THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 49, Issue of December 3, pp. 51554–51560, 2004


# Quantitative Proteomic Analysis of Metabolic Regulation by Copper Ions in *Methylococcus capsulatus* (Bath)*[S]

Received for publication, July 15, 2004, and in revised form, September 22, 2004
Published, JBC Papers in Press, September 22, 2004, DOI 10.1074/jbc.M408013200

**Wei-Chun Kao‡§, Yet-Ran Chen§, Eugene C. Yi¶, Hookeun Lee¶, Qiang Tian¶, Keh-Ming Wu‖**, Shih-Feng Tsai‖**, Steve S.-F. Yu§, Yu-Ju Chen§, Ruedi Aebersold¶, and Sunney I. Chan‡§‡‡**

*From the ‡Department of Chemistry, National Taiwan University, Taipei 106, Taiwan, the §Institute of Chemistry, Academia Sinica, Nankang, Taipei 115, Taiwan, the ¶Institute for Systems Biology, Seattle, Washington 98103-8904, the ‖Institute of Genetics and Genome Research Center, National Yang-Ming University, Taipei 112, Taiwan, and the **Division of Molecular and Genomics Medicine, National Health Research Institute, Taipei 115, Taiwan*

**Copper ions switch the oxidation of methane by soluble methane monooxygenase to particulate methane monooxygenase in *Methylococcus capsulatus* (Bath). Toward understanding the change in cellular metabolism related to this transcriptional and metabolic switch, we have undertaken genomic sequencing and quantitative comparative analysis of the proteome in *M. capsulatus* (Bath) grown under different copper-to-biomass ratios by cleavable isotope-coded affinity tag technology. Of the 682 proteins identified, the expressions of 60 proteins were stimulated by at least 2-fold by copper ions; 68 proteins were down-regulated by 2-fold or more. The 60 proteins overexpressed included the methane and carbohydrate metabolic enzymes, while the 68 proteins suppressed were mainly responsible for cellular signaling processes, indicating a role of copper ions in the expression of the genes associated with the metabolism of the organism downstream of methane oxidation. The study has also provided a complete map of the $C_1$ metabolism pathways in this methanotroph and clarified the interrelationships between them.**

Methanotrophs are a unique group of Gram-negative bacteria that grow aerobically on methane and utilize methane as the sole source of carbon and energy. There has been considerable interest in methanotrophs over the past 30 years, since they can be used to produce single-cell bulk chemicals such as propylene oxide. The ability of these bacteria to co-oxidize a wide range of alkanes, alkenes, and substituted aliphatic compounds has been exploited in bioremediation processes, for example, in the degradation of key pollutants such as trichloroethylene in soil and groundwater. Methanotrophs also play an important role in the global methane cycle (1, 2).

Methanotrophs are classified into three distinct types. Type I methanotrophs show disk-like intracellular membranes and use the RuMP (ribulose 5-phosphate) pathway for carbon assimilation. In contrast, Type II methanotrophs possess intracellular membranes in the cell periphery and use the serine pathway to degrade the formaldehyde produced. *Methylococcus capsulatus* (Bath) is classified as a Type X methanotroph (1), since it shows the physiological properties of both Type I and II methanotrophs, but it develops Type I intracytoplasmic membranes.

Two methane monooxygenases (MMO)[1] catalyze the methane oxidation process in methanotrophs, converting methane to methanol. All methanotrophs express a membrane-bound copper-containing particulate MMO (pMMO) (1, 3), while Type X and a few Type II methanotrophic bacteria are capable of producing a second, soluble form (soluble methane monooxygenase, sMMO). Copper ions are known to switch the methane oxidation from sMMO to pMMO. At low copper-to-biomass ratios, the cytoplasmic sMMO is the dominant MMO. When the cells are grown at high copper-to-biomass ratios, pMMO is expressed and produced instead in the plasma membrane.

Because of the unique role of copper ions in regulating the switch between sMMO and pMMO, we have applied cICAT (cleavable isotope coded affinity tag) in combination with off-line two-dimensional LC-MS/MS to quantitatively compare the expression levels of the unique enzymes involved in the featured metabolic pathways under different copper environments. The ICAT technology has been developed for quantitative analysis of proteomic changes in response to cellular perturbations with wide dynamic range and quantitation accuracy (4, 5). More recently, the cICAT reagent has also been developed to eliminate chromatographic isotope effects caused by hydrogen and deuterium (6). Two groups (the National High-Throughput Genome Sequencing Center at the National Yang-Ming University in Taiwan and the Institute for Genomic Research (Rockville, MD) (www.tigr.org/tdb/mdb/mdbinprogress.html)) have been independently making substantial progress toward completing the sequencing of the genome of this methanotroph. These data, although incomplete, have been indispensable to the present proteomic study.

## EXPERIMENTAL PROCEDURES

*Sample Preparation and cICAT Labeling*—The culturing and growth of *M. capsulatus* (Bath) and the separation of the cellular materials into cytoplasmic and membrane fractions were carried out following the

* This work was supported by Academia Sinica and grants from the National Science Council of Taiwan (NSC 91-2113-M-001-045 and 92-2113-M-001-057). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 1–3.

‡‡ To whom correspondence should be addressed: Inst. of Chemistry, Academia Sinica, 128 Academia Rd. Section 2, Nangkang 115, Taipei, Taiwan. Tel./Fax: 886-2-2789-8654; E-mail: chans@chem.sinica.edu.tw.

[1] The abbreviations used are: MMO, methane monooxygenase; pMMO, particulate methane monooxygenase; sMMO, soluble methane monooxygenase; BLAST, basic local alignment search tool; $C_1$, one carbon; ICAT, isotope-coded affinity tag; cICAT, cleavable isotope-coded affinity tag; COG, clusters of orthologous groups of proteins; FBP, fructose 1,6-bisphosphate; $H_4$MPT, tetrahydromethanopterin; KDPG, 2-keto-3-deoxy-6-phosphogluconate; L/H ratio, light-to-heavy ratio; ORF, open reading frame; RuMP, ribulose 5-phosphate; RuBP, ribulose bisphosphate; TMD, transmembrane domain; $\mu$LC-MS/MS, micro-capillary liquid chromatography tandem mass spectrometry; contig, group of overlapping clones.

procedure developed recently (7). Four fractions of proteins were obtained: two cytosolic fractions and two intracellular membrane fractions corresponding to the protein samples derived from cells grown in 0 and 30 $\mu$M copper. All samples were lyophilized for quantitation. Protein samples were prepared at a concentration of 1 mg/ml using the extraction buffer (0.3% SDS, 50 mM Tris, pH 8.3, 5 mM EDTA, and 6 M urea) by weighing out samples of the lyophilized powder. The protein concentration was verified by Bio-Rad protein assay. Labeling of the proteins by the cICAT reagents and following procedures before MS analysis were carried out as described in the literature (8).

*Peptide Separation and Purification*—The peptides were separated by cation exchange chromatography using a 4.6 × 200-mm polysulfoethyl A column (5-$\mu$m particles, 300-A pore size; Poly LC, Columbia, MD) at a flow rate of 200 $\mu$l/min. Peptides were eluted by a gradient of 0–25% Buffer B over 30 min, followed by 25–100% Buffer B over 20 min (Buffer A: 5 mM $K_2HPO_4$, 25% $CH_3CN$, pH = 3.0; Buffer B: 5 mM $K_2HPO_4$, 25% $CH_3CN$, 350 mM KCl, pH = 3.0). The elution profile of the cation exchange chromatography determined which fractions were to be analyzed further, and each were individually processed over avidin cartridges (Applied Biosystems), and the affinity tags were cleaved according to the manufacturer's protocol (cICAT kit for protein labeling; Applied Biosystems) to isolate the labeled Cys-containing peptides for micro-capillary liquid chromatography tandem mass spectrometry ($\mu$LC-MS/MS) analysis.

*$\mu$LC-MS/MS Analysis*—Samples were loaded using an auto-sampler and sequentially analyzed by $\mu$LC-MS/MS. Injections were made on a 1.5 cm × 100 $\mu$m trapping and 12 cm x 75 $\mu$m separation column packed in-house (Magic C18; Micromass BioResources, Auburn, CA). Peptides were eluted with a linear gradient of 5–40% Buffer B over 180 min at an elution rate of ~200 nl/min (Buffer A: 0.1% formic acid in $H_2O$; Buffer B: 0.1% formic acid in acetonitrile). A Hewlett Packard 1100 solvent delivery system with flow splitting (Hewlett Packard, Palo Alto, CA) was used. An LCQ-Deca ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) with an in-house built microspray device was used for all analyses. Peptide fragmentation by collision-induced dissociation was carried out in an automated fashion using the dynamic-exclusion option, and the resultant MS/MS spectra were recorded. The interpretation of the MS/MS data were finally submitted to a suit of software tools for automated data base searching and statistical interpretation of the search results.

*Data Analysis and Bioinformatics*—Genomic sequencing of *M. capsulatus* (Bath) was performed by the National High-Throughput Genome Sequencing Center at the National Yang-Ming University. The entire shotgun sequencing is not finished; however, we have obtained both the 4× and 8× coverage contigs. EMBOSS 2.8.0 (9), NCBI BLAST 2.2.6 (10), and HMMER 2.3.2 (11) were installed on FreeBSD 4.8 RELEASE with one Intel® Pentium® 4 processor operating in the batch mode. Open reading frames (ORFs) were predicted and translated into protein sequences by the GETORF function of the EMBOSS suite, and their homologues protein names were searched against the Swiss-Prot, TrEMBL, and the NCBI non-redundant data bases with the expectation value of 1E-05 as the criterion. 5485 protein homologues were predicted. Since the homologues of the ORFs were not exactly the ones that gave the best similarity (for example, some hypothetical proteins were predicted by genomic sequencing), all the entries were checked manually.

Automated data base searching using SEQUEST (12) software was performed to identify peptide and protein sequence matches for each recorded MS/MS spectrum. Uninterrupted MS/MS spectra were searched against the *M. capsulatus* (Bath) genomic sequencing ORF data base. SEQUEST search parameters for light-cICAT-labeled cysteine were set to 277.13, with a +9 differential modification for heavy cICAT-labeled cysteine and +16 for oxidized methionine, mass tolerance 3 Da. Tryptic digested peptides were specified to limit the search results. Identified peptide scores were validated by Peptide Prophet, and peptides that displayed Peptide Prophet scores >0.9 were further analyzed by Protein Prophet to validate identified open reading frames (13, 14). The light to heavy (L/H) ratios resulting from the differential exprssion were calculated and normalized by ASAPRatio (15) (Automated Statistical Analysis on Protein Ratio) program.

Additional function annotation was performed with the NCBI COGNITOR program on the web (www.ncbi.nlm.nih.gov/COG/), and each protein entry was placed into the highest scoring COG group (16). Protein families were classified by HMMER with the Pfam HMM library. Protein domains were analyzed by the PATMATMOTIFS function of EMBOSS 2.8.0 with the PROSITE data base (17). Protein locations were predicted by the PSORT WWW server (18), and the number of transmembrane domains was predicted by the TMAP function of EMBOSS 2.8.0 (19).

## RESULTS

*Choice of Copper Concentration for the Present Study*—According to our recent study (7), 30 $\mu$M is the optimal copper concentration for the growth and harvesting of cells of *M. capsulatus* (Bath). For the present study, the growth condition in the fermenter was controlled by a hollow-fiber membrane bioreactor (7), and the cells were grown continuously at a copper ion concentration of 30 $\mu$M and kept in the mid-log phase so that a steady state was reached between the copper in the enzyme and the copper concentration in the growth medium. While the cells would grow at lower concentrations of copper ions, the pMMO in the membranes was not homogeneous and did not all contain the full complement of copper ions and exhibit the maximum specific activity. Higher concentrations of copper ions were not used as no additional copper ions would be taken up by the cells (7).

*682 Proteins Revealed in Proteomic Study*—Proteins expressed in the absence and presence of copper ion, 0 and 30 $\mu$M, were labeled with isotopically light ($^{12}$C: L) and heavy ($^{13}$C: H) cICAT reagents, respectively. In summary, a total of 682 proteins of *M. capsulatus* (Bath) were identified with at least one peptide of high confidence value: 429 were from the cytosolic fraction and 253 from the membrane fraction (supplemental Table I). Moreover, it was possible to quantify the change in expression levels of 533 of the 682 identified proteins from the L/H mass ratios of the protein fragments.

87.8% of the identified proteins has significant sequence homology to known proteins in the Swiss-Prot and TrEMBL data bases, including 59 entries (8.7% of the identified proteins) sequenced from *M. capsulatus* (Bath). The remaining 12.2% identified proteins show very low or no detectable homology to other proteins in the data base and are considered poorly characterized. They may represent unique proteins expressed in *M. capsulatus* (Bath) or gene products that have not been discovered previously.

To obtain information regarding the subcellular localization of the 682 identified proteins, the identified ORFs were further analyzed by different algorithms. To predict outer and inner membrane proteins in *M. capsulatus* (Bath), the PSORT algorithm was used. Transmembrane domains (TMDs) were also predicted by the TMAP function of EMBOSS. The analysis showed that the identified proteins were mainly localized in the cytoplasm (412 soluble proteins). Of the 236 proteins predicted to be in the inner membrane fraction, 84 proteins (36%) were predicted to contain two or more transmembrane domains. The fraction of cytosolic proteins with two or more transmembrane domains was 8.7% (36 proteins), indicating that there was some contamination of the cytosolic pool with membrane proteins. The remaining 27 cytosolic proteins were periplasmic, and seven were Gram-negative outer membrane proteins. These results are summarized in Fig. 1.

Aside from pMMO subunit C (see below), other proteins identified in the membrane fraction included the $\alpha$, $\beta$, and $\gamma$ subunits of ATP synthase (L/H = 1.27, 1.1, and 0.99, respectively), NADH dehydrogenase C (L/H = 1.09) and D, F, and G (L/H = 1.43) chains, the proton-translocating pyrophosphate synthase (15 TMDs), and the sodium bile acid symporter family protein (10 TMDs, L/H = 0.58). The cation-transporting ATPase (10 TMDs, L/H = 0.84), magnesium-transporting ATPase (E1 E2 family) (8 TMDs), CopF Cu-ATPase (8 TMDs, L/H = 0.66), the *O*-antigen acetylase (10 TMDs), and the hypothetical protein PA0575 (10 TMDs, L/H = 1.84) were also found, with a large number of transmembrane domains predicted.

All unique protein sequences identified were analyzed by the functional annotation program NCBI COGnitor (Fig. 2). 170 proteins (25%), including all the sMMO and pMMO sub-

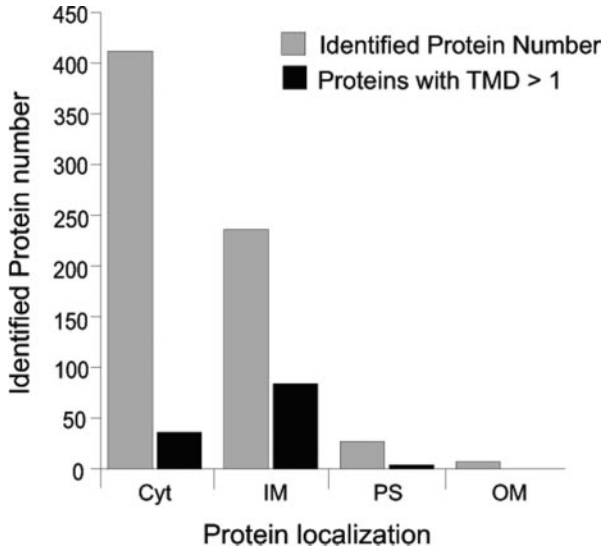


FIG. 1. **Subcellular localization of the 682 identified proteins.** The proteins identified were sorted according to their subcellular localizations and the totals tallied up (*gray columns*). *Cyt*, cytoplasm; *IM*, inner membrane; *PM*, periplasmic space; *OM*, outer membrane of Gram-negative bacteria. The numbers of proteins predicted with TMDs ≥2 in the various localizations are presented as *black columns*.

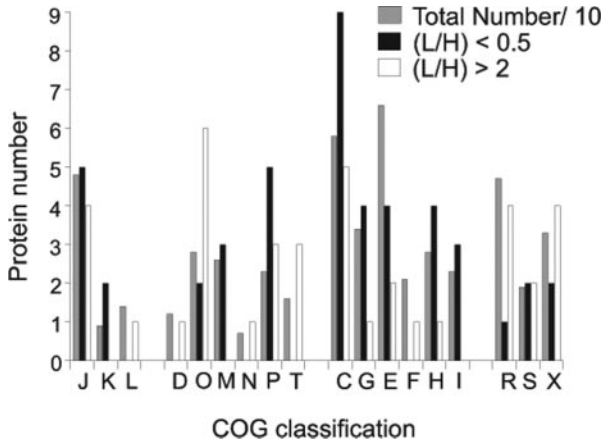


FIG. 2. **COG classification of the identified proteins.** The 682 identified unique proteins were sorted, according to the functional classification of the COG data base, into the following five categories and 17 subcategories: 1) information storage and processing (*J*: translation, ribosomal structure, and biogenesis; *K*: transcription; *L*: DNA replication, recombination, and repair); 2) cellular processes (*D*: cell division and chromosome partitioning; *O*: post-translational modification, protein turnover, chaperones; *M*: cell envelope biogenesis, outer membrane; *N*: cell motility and secretion; *P*: inorganic ion transport and metabolism; *T*: signal transduction mechanisms; 3) metabolism (*C*: energy production and conversion; *G*: carbohydrate transport and metabolism; *E*: amino acid transport and metabolism; *F*: nucleotide transport and metabolism; *H*: coenzyme metabolism; *I*: lipid metabolism); 4) poorly characterized (*R*: general function prediction only; *S*: function unknown; X: multiple hits); and 5) unidentified functions (total 170 entries, with 14 and 29 corresponding to (L/H) <0.5 and (L/H) >2, respectively). *Gray columns* denote the total number of proteins of a given COG category (in units of 10); *white columns* give the corresponding number of proteins with (L/H) >2; and *black columns* show the corresponding number with (L/H) <0.5.

units (see below), could not be classified in this manner, because COG relies on sequenced genomes as the classification data base, and there has been no systematic research and genomic sequencing of methanotrophs until the present study. Although TIGR is sequencing the genome of *M. capsulatus* (Bath) independently, the data available from this source is too limited and inadequate for the large scale proteomic study undertaken here.

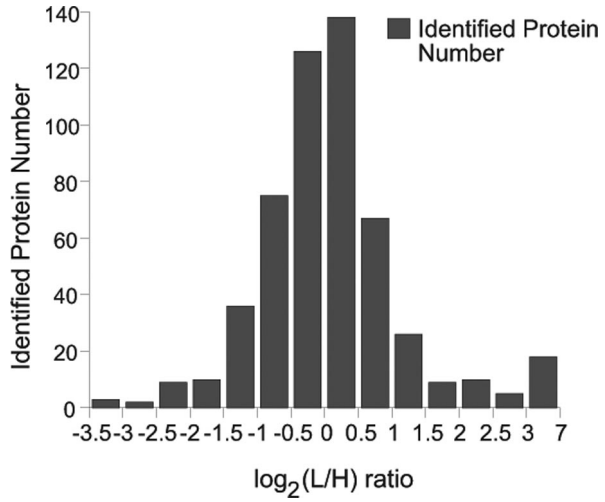


FIG. 3. **Distribution of proteins according to their differential expression levels.** A histogram denoting the number of identified proteins with different ranges of (L/H) ratios is shown.

Although the expression levels (L/H ratios) vary within 2-fold for 82% of the proteins in the present cICAT proteomic study, the composition of the remaining 18% proteins were found to be quite different (Fig. 3). It is clear from the nature of the proteins affected that the metabolism of the cell must be affected by copper ion concentration change. The entire metabolic structure remains the same, but the contributions of the various metabolic pathways are different.

*Proteins Stimulated at 30 μM Copper*—Of the 60 proteins with L/H ratio smaller than 0.5, most of them could be classified as enzymes involved in the biosynthesis of carbohydrates, lipids, cell wall, membrane, envelope biogenesis, coenzymes, metabolism, and inorganic ion transporters (Fig. 2 and supplemental Table IIa). Involved in the direct oxidation of methane to carbon dioxide are pMMO, tungsten-containing aldehyde ferredoxin oxidoreductase (L/H = 0.08), aldehyde dehydrogenase (L/H = 0.23), and tungsten-containing formate dehydrogenase (L/H = 0.23). pMMO (including its subunits PmoA, PmoB, and PmoC) is the most abundant protein produced at 30 μM copper as observed by SDS-PAGE (7). Although cICAT relies on the chemical modification of cysteine for identification, the cysteine content in pMMO is very low. There is no cysteine residue in PmoB, only one in PmoA, and there are three in PmoC. Only the PmoC cysteine-containing C terminus peptide was detected (L/H = 0.09), and we were unsuccessful in observing other cysteine-containing pMMO proteolytic fragments from the transmembrane domains of PmoA and PmoC. The L/H ratio observed was 0.09, indicating that pMMO was overexpressed by about 11-fold at 30 μM copper relative to the level that was detected in the proteome derived from cells grown without added copper, as expected from earlier results (7). Transmembrane domains are difficult to label by the cICAT reagent. In addition, the extraction and digestion of transmembrane peptides remains a technical challenge.

The tungsten-containing aldehyde ferredoxin oxidoreductase (L/H = 0.08) was another protein that was strongly up-regulated by copper. Another strongly induced ORF was identified by BLAST searching as aldehyde dehydrogenase (L/H = 0.23), which is involved in substrate assimilation of a wide variety of aldehydes. The diversity of formaldehyde-related redox proteins may reflect the high amounts of formaldehyde produced at 30 μM copper from methane oxidation by pMMO, due to the high levels of pMMO in the membranes as well as the higher solubility of methane in the intracytoplasmic membranes. The tungsten-containing formate dehydrogenase (L/H = 0/23), the

last enzyme in the pathway of methane oxidation, was also identified in the membrane fraction. Since formaldehyde is toxic to the cell, presumably, these enzymes are up-regulated to remove the formaldehyde efficiently. However, the metal content of these "tungsten" enzymes remains to be addressed.

From the genome sequencing data available, we have located all the genes involved in the tricarboxylic acid cycle. Seven of the enzymes, except succinyl-CoA dehydrogenase, were identified in the cICAT study. Succinyl-CoA synthetase $\beta$ chain (L/H = 0.14) and $\alpha$ chain (L/H = 0.19) (another fragment on another contig was also identified, with a different L/H ratio of 0.43), fumarate hydratase (L/H = 0.31), dihydrolipoamide acetyltranferase (L/H = 0.42), and pyruvate dehydrogenase (the E1 component, L/H = 0.46) were stimulated at 30 $\mu$M copper. It has been reported that *M. capsulatus* (Bath) lacks a complete tricarboxylic acid cycle due to the absence of 2-oxoglutarate dehydrogenase (20); however, the latter has been identified (L/H = 0.59). These results provide further evidence that Type X methylotrophs share some metabolic properties of both Type I and Type II bacteria. Type II methanotrophs use the serine pathway as its major $C_1$ assimilation route with a complete tricarboxylic acid cycle, while Type I methanotrophs do not (1, 20).

It is interesting that some cell wall and capsule synthesis related proteins such as UDP-*N*-acetylglucosamine acyltransferase (L/H = 0.45), UDP-2,3-diacyl-glucosamine hydrolase (L/H = 0.49), 2-dehydro-3-deoxyphospho-octonate aldolase (L/H = 0.19), UDP-glucose/GDP-mannose dehydrogenase (L/H = 0.27), and 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (L/H = 0.44) are stimulated at the higher copper ion concentration. With the cell volume increasing to accommodate the intracytoplasmic membranes, it is possible that the synthesis of additional cell wall material is necessary. Methanotrophic bacteria with pMMO also have higher growth yields than those with sMMO (1). The high copper environment is presumably more compatible with the survival of the *M. capsulatus* (Bath).

Four ion transport proteins are also overexpressed. Three enzymes provide the coenzyme source for other proteins: porphobilinogen deaminase (L/H = 0.39), the fourth enzyme in porphyrin biosynthesis; 6,7-dimethyl-8-ribityllumazine synthase (L/H = 0.43), the enzyme involved in the last step of riboflavin biosynthesis; and quinolinate synthetase (L/H = 0.36), which is involved in NAD biosynthesis.

Hemerythrin (L/H = 0.26) was observed at 30 $\mu$M copper in both the cytosolic and membrane fractions. Hemerythrin is an iron-containing oxygen carrier typically found in marine organisms, but the bacterial fusion domain homologue is also known (21). A BLAST search showed that the hemerythrin domain is conserved. We have successfully purified this hemerythrin, and its UV-visible spectrum is identical to hemerythrins from other sources, indicating the presence of iron.[2] Interestingly, this is the first bacterial hemerythrin isolated. Since the level of expression of the hemerythrin is significantly enhanced at the higher copper concentration, we surmise that hemerythrin functions as oxygen carrier to transport the dioxygen efficiently throughout the cell body for consumption by the pMMO, particularly at the high levels of pMMO produced at 30 $\mu$M copper.

Finally, at 30 $\mu$M copper, there are three protein sequences with no significant similarity in the current data base and six hypothetical proteins. The patterns matched by similarity search are summarized in supplemental Table IIb. An ATP/GTP-binding site motif A (P-loop) motif was found in MCB-2007_160 (predicted open reading frame name) (L/H = 0.39); and MCB-2065_178 (L/H = 0.23) possesses three amidation site motifs and one RGD (arginine, glycine, and aspartic acid) tripeptide pattern (cell attachment sequence). Although their exact functions are unclear, these proteins must be quite important, because their expressions are stimulated at the higher copper concentration.

*Proteins with Higher Expression Levels at 0 $\mu$M Copper*—For comparison, we highlight here the proteins that were expressed at higher levels in the absence of copper. The 68 proteins with L/H ratio >2 appear to affect different parts of the cell physiology than the proteins with ratio <0.5. They tend to be proteins related to post-translational modification, protein turnover, chaperones and signal transduction mechanisms (Fig. 2 and supplemental Table IIIa). In fact, no enzymes related to carbohydrate metabolism were found except sMMO (MmoX, MmoY, MmoB, MmoC, with L/H = 7.23, 4.36, 42.96, and 14.6, respectively; MmoD was also identified, but its L/H ratio could not be determined) and glucokinase (L/H = 2.62). The former is responsible for the first step of methane oxidation and has already been well studied (22). Since *M. capsulatus* (Bath) does not rely on hexose as the carbon source, and there are no reports supporting the ability of *M. capsulatus* (Bath) to grow on glucose, the function of glucokinase might be involved in glycogen biosynthesis or gluconeogenesis.

Two of the four gene translations (MmoG, MmoS, MmoR, and MmoQ) in the sMMO cluster were also found. MmoG (L/H = 18.71) is a 60-kDa chaperonin homologue involved in the assembly of the sMMO complex, while MmoS (L/H = 0.84) is a histidine kinase hypothesized to function as a copper sensor (23). No significant change was observed in the expression level of the histidine kinase.

Some cell signaling and regulatory proteins were highly expressed when the cells were grown without added copper, including the transcriptional regulatory protein zraR (L/H = 3.23), sensory box protein (L/H = 2.44), LacZ expression regulator (L/H = 2.45), and the probable universal stress protein homologue (L/H = 2.42). Two proteins that control cell division were identified, including the cell division inhibitor MinD (L/H = 2.27) and the cell division protease ftsH (L/H = 4.02). Also observed was the peptide methionine sulfoxide reductase (L/H = 15.22) with its SelR domain, which repairs methionine sulfoxide damaged peptides by reducing them to form methionine in conjunction with reductants such as thioredoxin, thus protecting proteins against oxidative stress. Other redox proteins such as glutathione *S*-transferase (L/H = 2.75) involved in glutathione-dependent reactions and the seven-iron ferredoxin (L/H = 9.81) were also detected. Two molecular chaperone-related proteins responsible for folding-corrections were identified.

Several key enzymes involved in metabolism were identified with enhanced abundances at 0 $\mu$M copper. The formylmethanofuran hydrolase (L/H = 2.18) is responsible for the conversion of formyl-$H_4$MPT (tetrahydromethanopterin) to formic acid in the $H_4$MPT pathway (24). GTP cyclohydrolase I (L/H = 4.02) is involved in the first step of the biosynthesis of tetrahydrofolate, which is a $C_1$ transfer intermediate found in bacteria, some archaea and eukaryotes (24). The ATP phosphoribosyltransferase regulatory subunit (L/H = 8.67), imidazoleglycerol phosphate synthase subunit hisF1 (L/H = 2.81) and histidinol dehydrogenase (L/H = 19.95), are enzymes involved in the first, fifth, and the last steps of histidine biosynthesis, respectively. Both denitrification enzymes were observed: nitrate reductase (large subunit, L/H = 2.88; small subunit, L/H = 2.22) and nitrite reductase (L/H = 3.84).

MCB-2060_183 (L/H = 2.15) was BLAST searched as the macrophage migration inhibitory factor found in various eu-

[2] W.-C. Kao and S. I. Chan, unpublished result.

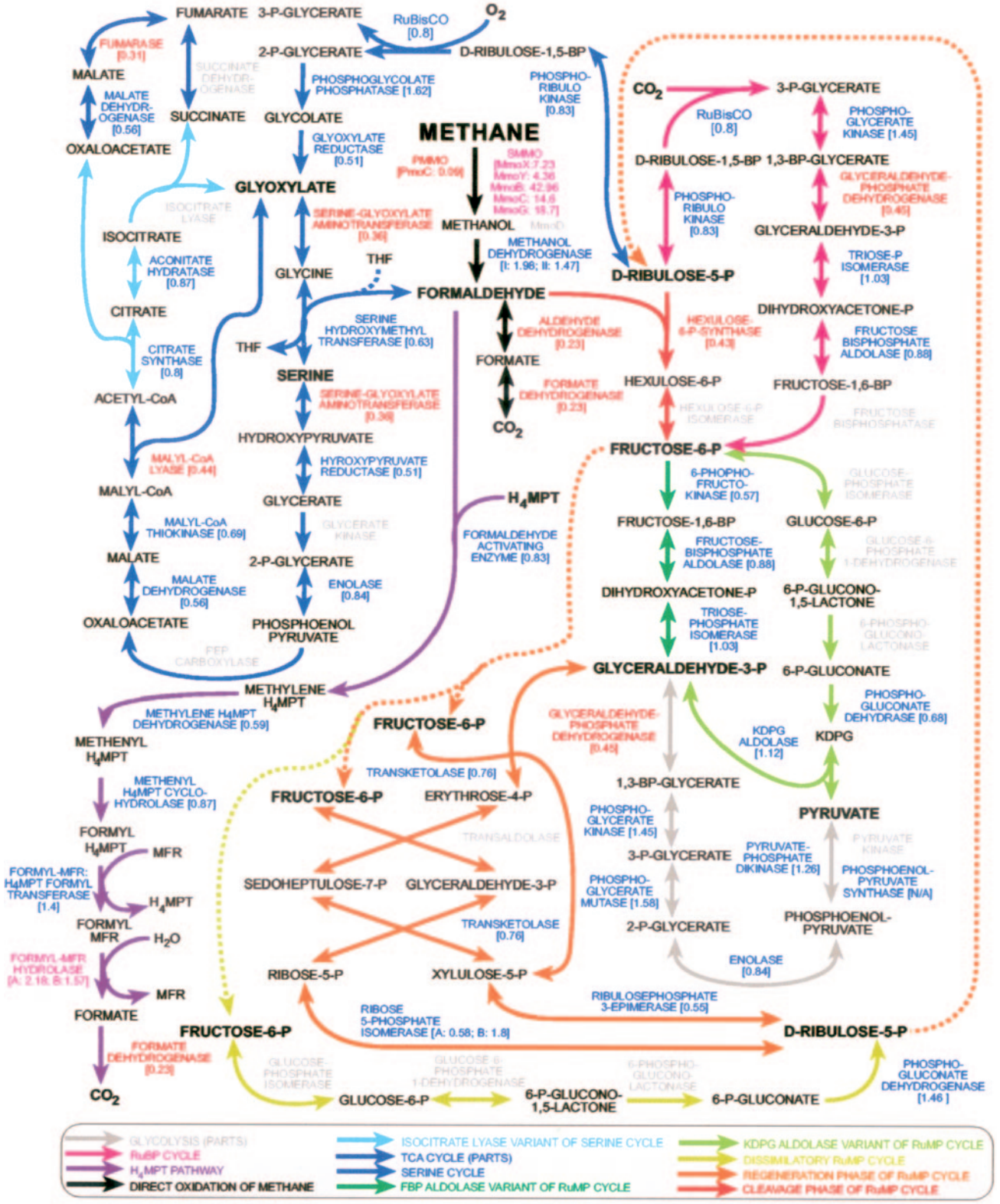


FIG. 4. **The $C_1$ metabolic pathways of *M. capsulatus* (Bath).** Four $C_1$ metabolic pathways assimilating formaldehyde are presented, including direct oxidation of methane, RuMP cycle, serine cycle, and $H_4$MPT pathway. All the enzymes shown were found in the genomic sequence. Proteins identified in the cICAT experiment are highlighted in *blue*. Proteins differentially expressed with L/H ratio <0.5 are shown in *red*, with L/H ratio >2 in *magenta*, otherwise they are shown in *gray*. The L/H ratio(s) are given in *square brackets* for each protein entry, including different subunits (A and B or I and II, according to the Swiss-Prot and TrEMBL data bases), if applicable. Different metabolic pathways are highlighted using different colors, with the color of the *arrows* consistent with the pathway. Some abbreviations are: *-P-*, -phospho-; *-P*, -phosphate; *-BP*, -bisphosphate; *MFR*, methanofuran.

karyotic species (25), but the role of this homologue in bacteria has not been reported and remains to be elucidated in future. We have also observed the surface-associated protein precursor (MopE) (L/H = 6.3) when the cells were grown at 0 $\mu$M copper concentration. The C-terminal of this protein has been reported to be secreted by *M. capsulatus* (Bath) into the growth media, with the possibility of some catalytic and/or copper binding function (26), but this protein was not found in enriched copper-containing environments. Peptidoglycan-associated lipoprotein (L/H = 3.39) is thought to play a role in Gram-negative bacterial envelope integrity. Whether or not copper also controls the outer membrane composition is unclear, but this presents an interesting issue.

Finally, we found 16 proteins with no BLAST hits as well as seven hypothetical proteins whose expression levels were enhanced without copper added to the growth medium (see supplemental Table IIIb). The most interesting one is MCB-2041_433 (L/H = 2.05), which consists of one amidation site motif, one ATP/GTP-binding site motif A (P-loop), and 13 cytochrome *c* motifs. MCB-2043_616 (L/H = 4.92) also shows three amidation sites and one cytochrome *c* motif. Their function(s) should be related to specific redox reactions when the organism is grown at low copper-to-biomass ratios.

## DISCUSSION

*Control of Protein Expression*—The expression patterns of the two MMOs are distinct (7). The outcome of the present proteomic study supports this picture, and further pinpoints the underlying metabolic and physiological differences. The highly expressed levels of the carbon assimilation enzymes in the 30 $\mu$M copper environment no doubt reflect the enhanced expressions of 3-hexulose-6-phosphate synthase (L/H = 0.43), glyceraldehyde-3-phosphate dehydrogenase (L/H = 0.45), pyruvate dehydrogenase (L/H = 0.46), serine-glyoxylate aminotransferase (L/H = 0.36), malyl-CoA lyase (L/H = 0.44); namely, the first enzyme in the RuMP pathway, the second phase of glycolysis, the tricarboxylic acid cycle, the serine pathway, and the first enzyme in isocitrate lyase variant of serine pathway, respectively. The common substrate of these pathways is formaldehyde. The amount of formaldehyde metabolized appears to be affected by which MMO is being used by the organism for methane oxidation. We have identified a pleiotropic regulatory protein (DegT) (L/H = 0.67) that may transfer environmental stimuli to the regulatory region of these genes and activate or repress their transcription, possibly functioning to enhance enzyme production (27).

*The $C_1$ Metabolic Pathways—M. capsulatus* (Bath) assimilates methane as its single carbon source. The methanol oxidation protein (MoxJ) (five TMDs, L/H = 1.25), the gene of which is located between the genes for the two methanol dehydrogenase subunits, might be involved in assembly of the active methanol dehydrogenase and/or its cofactor pyrrolquinoline quinine in the periplasm (28). Formaldehyde is the most important metabolic intermediate in the metabolism of these methanotrophs, because four metabolic pathways, namely, direct oxidation of formaldehyde, the RuMP pathway, the serine pathway, and the $H_4$MPT pathway, all branch from this metabolite (Fig. 4).

Several key enzymes in the RuMP pathway were identified. As part of the dissimilatory RuMP cycle, 6-phosphogluconate dehydrogenase (L/H = 0.68) regenerates RuMP. In the metabolism of methanotrophs, there are two variants of both the cleavage phase (via 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase or fructose-1,6-bisphosphate (FBP) aldolase) and the rearrangement phase (via transaldolase or sedoheptulose bisphosphatase) of the RuMP pathway (20), thus four combinations are theoretically possible. We have identified both the KDPG aldolase (L/H = 1.12) and the FBP aldolase (L/H = 0.88), as well as all the enzymes involved in glycolysis. No sedoheptulose bisphosphatase was found in twice (4x and 8X coverage) of the unfinished genomic sequencing. Thus, both the FBP aldolase/transaldolase and KDPG aldolase/transaldolase variants are possible here, although it is reported that most obligate methanotrophs usually use the latter only.

Both ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco) large chain (L/H = 0.8) and phosphoribulokinase (L/H = 0.83) are key enzymes of the RuBP pathway, but *M. capsulatus* (Bath) does not seem to be able to grow autotrophically (29). The other RuBP pathway enzymes could be found in the FBP aldolase variant of the RuMP pathway. This observation is interesting because it might explain why *M. capsulatus* (Bath) possesses both cleavage phase variants in the RuMP cycle. It is possible that the RuMP pathway is the evolutionary precursor of the RuBP pathway (30). In methanotrophs, the only carbon source is methane rather than glucose. However, the RuMP pathway provides the fundamental framework for carbohydrate metabolism. All metabolites are assimilated to pyruvate, and the tricarboxylic acid cycle is used to complete the process. So, we can assume that glycolysis is immersed in the entire carbohydrate metabolism network.

It is reported that *M. capsulatus* (Bath) has an incomplete serine pathway, and on this basis, it has been classified as a Type X methanotroph (1, 20). However, we have located all the genes of the serine pathway and identified all the proteins as well in the present cICAT study, except glycerate kinase. There is supporting evidence that the serine pathway is connected to the RuBP pathway (20), since phosphoglycolate could be generated by the oxygenase activity of Rubisco, and further converted by phosphoglycolate phosphatase (L/H = 1.62) to glycolate, which is then oxidized to glyoxylate by glyoxylate reductase (L/H = 0.51). The ORF for the glyoxylate reductase is downstream of the neighboring gene of serine-glyoxylate aminotransferase.

Finally, all enzymes of the $H_4$MPT pathway were identified. This is the latest $C_1$ transfer pathway reported in proteobacteria since its original discovery in methanogenic archaeon (24, 31, 32). According to our genomic sequencing, the formylmethanofuran hydrolase (L/H = 2.18 and 1.57 for subunit A and B, respectively) and formylmethanofuran-$H_4$MPT formyltransferase (L/H = 1.41) gene cluster is located next to the pmoC2 of the second pmo gene cluster (33), while the genes of the formaldehyde-activating enzyme (L/H = 0.83), putative tetrahydromethanopterin biosynthesis protein (BLAST searched as Swiss-Prot number AAR90367), and methenyl-$H_4$MPT cyclohydrolase (L/H = 0.93) form another independent cluster in the genome. Since the "formose reaction" was discovered a century and half ago (30), formaldehyde-related compounds have been hypothesized to play a role in early chemical evolution (34–36), and there has been considerable interest in organisms that utilize $C_1$ compounds in biological evolution. The present genomic and proteomic study has allowed us to integrate the four metabolic pathways in methanotrophic carbon assimilation.

*Acknowledgment*—We are grateful to Dr. Lan-Yang Ch'ang for a number of helpful discussions.

## REFERENCES

1. Hanson, R. S., and Hanson, T. E. (1996) *Microbiol. Rev.* **60,** 439–471
2. Trotsenko, Y. A., Ivanova, E. G., and Doronina, N. V. (2001) *Mikrobiologiya* **70,** 725–736
3. Chan, S. I., Chen, K. H.-C., Yu, S. S.-F., Chen, C.-L., and Kuo, S. S.-J. (2004) *Biochemistry* **43,** 4421–4430
4. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) *Nat. Biotechnol.* **17,** 994–999
5. Aebersold, R., Mann, M. (2003) *Nature* **422,** 198–207
6. Li, J., Steen, H., Gygi, S. P. (2003) *Mol. Cell. Proteomics* **2,** 1198–1204
7. Yu, S. S.-F., Chen, K. H.-C., Tseng, M. Y.-H., Wang, Y.-S., Tseng, C.-F., Chen, Y.-J., Huang, D.-S., and Chan S.-I. (2003) *J. Bacteriol.* **185,** 5915–5924

8. Yi, E. C., and Goodlett, D. (2004) *Current Protocol in Protein Science Online* (http://www.wiley.com/legacy/cp/cpps/), Unit 23.2, John Wiley & Sons, Inc., Edison, NJ
9. Rice, P., Longden, I., Bleasby, A. (2000) *Trends Genet.* **16,** 276–277
10. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
11. Eddy, S. R. (1998) *Bioinformatics* **14,** 755–763
12. Eng, J. K., McCormack, A. L., and Yates, J. R., III (1994) *J. Am. Soc. Mass. Spectrom.* **5,** 976–989
13. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) *Anal. Chem.* **74,** 5383–5392
14. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) *Anal. Chem.* **75,** 4646–4658
15. Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) *Anal. Chem.* **75,** 6648–6657
16. Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2001) *Nucleic Acids Res.* **29,** 22–28
17. Hulo, N., Sigrist, C. J., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P., and Bairoch, A. (2004) *Nucleic Acids Res.* **32,** D134–D137
18. Nakai, K., and Kanehisa, M. (1991) *Proteins Struct. Funct. Genet.* **11,** 95–110
19. Persson, B., and Argos, P. (1994) *J. Mol. Biol.* **237,** 182–192
20. Anthony, C. (1982) *The Biochemistry of Methylotrophs*, pp. 296–379, Academic Press Inc. Ltd., London
21. Xiong, J., Kurtz, Jr. D. M., Ai, J., and Sanders-Loehr, J. (2000) *Biochemistry* **39,** 5117–5125
22. Whittington, D. A., and Lippard, S. J. (2001) *J. Am. Chem. Soc.* **123,** 827–838
23. Csáki, R., Bodrossy, L., Kelm, J., Murrell, J. C., and Kovács, K. L. (2003) *Microbiology* **149,** 1785–1795
24. Pomper, B. K., Saurel, O., Milon, A., and Vorholt, J. A. (2002) *FEBS Lett.* **523,** 133–137
25. Sun, H. W., Bernhagen, J., Bucala, R., and Lolis, E. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5191–5196
26. Fjellbirkeland, A., Kruger, P. G., Bemanian, V., Hogh, B. T., Murrell, J. C., and Jensen, H. B. (2001) *Arch. Microbiol.* **176,** 197–203
27. Takagi, M., Takada, H., and Imanaka, T. (1990) *J. Bacteriol.* **172,** 411–418
28. Anderson, D. J., Morris, C. J., Nunn, D. N., Anthony, C., and Lidstrom, M. E. (1990) *Gene* (*Amst.*) **90,** 173–176
29. Baxter, N. J., Hirt, R. P., Bodrossy, L., Kovacs, K. L., Embley, T. M., Prosser, J. I., and Murrell, J. C. (2002) *Arch. Microbiol.* **177,** 279–289
30. Quayle, J. R., and Ferenci, T. (1978) *Microbiol. Rev.* **42,** 251–273
31. Vorholt, J. A., Chistoserdova, L., Stolyar, S. M., Thauer, R. K., and Lidstrom, M. E. (1999) *J. Bacteriol.* **181,** 5750–5757
32. Chistoserdova, L., Chen, S.-W., Lapidus, A., and Lidstrom, M. E. (2003) *J. Bacteriol.* **185,** 2980–2987
33. Stolyar, S., Costello, A. M., Peeples, T. L., and Lidstrom, M. E. (1999) *Microbiology* **145,** 1235–1244
34. Mark, A. S. (2001) *Nature* **414,** 857–858
35. Cooper, G., Kimmich, N., Belisle, W., Sarinana, J., Brabham, K., Garrel, L. (2001) *Nature* **414,** 879–883
36. Kerr, R. A. (2004) *Science* **303,** 1953