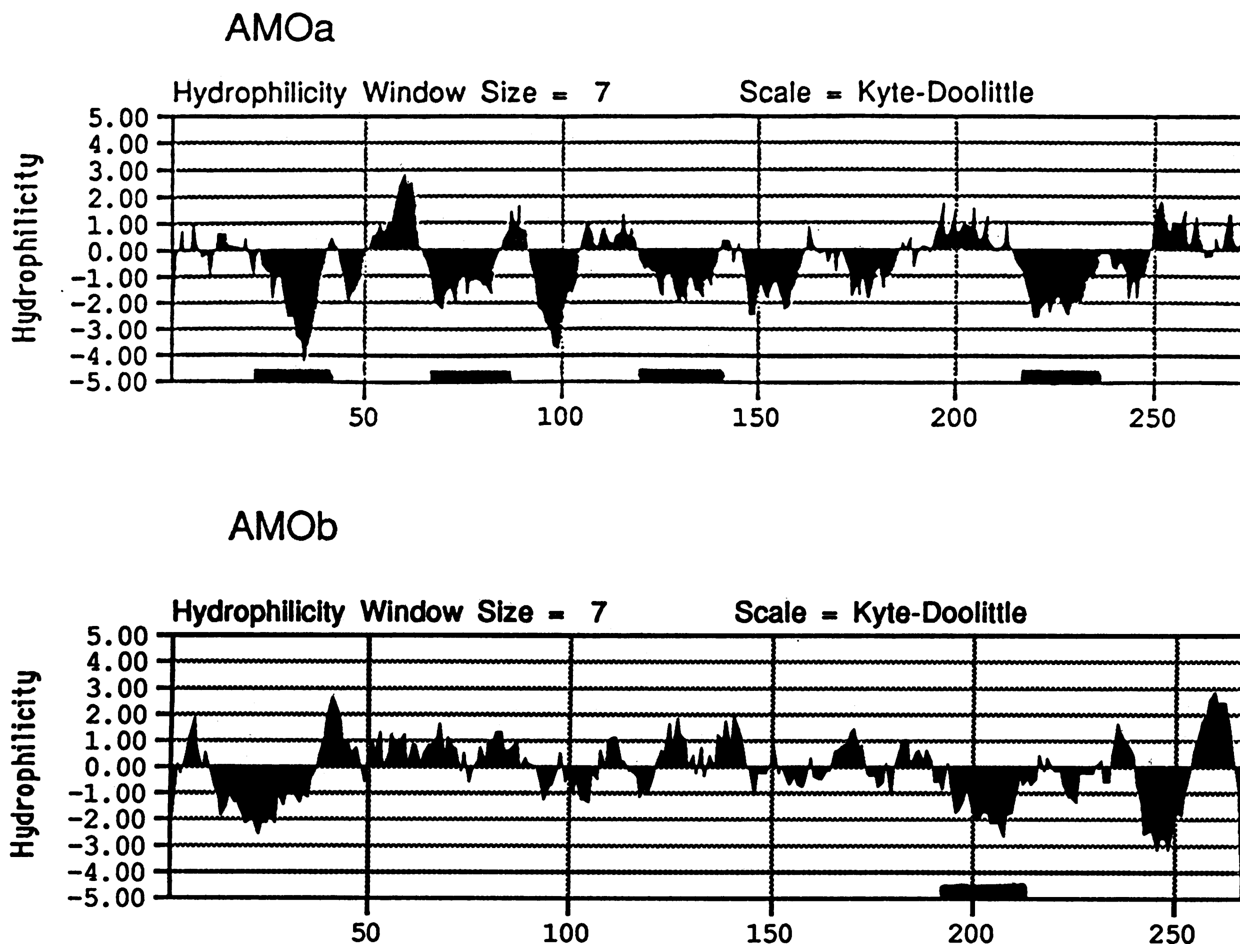


FIG. 4. Hydropathy plots of the *amoA* and *amoB* gene products. The underlined regions are possible membrane-spanning regions. Hydropathy values were calculated over a span of seven residues as described by Kyte and Doolittle (19).

ported in this paper is that the gene for the 40-kDa polypeptide that copurifies with the AMO acetylene-binding polypeptide immediately follows the gene for the acetylene-binding polypeptide and is in the same operon. This is strong evidence that the 40-kDa polypeptide is involved in AMO activity.

From the DNA sequence, AMO appears to be a totally novel type of monooxygenase. It apparently has no homology with any of the previously sequenced monooxygenases, including the soluble methane monooxygenase (32, 33). Since AMO is thought to be homologous to the membrane methane monooxygenase, this supports the idea that the membrane methane monooxygenase is also not homologous to the soluble form of the enzyme. Interestingly, despite the fact that they are thought to contain copper, the AMO polypeptides exhibit no overall homology to the two previously sequenced copper-containing monooxygenases, tyrosinase (11, 21) and dopamine B-hydroxylase (38), and no local homology to the CuA and CuB binding sites of tyrosinase (22). Thus, if AMO contains copper, it may be a novel type of copper center.

Hyman and Arp (13) have suggested that the AMO acetylene-binding polypeptide may have a cleaved leader sequence that directs membrane insertion. When cells were exposed to [$^{14}$C]acetylene during recovery from prior inhibition with acetylene or light, SDS-PAGE and fluorography of cellular proteins revealed two distinct labelled bands migrating near 27 kDa (13). Hyman and Arp speculated that the higher-molecular-weight band might be processed into the lower-molecular-weight band by proteolytic cleavage of a signal sequence. However, we found in this study from the DNA sequence that there is only one possible start codon for *amoA*, which immediately precedes the codon for the first amino acid of the mature protein. This precludes the possibility of a cleaved signal sequence, at least from the product of this copy of the gene. It is conceivable that the second copy of *amoA* encodes a signal sequence.

The calculation of AMO's $k_{cat}$ for ammonia (20 $s^{-1}$) allows a comparison of the kinetic properties of the enzyme with those of the soluble methane monooxygenase. By using other published data this $k_{cat}$ can be used to estimate a turnover number for AMO for methane. In the study of Hyman and Wood (15) the $V_{max}$ of *N. europaea* for ammonia oxidation at a cell concentration of 5.5 mg/ml was 184 μM/min. On the basis of the difference in $O_2$ consumption rates between cells with hydrazine as the substrate and cells with hydrazine plus methane as the substrate, the data of Hyman and Wood show that at the same cell concentration and in the presence of 0.6 mM methane, methane is oxidized at a rate of 28 μM/min. However, this does not represent the

$V_{max}$ because this concentration of methane is below the $K_m$ of 2 mM. When this $K_m$ value is used, the $V_{max}$ for methane (which would require unattainable methane concentrations) is 121 μM/min. This means that the $k_{cat}$ for methane is 121/184 of the $k_{cat}$ for ammonia, or 14 $s^{-1}$. This compares favorably with the $k_{cat}$ of 2.6 $s^{-1}$ for the soluble methane monooxygenase from *Methylosinus trichosporium* OB3b (6). However, the $K_m$ for methane of the soluble methane monooxygenase is much lower than the $K_m$ for methane of AMO (160 μM for the soluble methane monooxygenase from *Methylococcus capsulatus* versus 2 mM for AMO), which makes the velocities of the two enzymes more comparable at lower methane concentrations. At 160 μM methane (if we assume that the $K_m$ of the *Methylosinus trichosporium* enzyme is the same as that of the *Methylococcus capsulatus* enzyme), the soluble methane monooxygenase would turn over at a rate of 1.3 $s^{-1}$, while the AMO rate would be 1.0 $s^{-1}$. Still, this leads to the surprising conclusion that AMO is nearly as good at oxidizing methane as an enzyme evolved for that purpose.

## ACKNOWLEDGMENTS

We thank Myke Logan for suggesting the use of a chromophorically labelled suicide substrate. We thank Frank LaQuier for creating the *Sau*3A lambda library, Greg Pazour and Anath Das for helpful discussions, and Candace Pilon for growing *N. europaea* cells.

This work was supported by the Minnesota Sea Grant Program (grant USDOC/NA90AA-D-56149) and the U.S. Department of Agriculture National Research Initiative Competitive Grants Program (grant 9208667).

## REFERENCES

1. **Andersson, K. K., and A. B. Hooper.** 1983. $O_2$ and $H_2O$ are each the source of one O in $NO_2^-$: $^{15}$N-NMR evidence. FEBS Lett. **164:**236–240.
2. **Arciero, D. M., C. Balny, and A. B. Hooper.** 1991. Spectroscopic and rapid kinetic studies of reduction of cytochrome *c*-554 by hydroxylamine oxidoreductase from *Nitrosomonas europaea*. Biochemistry **30:**11466–11472.
3. **Bedard, C., and R. Knowles.** 1989. Physiology, biochemistry, and specific inhibitors of $CH_4$, $NH_4^+$, and CO oxidation by methanotrophs and nitrifiers. Microbiol. Rev. **53:**68–84.
4. **Colby, J., D. I. Stirling, and H. Dalton.** 1977. The soluble methane mono-oxygenase of *Methylococcus capsulatus* (Bath): its ability to oxygenate n-alkanes, n-alkenes, ethers, and alicyclic, aromatic, and heterocyclic compounds. Biochem. J. **165:** 395–407.
5. **DiSpirito, A. A., J. D. Lipscomb, and A. B. Hooper.** 1986. Cytochrome $aa_3$ from *Nitrosomonas europaea*. J. Biol. Chem. **261:**17048–17056.
6. **Fox, B. G., and J. D. Lipscomb.** 1988. Purification of a high specific activity methane monooxygenase hydroxylase component from a type II methanotroph. Biochem. Biophys. Res. Commun. **154:**165–170.
7. **Gren, E. J.** 1984. Recognition of messenger RNA during translational initiation in *E. coli*. Biochimie **66:**1–29.
8. **Hollocher, T. C., M. E. Tate, and D. J. D. Nicholas.** 1981. Oxidation of ammonia by *Nitrosomonas europaea*: definitive $^{18}$O-tracer evidence that hydroxylamine formation involves a monooxygenase. J. Biol. Chem. **256:**10834–10836.
9. **Hooper, A. B., P. C. Maxwell, and K. R. Terry.** 1978. Hydroxylamine oxidoreductase from *Nitrosomonas*: absorption spectra and content of heme and metal. Biochemistry **17:**2984–2989.
10. **Hooper, A. B., and K. R. Terry.** 1973. Specific inhibitors of ammonia oxidation in *Nitrosomonas*. J. Bacteriol. **115:**480–485.
11. **Huber, M., G. Hintermann, and K. Lerch.** 1985. Primary structure of tyrosinase from *Streptomyces glauascens*. Biochemistry **24:**6038–6044.
12. **Hyman, M. R., and D. J. Arp.** 1990. Small-scale production of [U-$^{14}$C]acetylene from $Ba^{14}CO_3$: application to labelling of ammonia monooxygenase in autotrophic nitrifying bacteria. Anal. Biochem. **190:**348–353.
13. **Hyman, M. R., and D. J. Arp.** 1992. $^{14}C_2H_2$- and $^{14}CO_2$-labelling studies of the *de novo* synthesis of polypeptides by *Nitrosomonas europaea* during recovery from acetylene and light inactivation of ammonia monooxygenase. J. Biol. Chem. **267:**1534–1545.
14. **Hyman, M. R., I. B. Murton, and D. J. Arp.** 1988. Interaction of ammonia monoxygenase from *Nitrosomonas europaea* with alkanes, alkenes, and alkynes. Appl. Environ. Microbiol. **54:** 3187–3190.
15. **Hyman, M. R., and P. M. Wood.** 1983. Methane oxidation by *Nitrosomonas europaea*. Biochem. J. **212:**31–37.
16. **Hyman, M. R., and P. M. Wood.** 1985. Suicidal inactivation and labelling of ammonia mono-oxygenase by acetylene. Biochem. J. **227:**719–725.
17. **Kok, M., R. Oldenhuis, M. P. G. van der Linden, C. H. C. Meulenberg, J. Kingma, and B. Witholt.** 1989. *Pseudomonas oleovorans* alkBAC operon encodes two structurally related rubredoxins and an aldehyde dehydrogenase. J. Biol. Chem. **264:**5442–5451.
18. **Kok, M., R. Oldenhuis, M. P. G. van der Linden, P. Raatjes, J. Kingma, P. H. van Lelyveld, and B. Witholt.** 1989. *Pseudomonas oleovorans* alkane hydroxylase gene: sequence and expression. J. Biol. Chem. **264:**5435–5441.
19. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
20. **Lees, H.** 1952. The biochemistry of nitrifying organisms: the ammonia oxidizing systems of *Nitrosomonas*. Biochem. J. **52:**134–139.
21. **Lerch, K.** 1978. Amino acid sequence of tyrosinase from *Neurospora crassa*. Proc. Natl. Acad. Sci. USA **75:**3635–3639.
22. **Lerch, K., M. Huber, H.-J. Schneider, R. Drexel, and B. Linzen.** 1986. Different origins of metal binding sites in binuclear copper proteins, tyrosinase and hemocyanin. J. Inorg. Biochem. **26:** 213–217.
23. **Logan, M. S. P.** 1991. Ph.D. thesis. University of Minnesota, St. Paul.
24. **Matusdaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262:**10035–10038.
25. **Moos, M., N. Y. Nguyen, and T.-Y. Liu.** 1988. Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. J. Biol. Chem. **263:**6005–6008.
26. **Porzio, M. A., and A. M. Pearson.** 1977. Improved resolution of myofibrillar proteins with sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Biochim. Biophys. Acta **490:**27–34.
27. **Prior, S. D., and H. Dalton.** 1985. Effect of copper ions on membrane content and methane monooxygenase activity in methanol-grown cells of *Methylococcus capsulatus* (Bath). J. Gen. Microbiol. **131:**155–163.
28. **Reid, E. E.** 1963. Organic chemistry of bivalent sulfur, vol. 4 and 5. Chemical Publishing Co., Inc., New York.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Shears, J. H., and P. M. Wood.** 1985. Spectroscopic evidence for a photosensitive oxygenated state of ammonia monooxygenase. Biochem. J. **226:**499–507.
31. **Smith, D. D. S., and H. Dalton.** 1989. Solubilisation of methane monooxygenase from *Methylococcus capsulatus* (Bath). Eur. J. Biochem. **182:**667–671.
32. **Stainthorpe, A. C., V. Lees, G. P. C. Salmond, H. Dalton, and J. C. Murrell.** 1990. The methane monooxygenase gene cluster of *Methylococcus capsulatus* (Bath). Gene **91:**27–34.
33. **Stainthorpe, A. C., J. C. Murrell, G. P. C. Salmond, H. Dalton, and V. Lees.** 1989. Molecular analysis of methane monooxygenase from *Methylococcus capsulatus* (Bath). Arch. Microbiol. **152:**154–159.
34. **Stanley, S. H., S. D. Prior, D. J. Leak, and H. Dalton.** 1983.

Copper stress underlies the fundamental change in intracellular location of methane monooxygenase in methane-oxidizing organisms: studies in batch and continuous culture. Biotechnol. Lett. **5**:487–492.

35. **Suzuki, I., and S.-C. Kwok.** 1981. A partial resolution and reconstitution of the ammonia oxidizing system of *Nitrosomonas europaea*: role of cytochrome c554. Can. J. Biochem. **59**:484–488.
36. **Suzuki, I., S.-C. Kwok, U. Pular, and D. C. Y. Tsang.** 1981. Cell-free ammonia-oxidizing system of *Nitrosomonas europaea*: general conditions and properties. Can. J. Biochem. **59**:477–483.
37. **Suzuki, M., T. Haychawa, J. P. Shaw, M. Rekik, and S. Harayama.** 1991. Primary structure of xylene monooxygenase: similarites to and differences from the alkane hydroxylation system. J. Bacteriol. **173**:1690–1695.
38. **Taljanidisz, J., L. Stewart, A. J. Smith, and J. P. Klinman.** 1989. Structure of bovine adrenal dopamine B-monooxygenase, as deduced from cDNA and protein sequencing: evidence that the membrane-bound form of the enzyme is anchored by an uncleaved signal peptide. Biochemistry **28**:10054–10061.
39. **Vanelli, T., and A. B. Hooper.** 1992. Oxidation of nitrapyrin to 6-chloropicolinic acid by the ammonia-oxidizing bacterium *Nitrosomonas europaea*. Appl. Environ. Microbiol. **58**:2321–2325.
40. **Vanelli, T., M. Logan, D. M. Arciero, and A. B. Hooper.** 1990. Degradation of halogenated aliphatic compounds by the ammonia-oxidizing bacterium *Nitrosomonas europaea*. Appl. Environ. Microbiol. **56**:1169–1171.
41. **Youvan, D. C., E. J. Bylina, M. Alberti, H. Begusch, and J. E. Hearst.** 1984. Nucleotide and deduced polypeptide sequences of the photosynthetic reaction center, B870 antenna, and flanking polypeptides from *R. capsulata*. Cell **37**:949–957.