Methanogenic archaea are unique in their ability to grow on low-energy substrates, such as acetic acid, by converting them into methane and other by-products. Methanogens are a critical part of the global carbon cycle, consuming by-products of other natural bioprocesses that would otherwise be recalcitrant in sulfate-poor, anaerobic environments (12). They also play an important role in global warming, since methane is a greenhouse gas 20 times as potent as carbon dioxide (42) and methanogenesis is the primary mechanism for the emission of methane into the atmosphere (2).

*Methanosarcina* is the only known genus of methanogens with members that can utilize all of the known methanogenic pathways (acetotrophic, methylo trophic, hydrogenotrophic, and methyl reducing) (71). This metabolic diversity makes these species more permissive to metabolic and genetic manipulations than other methanogens. To capitalize on this characteristic, the genomes of three *Methanosarcina* species have been sequenced (15, 22, 38). In addition, genetic tools have been developed for several of these species, including methods for directed mutagenesis and regulated expression of specific genes (3, 34, 73, 74).

The constraint-based reconstruction and analysis (COBRA) strategy is a powerful paradigm for consolidating large amounts of metabolic knowledge and synthesizing that knowledge into quantitative phenotypic predictions (45, 51). For the performance of constraint-based analysis on an individual organism, its metabolic network is reconstructed from the bottom up, beginning with a sequenced and annotated genome and ending with a network of reactions and reaction-gene associations that directly link genotype and phenotype (68). Many metabolic reconstructions have been curated by hand and have been used to make useful predictions, such as the identification of putative drug targets and the design of novel strains for enhanced biofuel production (45).

In recent years, there have been major advances toward automating much of the reconstruction process (26). Automation is needed to continue the exponential increase in the number of genome-scale metabolic models (43, 45). However, automated reconstructions for methanogens are still particularly problematic for three major reasons: (i) automated predictions tend to be overly homogenized due to their strong reliance on homology-based methods; (ii) reaction and gene databases have more limited coverage of the *Archaea* than of the other domains of life; and (iii) the energy conservation mechanisms of methanogens are highly specialized (14). Hence, manual curation is necessary in order to obtain reliable predictions from metabolic models of these organisms.

Among methanogens, *Methanosarcina acetivorans* is notable for its substrate utilization. It can grow and produce methane by using methylated substrates, carbon monoxide, or acetate, but it cannot grow with hydrogen as its primary energy source (58). Also, unlike most methanogens, *M. acetivorans* is genetically tractable. Therefore, this organism offers opportunities to learn about novel energy conservation mechanisms.

An independent reconstruction for *M. acetivorans* strain C2A has been reported recently (32). The previously reported reconstruction was curated primarily by using an automated curation pipeline, including the GapFind, GapFill, and GrowMatch algorithms (31, 33). Here we present iMB745, an extensively curated manual reconstruction that differs significantly from the other model published. As many literature sources as were available were integrated to generate a highly accurate list of metabolic reactions, making this reconstruction a valuable knowledge base for this organism. In addition, curation has enabled us to make quantitative phenotypic predictions using constraint-based modeling. We demonstrate the usefulness of this model for probing hypo-
MATERIALS AND METHODS

Model reconstruction. An initial list of potential reaction-gene associations in Methanosarcina acetivorans strain C2A was generated based on the integration of data from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (30), MetaCyc (13), the Model SEED reconstruction (26), the Transport Protein Analysis Database (TransportDB) (53), and UniProt (70). Reactions from the existing Methanosarcina Barkeri strain Fusaro reconstruction (19) and the BiGG database (57) were added if there was sufficient evidence for their inclusion, based on sequence homology and/or literature-based curation, or to fill gaps in the annotation. Some gene suggestions from EFIACZ, which includes evidence from other bioinformatics tools, such as PFAM, were also incorporated (4). Gene associations were verified whenever possible by using bidirectional BLASTp against archaean protein products with experimentally verified functions (11). In case of conflicts, metabolic functions suggested from the literature were chosen over those suggested in the databases, and inconsistent reactions were removed from the model (see the supplemental material for a comprehensive list). An SBML file containing the iMB745 model may be downloaded at http://price.systemsbiology.net/downloads_tmp.php.

Reaction and metabolite nomenclature consistent with the BiGG database was utilized whenever possible to facilitate comparison with existing manually curated metabolic models (57). Reactions and metabolites without BiGG identifiers were assigned abbreviations in a manner similar to that used for the database (see the supplemental material for a complete list).

Logical gene-protein reaction (GPR) relationships were constructed manually on the basis of literature or database evidence. For example, genes annotated or characterized as separate subunits of a complex were given an “AND” relationship. If there was no evidence of a protein complex catalyzing a reaction with multiple genes, the genes were all assigned an OR relationship.

All intracellular and transport reactions were computationally mass and charge balanced at a pH of 7 based on charges and formulas computed with ACD/Labs software (version 12; Advanced Chemistry Development, Inc.). Charges and formulas are available in the supplemental material.

Construction of the biomass reaction. The biomass reaction is a sink for essential cell components that represents the consumption of molecular building blocks (such as amino acids and nucleotides) required for cell division. The biomass reaction for Methanosarcina acetivorans strain C2A was modified from that of the closest relative for which a biomass reaction had previously been built, Methanosarcina Barkeri strain Fusaro (19). This biomass objective function was first expanded by incorporating more-detailed carbohydrate data from M. Barkeri (29) and adding methanofuran-B to the list of required cofactors (40). Then coefficients for lipids were modified on the basis of available data on the unique lipid composition of M. acetivorans (61). Nucleotide and amino acid coefficients specific to M. acetivorans were calculated based on the published genome sequence according to established procedures (68). The coefficients of compounds in the soluble pool were assumed to be the same as those in the M. Barkeri biomass equation.

FBA. Exponential-growth phenotypes were predicted using flux balance analysis (FBA), which has been reviewed previously (49). Briefly, all reactions in the model were represented in a stoichiometric matrix, S, in which each column represented a reaction and each row a metabolite. Hence, the entry (ij) of S contained the stoichiometric coefficient of metabolite i in reaction j. If metabolite concentrations are assumed to be constant (steady state), conservation of mass requires that Sv be equal to zero, where v is the vector of reaction fluxes (reaction rates). Because there were more reactions than metabolites in the model (as is typical), multiple flux distributions, all of which satisfied the mass balance, were possible. Reaction fluxes were also constrained by setting minimum and maximum fluxes. In the current study, the reversibility of each reaction was determined based on the literature, database evidence, and thermodynamic calculations. The flux through reversible reactions was unconstrained, while that of irreversible reactions was set to have a minimum v (v_min) of zero. Substrate uptake rates were set to experimentally measured values for purposes of simulations (see the supplemental material for values and references). The reaction rate through the non-growth-associated ATP maintenance reaction (ATPM) was set to 2.5 mmol/g (dry weight)/h to account for upkeep energy costs. This value is somewhat lower than the experimental value of 8.39 mmol/g (dry weight)/h used in the current Escherichia coli model and larger than that in the published M. Barkeri model (19). Growth-associated ATP maintenance was included in the biomass equation and was set to 65 mmol/g (dry weight), similar to the values in the M. Barkeri and E. coli FBA models (16, 19), to account for energy costs for growth (such as the production of macromolecules from biomass components). Both growth-associated maintenance (GAM) and non-growth-associated maintenance (NGAM) costs were chosen to best match experimentally measured growth and secretion rates. The NGAM is about 1.5 mmol/g (dry weight)/h more than that of the previously published M. Barkeri model (19), primarily because ion-pumping inefficiencies for membrane-bound pumps such as Fpo were lumped into the NGAM rather than explicitly stated in the reaction stoichiometry. Detailed calculations and references related to the biomass equation are available in the supplemental material.

Under the assumption that the cell seeks to maximize its growth potential, the specific growth rate was predicted by maximizing the flux through the biomass reaction (max βbiomass) subject to the aforementioned constraints: Sv = 0 and v_min ≤ v ≤ v_max.

Reaction fluxes were predicted in millimoles per gram (dry weight) per hour, and growth rates were predicted per hour. FBA problems were solved using the COBRA toolbox in MATLAB (7) linked to the GLPK linear program solver. Simulations were also repeated using the CPLEX package via the TOMLAB (version 7.0) interface, with identical results. A defined high-salt (HS) medium without a vitamin supplement was used for all simulations. The complete medium composition is listed in the supplemental material; the medium was defined on the basis of the work of Sowers et al. (59).

FVA. FBA does not necessarily yield a unique flux distribution, although it will yield a unique optimal value for the objective function. Flux variability analysis (FVA) was thus used to calculate the possible range of each flux under optimal growth conditions. Mathematically, the possible range of flux through each reaction i was calculated by maximizing and minimizing its flux (v_i) while constraining the objective to be larger than a certain threshold (39). This is performed by minimizing or maximizing v_i subject to the following constraints: Sv = 0, v_min ≤ v ≤ v_max, and βbiomass ≥ pct*βbiomass_MAX, where pct is a percentage of maximum biomass production within which the flux range is to be calculated.

All flux variability analyses in this paper were performed with a pct of 100% (the biomass objective function was fixed to its maximum value, within rounding error). Flux variability analysis was performed using the fluxVariability function in the COBRA toolbox (7).

Knockout lethality studies. For knockout lethality studies, knocked out genes were assigned a value of zero (false) and other genes were assigned a value of 1 (true). The Boolean GPRs were evaluated for every reaction, and reactions with a GPR evaluated as false were removed from the model. After the network was modified in this way, FBA was used to make a growth or no-growth prediction (growth was defined as a predicted βbiomass of >10^-7 m^2 h^-1) for all knockouts of metabolic genes. Lethality predictions were compared to published gene knockout phenotype data (see the supplemental material for references). For substrates with unknown uptake rates (such as monomethylamine), the uptake rate was assumed to be 15 mmol/g (dry weight)/h, similar to the calculated rate for growth on methanol (63), for purposes of FBA simulations.
Calculation of theoretical ATP yield and thermodynamic efficiency. To calculate the theoretical (maximum possible) ATP yield under various conditions, a flux balance analysis problem was solved, but the flux through the non-growth-associated maintenance (ATPM) reaction was maximized instead of flux through the biomass equation. If necessary, the model was forced to carry flux through a specific ATP-generating pathway (such as acetogenesis) by adding constraints on other ATP-generating pathways (such as methanogenesis). The theoretical ATP yield was calculated by dividing the maximum flux through the ATP maintenance reaction by the CO uptake rate. Thermodynamic efficiency was calculated as the ratio of the theoretical ATP yield predicted by the model and the ATP yield that would be observed if the entire Gibbs energy available from the production of methane and CO₂ from a given substrate (under standard conditions) were used to generate ATP (66).

Thermodynamics. Experimental standard Gibbs free energy changes (∆G) were unavailable for most of the reactions in the network, but they were available and were included for some methanogenesis reactions (66, 67) and some reactions involved in central metabolism (1) based on experimental Gibbs free energies of formation. When available, the free energy changes were used to help make decisions about reaction reversibility (in combination with direct experimental or modeling evidence).

To estimate the standard Gibbs free energy changes of reactions for which no experimental data were available, Mol files were generated to represent all compounds in the model containing charged structures at a pH of 7. The Mol files contain the structures and locations of charged moieties in each compound in the model. Charges were computed, and charged Mol files were exported, using ACD/Labs software (version 12; Advanced Chemistry Development, Inc.). The Gibbs energy of each compound was estimated using a previously published group contribution method (28). Standard Gibbs free energies of formation and reaction are reported in the supplemental material under the following conditions: temperature, 25°C; pH 7; ionic strength, zero; water in the liquid phase; and all other compounds in the aqueous phase at a concentration of 1 M.

RESULTS

Model reconstruction. The metabolic network of Methanosarcina acetivorans was curated and validated as described in Materials and Methods. The final network accounts for the activity of 745 metabolic genes and contains 715 intracellular metabolites and 756 reactions (excluding exchange and biomass reactions). The network is considerably larger than that found in the manually curated model of M. barkeri (19) and is comparable in size to that found in the previously existing reconstructions. The iMB745 network includes pathways and genes for several other cofactors, such as NAD, biotin, flavins, and folate.

The metabolic network of M. acetivorans consists mostly of reactions required for the synthesis of amino acids, nucleotides, and cofactors (Fig. 1A). This was not surprising given the relatively small number of growth factors required for the growth of this organism. The reconstructed network includes pathways for synthesizing most of the cofactors required for methanogenesis in Methanosarcina species. The exception is methanophenazine, for which, to our knowledge, no complete synthesis pathway has been proposed in any organism. Synthesis pathways were also included for several other cofactors, such as NAD, biotin, flavins, and folate.

There are still significant gaps in the knowledge of central metabolic pathways. For example, no homologues to currently known IMP dehydrogenase genes could be found in the genome of M. acetivorans, but the reaction catalyzed by this enzyme is predicted to be essential for nucleic acid synthesis. In addition, several of the methanogenic cofactor synthesis pathways that are completely or partially characterized in Methanocaldococcus jannaschii seem to have diverged in Methanosarcina. Although this is perhaps not surprising given the great evolutionary distance between Methanocaldococcus and Methanosarcina, the identification of these differences could provide a motivation for further investigation into the evolution of these species. A detailed discussion of these differences and other identified gaps in metabolic pathways is provided in the supplemental text.

Comparison to existing genome-scale reconstructions of Methanosarcina species. The iMB745 model is the third model of members of the genus Methanosarcina to be published. The first was iAF692, a manually curated model of M. barkeri strain Fusaro (19). The iAF692 model was used to estimate the ion-pumping stoichiometry of the Ech hydrogenase and the ATP requirements of nitrogenase in that organism. The second was iVS941, a genome-scale model of M. acetivorans C2A heavily based on automated approaches (32). The iVS941 model was used to study the essentiality of methanogenesis pathways during growth on CO, acetate, and methanol and to predict ways to reconcile simulated knockout lethality predictions with available data.

Many of the differences between iMB745 and each of the previously published models are due to new literature sources for novel metabolic paths unique to the archaea. For example, both of the previously published models include the nonoxidative portion of the pentose phosphate pathway for the synthesis of five-carbon sugars. However, the genes encoding proteins contributing to reactions in that pathway are apparently absent in many methanogens, including both M. acetivorans and M. Barkeri. An alternative pathway for the synthesis of ribulose-5-phosphate was recently characterized in Methanocaldococcus jannaschii (23); unlike the pentose phosphate pathway, this pathway generates formaldehyde as a by-product. The genes involved in this pathway had strong homology to genes in M. acetivorans. Therefore, the reactions in the pentose phosphate pathway were excluded from the iMB745 model, and the new pathway was added to the model. Other gaps in the previously existing models were also filled based on recent literature (see the supplemental text for details).

The iMB745 model accounts for key differences in methanogenesis pathways between M. acetivorans and M. Barkeri (Fig. 1B to D). Critically, the Fpo and Vht hydrogenases are not functional in these reactions in the model. The automated reconstruction approach for the previous M. acetivorans reconstruction incorrectly included these reactions in the model. The automated model also did not include the recently characterized soluble heterodisulfide reductase (HdrABC), which plays an important role in methanogenesis during growth on methylated substrates (10).

Other important differences also are found between existing models. For example, M. acetivorans cannot grow on H₂ and CO₂, and it grows on CO by using a completely different pathway that involves the secretion of acetate, methyl sulfides, and formate (48, 55). M. acetivorans is also able to grow on dimethyl sulfide (DMS), whereas M. Barkeri can only perform methanogenesis from that substrate (65). The iMB745 network includes pathways and genes necessary to perform these functions, which are not found in either of the previously existing reconstructions. The iMB745 network also includes novel pathways for the synthesis of methano-
furan and cell wall polymers that were newly added to the biomass equation (see Materials and Methods) and accounts for experimentally determined differences in lipid composition from \textit{M. barkeri} (61). Therefore, the description of essential biochemistry is more complete in this model than in those published previously.

**Estimation of Rnf and Mrp ion-pumping stoichiometry.** The stoichiometry of ion pumps in the electron transport chain can have a significant effect on the quantitative predictions of metabolic models (16). In lieu of experimental data, it becomes necessary to estimate the stoichiometry by simulation. In the \textit{M. acetivorans} model, flux balance analysis was used to estimate the stoichiometry of ion exchange due to the H\textsuperscript{+}/H\textsubscript{2}-exchanging complex Mrp and the Rnf complex, two methanogenesis enzymes found in \textit{M. acetivorans} but not in \textit{M. barkeri} (54). The Rnf complex is thought to catalyze the reduction of methanophenazine by ferredoxin and to generate either a proton or a sodium motive force (37). The specific ion pumped by Rnf is unknown, but Mrp is strongly upregulated on acetate compared to methylotrophic substrates, suggesting an increased importance of H\textsuperscript{+}/Na\textsuperscript{+} exchange across the membrane on that substrate (54). Since the ion-pumping activity of Rnf is essential for growth on acetate, and since Mrp and Rnf are coregulated (54), it was assumed for modeling purposes that Rnf pumps sodium ions.

The ATP yield of methanogenesis on substrates that utilize Rnf is strongly dependent on both the number of sodium ions pumped by Rnf and the H\textsuperscript{+}/Na\textsuperscript{+} exchange ratio of Mrp, neither of which has been determined experimentally. H\textsuperscript{+}/Na\textsuperscript{+} exchange proteins are known in other organisms that pump protons and sodium ions in 2:1 (64), 3:2 (50), or 1:1 (5) ratios. To find out which combination was most likely in light of experimental growth data, a sensitivity study was performed, varying the number of sodium ions pumped by Rnf as well as the H\textsuperscript{+}/Na\textsuperscript{+} ratio of Mrp. The closest match between the predicted and experimental growth yields occurred when Rnf was set to pump 3 Na\textsuperscript{+}/2 e\textsuperscript{-} and when Mrp was set to pump 1 H\textsuperscript{+}/Na\textsuperscript{+} (Fig. 2) (see also the supplemental material). Changes in these ratios had only a minimal effect on predicted product secretion rates (see the supplemental material). Due to the close match with experimental growth yields, these ratios were chosen for all further simulations.

The calculated ratios are consistent with thermodynamic data.
The Rnf complex utilizes the same electron acceptor (methanophenazine) as the F_{420} dehydrogenase (Fpo) but uses ferredoxin instead of F_{420} as the electron donor. The reaction catalyzed by Fpo is coupled to the pumping of two protons across the membrane (6). Since ferredoxin has a lower redox potential than F_{420}, it is reasonable to expect that Rnf can pump more proton equivalents across the membrane than Fpo.

**F_{420} regeneration during growth on carbon monoxide.** Both *M. acetivorans* and *M. barkeri* grow on carbon monoxide by oxidizing it to CO$_2$ and subsequently reducing CO$_2$ to methane (21). In methanogens, the reduction of carbon dioxide to methane requires the oxidation of two equivalents of reduced coenzyme F$_{420}$. Therefore, in order for a methanogen to grow on carbon monoxide, a mechanism for generating reduced coenzyme F$_{420}$ must be present in the cell. In CO-grown *M. barkeri*, reduced F$_{420}$ is probably regenerated by the generation of molecular hydrogen via the reverse action of Ech hydrogenase, followed by H$_2$-dependent reduction of F$_{420}$ via the F$_{420}$-reducing hydrogenase Frh (41, 46). Ech hydrogenase is not present in the *M. acetivorans* genome, and although an *frh* operon is present, it does not encode a functional enzyme (24). Thus, the mechanism for F$_{420}$ regeneration in *M. acetivorans* during growth on CO remains unknown (21).

Our model suggests that F$_{420}$ is regenerated by the combined action of F$_{420}$ dehydrogenase (Fpo) and the Rnf complex (Fig. 3A). In the proposed pathway, Rnf would reduce methanophenazine with ferredoxin, and subsequently, reverse electron transport via Fpo would be used to generate reduced F$_{420}$. Reverse electron transport by Fpo has not been observed experimentally, but it is thermodynamically feasible in an environment containing excesses of oxidized F$_{420}$ and reduced methanophenazine. In addition, this hypothesis is consistent with the high levels of Fpo protein and transcript measured during growth on CO (35). Finally, the estimated proton-pumping stoichiometries for Rnf and Fpo (3 and 2 proton equivalents, respectively) suggest that *M. acetivorans* would conserve one proton for each unit of F$_{420}$ reduced. This is consistent with the level of conservation in *M. barkeri*, in which the Ech hydrogenase pumps at least one proton out of the cell (72).

As an alternative hypothesis, we also tried to implement a F$_{420}$-ferredoxin oxidoreductase reaction (FFRED reaction) for the purposes of regenerating coenzyme F$_{420}$ during growth on CO (10): $\text{F}_420-2(\text{aq}) + 2\text{fdred}(\text{aq}) + h(\text{aq}) \rightarrow \text{F}_420-2\text{h}_2(\text{aq}) + 2\text{fdox}(\text{aq})$, where $\text{F}_420-2$ is oxidized F$_{420}$, $\text{F}_420-2\text{h}_2$ is reduced F$_{420}$, fdox is ferredoxin, fdred is reduced ferredoxin, and $h(\text{aq})$ indicates the aqueous phase. Growth on CO was predicted to be possible if reaction FFRED was added to the model (data not shown). However, the presence of FFRED was also predicted to make a Δrnf mutant viable on acetate, contrary to experimental evidence (N. R. Baun, A. M. Guss, G. Kulkarni, and W. W. Metcalf, unpublished data), and therefore, the reaction was not included in the model. According to the model, a Δrnf mutant growing on acetate could survive with a lower growth rate by reducing coenzyme F$_{420}$ with FFRED and then generating a proton gradient with Fpo and heterodisulfide reductase (Hdr).

It is possible that an enzyme catalyzing a reaction like FFRED really exists but that the ATP yield of this alternative path is insufficient to make it viable for growth on acetate. According to the model, the theoretical maximum wild-type ATP yield of methanogenesis from acetate (utilizing Rnf) is only about 0.75 mmol ATP/mmol acetate, corresponding to ~65% thermodynamic efficiency (at pH 7) or 30% efficiency (at pH 0) (66). The efficiency is considerably higher than that of methanogenesis from carbon monoxide (0.56 mmol ATP/mmol CO [34%]) or methanol (0.75 mmol ATP/mmol methanol [27%]), possibly because the cell is not viable with less-efficient ATP production from this substrate.

**Consistency of the model with knockout lethality data.** Comparison of knockout lethality predictions with available data indicates that the model correctly predicts the growth/no-growth phenotypes of 60/63 knockout mutants (Table 1). In all cases of incorrect predictions, genes were shown experimentally to be lethal but were predicted to be nonlethal. Two of the incorrect lethality predictions involved actogenesis during growth on carbon monoxide. The genes encoding Pta and Ack are essential for growth on CO (55). However, flux balance analysis incorrectly predicts that a Δpta Δack mutant can grow by producing methane.
and carbon dioxide as sole by-products. Although inhibition of Mtr and (if the Mtr bypass indeed exists) another reaction in methanogenesis would cause the \( \text{pta}^{-1} \text{ack}^{-1} \) mutation to be lethal, we cannot rule out the possibility that other mechanisms, such as regulatory constraints, are responsible for the essentiality of these genes. There is physiological evidence both supporting (55) and refuting (48) inhibition of methanogenesis reactions by carbon monoxide.

A \( mch^{-1} \) mutant was predicted to be viable on acetate (Table 1), but this knockout is known to be lethal on that substrate (25). The \( mch \) gene was hypothesized to be essential for growth on acetate because it generates reduced F420, which \( M. \text{acetivorans} \) requires for use in anabolic reactions, such as that involving the F420-dependent glutamate synthase (52). However, during growth on CO, Mch cannot be used to generate reduced F420 because it is required to carry flux in the direction of F420 oxidation. Therefore, for growth on CO, another enzyme (in the model, this is predicted to be Fpo) that is able to reduce F420 must be present. Flux balance analysis predicts that this other enzyme could also be used to reduce F420 during growth on acetate, thus making \( mch \) nonessential for growth on acetate. However, the incorrect prediction depended on the ability of \( M. \text{acetivorans} \) to secrete methyl sulfides during growth on acetate, which is unlikely given that the Mts methyltransferase required for methyl sulfide synthesis is downregulated during growth on acetate (9, 35). Hence, inclusion of the regulatory constraint would fix the phenotype prediction.

Knockout data were useful for refining the model and finding gene annotations in the \( M. \text{acetivorans} \) genome that are inconsistent with experimental data. For example, \( M. \text{acetivorans} \) uses Ack and Pta to activate acetate to acetyl coenzyme A (acetyl-CoA) during acetoclastic methanogenesis and cannot grow on acetate without the encoding genes (55). However, the \( M. \text{acetivorans} \) genome also encodes genes (MA3168 and MA3602) with high sequence identity to the ADP-forming acetyl-CoA synthase of \( \text{Methanocaldococcus} \) jannaschii, which catalyzes an alternative pathway for activating acetate (44). Inclusion of this reaction would make Pta and Ack nonessential for growth on acetate. On this basis, the reactions in the alternative pathway were excluded from the model.

**Consistency of the model with growth phenotype data.** The iMB745 model was used to predict growth phenotypes for wild-type strains of \( \text{Methanosarcina acetivorans} \) growing on acetate, methanol, and carbon monoxide, the three substrates for which growth and substrate uptake data are available (55, 60, 63). Predicted growth rates were highly dependent on substrate uptake.
rates, which differed as much as 2-fold depending on the data set used to perform the calculation (see the supplemental material). It was possible to select uptake rates within the experimentally feasible ranges for each substrate that matched the observed growth rates and growth yields within 20% (Table 2).

Both the growth rates and the secretion rates during growth on acetate closely matched the experimental values. On methanol, the rate of methanogenesis was predicted to be much lower than the experimental value, but the ratios of products are consistent with experimental data. When the ATP yield was maximized during growth on methanol, the predicted ratio of methane to CO₂ produced was exactly 3:1, as would be expected to balance redox potentials in the cell (10). When conditions were optimized for growth, the actual ratio of methane to CO₂ secreted was predicted in the model to be 3.8:1, because the carbon dioxide-fixing activity of carbon monoxide dehydrogenase (CODH)/acetyl-CoA synthase reduced the net secretion of carbon dioxide. The actual ratio in *M. barkeri* has been measured as 3.4:1 or 4.4:1 (8).

*M. acetivorans* produces acetate, formate, methane, and methyl sulfides as by-products when grown on carbon monoxide (in addition to CO₂) (48, 55), but the mechanisms of formate and methyl sulfide formation are unclear. Therefore, to make predictions about the conditions necessary to produce these by-products, it was necessary to hypothesize mechanisms for their production.

It is not currently known how or why *M. acetivorans* generates formate during growth on CO, although it is probably not coupled to methanogenesis (48). It is possible that formate is produced as a by-product of carbon monoxide dehydrogenase during growth on CO to prevent toxic CO accumulation in the cell (48, 56). The CO dehydrogenase enzyme from *Rhodospirillum rubrum* has been shown to create formate as a by-product, and formate may be formed by a similar mechanism in *M. acetivorans* (27), although the physiological substrate for the reaction is still unknown (56). In order to investigate formate production, the following reaction was tentatively included in the model: CODH₃_SIDERXN: co(g)/H₁₁₀₀₁ₓh₂o(l) ¡ for(aq)/H₁¹₀₀₁ₓh(aq), where co is carbon monoxide, h₂o is water, for is formate, h is H⁺/H₂o, (g) indicates the gaseous phase, and (l) indicates the liquid phase. This reaction implies that formate production from CO does not yield ATP, which is likely to be true, since this reaction is endergonic (ΔG° = +24 kJ/mol).

### Table 1 Knockout lethality predictions from FBA and agreement with experimental results

<table>
<thead>
<tr>
<th>Genotype</th>
<th>AC</th>
<th>CO</th>
<th>DMA</th>
<th>DMS</th>
<th>MeOH</th>
<th>MMA</th>
<th>TMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δackθ</td>
<td>L</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔtapDCXREFAG</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔcosS1</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔcosS2</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔhkvABC</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔhdrED</td>
<td>L</td>
<td>N</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Δsrmch</td>
<td>N</td>
<td>N</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>ΔmtaAI</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaB1C1ΔmtaB2C2ΔmtaB3C3</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaAIΔmtaB1C1ΔmtaB2C2ΔmtaB3C3</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaH</td>
<td>N</td>
<td>N</td>
<td>L</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaADmtaFSmtSH</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>L</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaXAl ΔmtaY two mts genes</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaYhHDCGABF</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaXAl ΔmtaY two mts genes</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ΔmtaYhHDCGABF</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

a L, lethal; N, nonlethal. AC, acetate; DMA, dimethylamine; DMS, dimethyl sulfide; MeOH, methanol; MMA, monomethylamine; TMA, trimethylamine. The ack gene encodes acetate kinase; the pta gene, phosphotransacetylase; hkd, heterodisulfide reductase; mch, methyl-tetrahydrofolate reductase; mts, dimethyl sulfide-coenzyme M methyltransferase. rnf, putative ferredoxin-methanophenazine oxidoreductase; mtr, sodium-pumping tetrahydrofolate reductase, coenzyme M methyltransferase. See the supplemental material for knockout data references.
b Green, correct prediction; red, incorrect prediction; no color, no experimental data available for that knockout under those conditions.

### Table 2 Growth and secretion rates and yields of *M. acetivorans* on methanol, acetate, and carbon monoxide using experimentally feasible uptake rates

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Measured substrate uptake rate (mmol/g dry wt/h)</th>
<th>Growth (h⁻¹)</th>
<th>Growth yield (g/dry wt/mol substrate)</th>
<th>CH₄ secretion rate (mmol/g dry wt/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Measured Predicted</td>
<td>Measured Predicted</td>
<td>Measured Predicted</td>
<td>Measured Predicted</td>
</tr>
<tr>
<td>Acetate</td>
<td>7</td>
<td>0.023</td>
<td>0.021</td>
<td>2.4</td>
</tr>
<tr>
<td>Methanol</td>
<td>20</td>
<td>0.098</td>
<td>0.086</td>
<td>5.2</td>
</tr>
<tr>
<td>CO</td>
<td>11.6</td>
<td>0.029</td>
<td>0.030</td>
<td>2.5</td>
</tr>
</tbody>
</table>

a See references 55, 60, and 63. For simulations on CO, no additional constraints to Hdr or Cdh were assigned. All measured values are averaged across literature sources (see the supplemental material for a complete list of references). Note that due to experimental variability, the measured methanol uptake rate and the secretion of methane on that substrate are inconsistent with mass balance constraints.
[or \( \Delta G^\circ = +6 \text{kJ/mol if CO is treated in the aqueous phase} \)] under standard conditions (1).

The recently characterized Mts enzymes are necessary for the production of DMS in *M. acetivorans*, although due to the very low ratio of the dimethyl sulfide production rate to the transcription levels of these enzymes, the *in vivo* function of these enzymes remains unclear (47). The source of methyl sulfide needed as a substrate for Mts to make dimethyl sulfide is unknown. However, due to the similar structures of sulfide \( (\text{HS}^-) \) and methyl sulfide, one could reasonably hypothesize that methyl sulfide is formed by the same Mts enzyme that produces dimethyl sulfide: \( \text{m5hbc(aq)} + \text{h2s(aq)} + \text{h(aq)} \rightarrow \text{ch4s(aq)} + \text{5hbc_red(aq)}. \) Here, \( \text{ch4s} \) is methyl sulfide, \( \text{h2s} \) is \( \text{H}_2\text{S}, \text{m5hbc} \) is the methylated form of a cobalamide cofactor utilized in methyltransferases in *Methanosarcina* (20), and \( \text{5hbc_red} \) \( [5\text{-hydroxybenzimidazolylcob(I)amide}] \) is the unmethylated form of \( \text{m5hbc}. \)

Despite the inclusion of reactions to make the observed by-products acetate, formate, and methyl sulfides, FBA predicted that only methane and CO\(_2\) would be produced as by-products during growth on CO. Consequently, the methane secretion rate was significantly higher than that observed in experiments (Table 2). To investigate the reason for this incorrect prediction, the theoretical ATP yield was calculated for the production of each by-product per mole of CO consumed, as described in Materials and Methods. The theoretical ATP yield from methanogenesis was significantly higher than that from acetogenesis, methyl sulfide production, or formate generation (Fig. 3B). Because most biosynthesis reactions were unconstrained in the direction necessary for biosynthesis, FBA predicted the utilization of pathways with greater ATP production efficiency, because if less substrate is needed to satisfy the ATP requirements of the cell, then more is available to produce biomass. We subsequently examined possible conditions under which these by-products could be produced in an FBA model.

**Formate production and regulation of CO levels in *M. acetivorans*.** *M. acetivorans* encodes at least two complete CODH operons, and their relative expression during growth on CO may depend on the concentration of CO in the medium (56). Since formate production may be a result of a side reaction of CODH (27), it is tempting to speculate that the level of carbon monoxide in the cell is regulated by the balance of the levels of these proteins, one of which produces formate as a by-product and one of which does not. FBA predicts that if the proposed mechanism for formate production is correct, reduction of the flux through the primary reaction catalyzed by CODH leads to formate production (Fig. 3C). This indicates that CO toxicity could be controlled by balancing the rates of the side reaction and the primary reaction of CODH enzymes.

**Inhibition by CO of the methyltransferase Mtr and its possible role in acetogenesis.** The sodium-pumping methyltransferase Mtr catalyzes the reversible transfer of methyl from methyl tetrahydrofolate to methyl coenzyme M (methyl-CoM) via a cobalamide intermediate. This enzyme is strongly downregulated during growth on CO compared to other substrates (56). It has been hypothesized that methanogenesis is kinetically limited during growth on CO (55), possibly including Mtr, although this hypothesis is controversial (48). Flux balance analysis predicted that if Mtr was not inhibited, the cell could still survive without methanogenesis or acetogenesis by producing methyl sulfides.

The theoretical ATP yield of methyl sulfide production, including ATP generation due to the sodium gradient created by Mtr (0.33 ATP/CO), is sufficient to fulfill the ATP maintenance requirement (0.22 ATP/CO) (Fig. 3B). Since acetogenesis is actually essential for growth on carbon monoxide (55) and methyl sulfide production is actually very low, this prediction suggests that either Mtr activity is strongly limited during growth on carbon monoxide or the production of methyl sulfides is kinetically limited.

Despite the downregulation and possible inhibition of Mtr during growth on CO, significant methanogenesis is still observed during growth on this substrate (48, 55). The combination of these observations has inspired the hypothesis that a Mtr bypass reaction exists that performs the same function but does not generate a sodium gradient (21): \( \text{MTR_BYPASS: mh4spt(aq)} + \text{5hbc_red(aq)} + \text{h(aq)} \rightarrow \text{h4spt(aq)} + \text{m5hbc(aq)}, \) where \( \text{mh4spt} \) is methyl tetrahydrofolate. This reaction, when coupled with methyl transfer to coenzyme M, would be strongly thermodynamically favored in the direction of methyl-CoM formation \( (\Delta G^\circ = -30 \text{kJ/mol}) \) (67). The presence of the bypass reaction could permit tolerance for a wider range of environmental CO concentrations by permitting the cell to balance the increased ATP potential of Mtr with a possibly greater kinetic capacity of the Mtr bypass reaction (21).

To test the effects of the Mtr bypass reaction on metabolism, the bypass reaction was added to the metabolic network, and the sodium-pumping methyltransferase Mtr was removed from the model. The modified model was still predicted to perform only methanogenesis, not acetogenesis, because the theoretical ATP yield of methanogenesis was still higher than that of acetogenesis (0.44 ATP/CO and 0.38 ATP/CO, respectively) (Fig. 3B). In addition, flux variability analysis indicated that no alternative optimal solutions led to acetate secretion. As a result, Mtr inhibition is probably not the sole reason for acetogenesis during growth on CO.

**Exploration of an alternate heterodisulfide reductase (HdrABC) on methanol.** HdrABC is a soluble heterodisulfide reductase typically found in methanogens without cytochromes (62). Most methanogens with cytochromes, including *Methanosarcina* species, use a membrane-bound heterodisulfide reductase, HdrABC, instead of the soluble HdrABC to couple methanogenesis to ATP production (67). Surprisingly, *M. acetivorans* was found to utilize both types of heterodisulfide reductase during growth on methylotrophic substrates (10).

Since the HdrABC complex is not a sodium or proton pump, it is not clear whether the activity of this complex is coupled to ATP synthesis in *M. acetivorans*. In *Methanothermobacter marburgensis*, a methanogen without cytochromes, HdrABC is coupled to ATP synthesis through its interaction with the MvhADG hydrogenase complex (62). The HdrABC-MvhADG complex in *M. marburgensis* uses an electron bifurcation mechanism, in which the electrons from two equivalents of molecular hydrogen are donated to ferredoxin and to the heterodisulfide (62). *M. acetivorans* lacks the genes encoding the MvhADG complex, but the similarity of the Hdr enzymes suggests that *M. acetivorans* HdrABC may also use an electron bifurcation mechanism (Fig. 4A), splitting the electrons of two fully reduced ferredoxins between heterodisulfide and coenzyme F\(_{430}\) (10). Alternatively, HdrABC may simply reduce heterodisulfide with ferredoxin, acting as a sink for excess ferredoxin produced during the oxidation of methanol to CO\(_2\).

To test the electron bifurcation hypothesis in a genome-scale
context, the phenotype of *M. acetivorans* was simulated with and without electron bifurcation in HdrABC. The addition of two additional constraints was necessary in order to obtain reasonable predictions. Reactions catalyzed by Pta and Ack were disabled so as to prevent acetate secretion, which has not been observed during growth on methanol (36), and the flux through pyruvate-acetyl-CoA oxidoreductase was set to be equal to the wild-type value so as to prevent the secretion of formate and other unobserved by-products during growth on methanol (60). In the presence of Rnf, flux variability analysis did not predict utilization of HdrABC regardless of the mechanism. However, a \(rnf\) mutant was predicted to utilize HdrABC to oxidize ferredoxin using any optimal flux distribution (Fig. 4B). The \(rnf\) mutant was predicted to grow 35% more slowly than the wild type without bifurcation and 20% more slowly with bifurcation. A \(rnf\) mutant actually grew about 25% more slowly on methanol than the wild type (W. Metcalf, unpublished data), so within experimental error it is difficult to tell which mechanism is correct. Further experiments could help elucidate the true mechanism.

**DISCUSSION**

We have built and manually curated a computable genome-scale model of metabolism in *M. acetivorans*, only the third methanogen species to be reconstructed (after *M. barkeri* [19] and *Methanocaldococcus jannaschii* [69]) and the second in the genus *Methanosarcina*. We have focused on three approaches to model-guided discovery using the *Methanosarcina acetivorans* model: (i) identification of knowledge gaps and “missing” reactions in metabolic pathways, (ii) detailed study of metabolic differences between closely related methanogenic species, and (iii) use of constraint-based modeling to study alternative hypotheses about the workings of the metabolic network and the implications of those hypotheses for predicted phenotypes.

The reconstruction endeavor has helped pinpoint gaps in our knowledge of the metabolic networks, both due to unknown differences between different archaeal species and due to differences between the *Archaea* and other domains of life. Interestingly, even some pathways for the synthesis of specialized methanogenic cofactors, such as those for tetrahydrosarcinopterin and coenzyme M, seem to have diverged from those observed in other methanogens, such as *Methanocaldococcus jannaschii*. However, the synthesis pathways for other methanogenic cofactors (such as coenzyme B) are conserved across these genera. This observation raises interesting questions about the evolution of these ecologically important organisms and the role of the environment in the selection of metabolic pathways.

Constraint-based modeling proved useful for integrating experimental data from different sources and identifying tensions between data sets. Some of these tensions, such as the disparity between the ability of *M. acetivorans* to grow on CO and the lethality of a *mch* knockout, may have been difficult to identify without a genome-scale model and its accompanying predictions. These findings highlight the usefulness of an integrative, genome-scale modeling approach both for validating model predictions and for identifying gaps in our knowledge of methanogen biology.

One of the strengths of genome-scale metabolic modeling is the ability to continually update the model as additional experimental data become available (18). Our investigations of alternative hypotheses for the mechanism of F420 regeneration during growth on carbon monoxide, pathways for synthesis of by-products observed during growth on CO, and the precise reaction catalyzed by the soluble heterodisulfide reductase HdrABC have yielded predictions that can be tested in the laboratory. As additional data become available, improved models may be constructed and used to provide further novel hypotheses in an iterative process that lies at the heart of systems biology.

**FIG 4** Studies on the soluble heterodisulfide reductase HdrABC during growth on methanol. (A) Hypothesized bifurcation mechanism of the soluble heterodisulfide reductase HdrABC during growth on methylotrophic substrates in *M. acetivorans*. The alternative hypothesis is that the reaction does not involve F420 (dashed lines). (B) Flux balance analysis predicts that during growth on methylotrophic substrates (methanol shown), HdrABC is not used when Rnf is available, but a \(\Delta rnf\) mutant is predicted to carry flux through HdrABC. The growth rate for the \(\Delta rnf\) mutant is predicted to be about 20% less than that for the wild type with bifurcation and 35% less without.
ACKNOWLEDGMENT
We gratefully acknowledge support from the Genomic Sciences Program within the Department of Energy, Office of Biological and Environmental Research (award DE-FG02-10ER64999).

REFERENCES

Benedict et al.


71. Welander PV, Metcalf WW. 2005. Loss of the nhaA gene in Methano-

72. Welte C, Krätzer C, Deppenmeier U. 2010. Involvement of Ech hydro-
