Molecular BioSystems

PAPER

View Article Online
View Journal | View Issue

Cite this: *Mol. BioSyst.*, 2014, **10**, 1043

A genome-scale metabolic model of *Methanococcus maripaludis* S2 for CO_2 capture and conversion to methane†

Nishu Goyal,[a] Hanifah Widiastuti,[a] I. A. Karimi*[a] and Zhi Zhou[b]

Received 20th September 2013,
Accepted 3rd February 2014

DOI: 10.1039/c3mb70421a

www.rsc.org/molecularbiosystems

Methane is a major energy source for heating and electricity. Its production by methanogenic bacteria is widely known in nature. *M. maripaludis* S2 is a fully sequenced hydrogenotrophic methanogen and an excellent laboratory strain with robust genetic tools. However, a quantitative systems biology model to complement these tools is absent in the literature. To understand and enhance its methanogenesis from CO_2, this work presents the first constraint-based genome-scale metabolic model (*i*MM518). It comprises 570 reactions, 556 distinct metabolites, and 518 genes along with gene-protein-reaction (GPR) associations, and covers 30% of open reading frames (ORFs). The model was validated using biomass growth data and experimental phenotypic studies from the literature. Its comparison with the *in silico* models of *Methanosarcina barkeri*, *Methanosarcina acetivorans*, and *Sulfolobus solfataricus P2* shows *M. maripaludis* S2 to be a better organism for producing methane. Using the model, genes essential for growth were identified, and the efficacies of alternative carbon, hydrogen and nitrogen sources were studied. The model can predict the effects of reengineering *M. maripaludis* S2 to guide or expedite experimental efforts.

Introduction

Increasing atmospheric carbon dioxide (CO_2) levels, leading to global warming,[1] is a serious concern facing the world today. Power, transport, and chemical sectors are the three major sources of CO_2 emissions. These emissions can be minimized by either storing or converting the emitted CO_2. At present, carbon capture and storage (CCS) technology[2] demands compression to 120–150 bar, and it is costly. In addition, it poses uncertain risks such as burping and ocean acidification. Thus, considerable research is in progress to convert CO_2 to valuable products and fuels *via* metal catalysts,[3] photocatalysts,[4] electrocatalytic reduction,[5] nanotechnology,[6] and biocatalysts.[7]

Biochemical conversion of CO_2 to methane *via* methanogens is a novel biocatalytic approach for carbon capture and utilization. Methanogens are archaebacteria that can produce methane as a catabolic product using three different pathways: CO_2-reducing, methylotrophic, and aceticlastic.[8] Hydrogenotrophic methanogens reduce CO_2 to methane in the presence of H_2/formate, and create an electrochemical gradient across the cell membrane to produce ATP. Ruminants (*e.g.* cattle, sheep, and goats) produce 86 million tonnes of methane per year due to the presence of hydrogenotrophic methanogens in the rumen.[9]

Several methanogens such as *Methanosarcina barkeri*,[10] *Methanococcus jannaschii*,[11] *Methanococcus voltae*,[12] and *Methanobacterium thermoautotrophicum*[13] have been studied in the literature. *M. maripaludis* S2 (mmp), also known as *M. maripaludis LL*, is a hydrogenotrophic, mesophilic, gram-negative, anaerobic archaebacterium that can utilize CO_2 as the sole carbon source.[14] Its ability[14,15] to use both CO_2 and N_2 makes it ideal for treating industrial flue gases comprising primarily of N_2 (70–80%) and CO_2 (8–14%). It is one of the fastest growing mesophilic methanogens with a doubling time of 2 h at 35–39 °C, and is an excellent laboratory strain for molecular and biochemical studies.[16–19] The presence of selectable resistance markers,[20] efficient transformation methods,[21] gene deletion or substitution strategies,[18,22] and integrative and shuttle expression vectors[23] makes it possible to manipulate the genome of *M. maripaludis* to enhance its methanogenesis. For instance, acetate auxotrophs were isolated using random insertional mutagenesis by transforming the wild type *M. maripaludis* with the pWDK104 vector.[24] Four mutations were made in and around the *nifH* gene using

[a] *Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576. E-mail: cheiak@nus.edu.sg*
[b] *Department of Civil and Environmental Engineering, National University of Singapore, 1 Engineering Drive 2, Singapore 117576*

† Electronic supplementary information (ESI) available: Supplementary file 1: Microsoft Excel® file listing all the reactions along with GPR associations, metabolites, exchange reactions, minimal media compositions, dead end metabolites, and essential genes/reactions. Supplementary file 2: Microsoft word file (.docx) containing information about major cellular components required for producing the biomass of *M. maripaludis* S2. See DOI: 10.1039/c3mb70421a

transposon insertion mutagenesis for studying nitrogen fixation.[25] However, most such studies have been experimental.

In silico genome-scale metabolic modeling[26,27] and analysis is a systematic, proven, and attractive approach for exploring metabolic manipulations in an organism to enhance its desirable traits. It helps to develop insights into active pathways, genes, proteins, and metabolites and their interactions in a microbe. The first genome-scale metabolic model was developed in 1995 for *Haemophilus influenzae.*[28] Since then, many reconstructed *in silico* models for bacteria (*Escherichia coli*,[29] *Rhodococcus erythropolis*,[30] *Zymomonas mobilis*[31]) archaea (*Methanosarcina barkeri*,[10] *Methanosarcina Acetivorans*[32]), and eukaryota (*Mus musculus*[33]) have been reported in the literature. The predictions from such models add a level of assurance and guidance for experimentation with organisms.

Only three genome-scale metabolic models are available to date for Archaea. These are *Methanosarcina barkeri*,[10] *Methanosarcina acetivorans*,[32] and *Sulfolobus solfataricus P2*.[34] While the genome of *M. maripaludis* S2 is fully sequenced with a size of 1.66 Mb and 1722 protein coding genes,[16] no genome-scale metabolic model exists. A reconstructed model would greatly aid experimental work by helping to predict various cellular phenotypes including growth and quantify the impact of complex intracellular interactions.

This study is the first (to our knowledge) to report a reconstructed genome-scale metabolic model for *M. maripaludis* S2 based on the available genomic, biochemical, and physiological data. It identifies several new genome annotations for proteins including those designated as ''unique proteins of unknown function'' by Hendrickson *et al.*[16] The model successfully predicts the specific growth rate (μ), the Methane Evolution Rate (MER), and substrate uptake rates. It presents a factual and quantitative framework for analyzing, simulating, and enhancing the metabolic phenotypes of *M. maripaludis* S2, which may eventually help in the capture and conversion of CO_2 to methane.

Results and discussion

Reconstructed metabolic network

The genome of *M. maripaludis* S2 is a circular chromosomal DNA with 1722 protein-coding genes (ORFs or Open Reading Frames). Our metabolic model comprises 605 metabolites (556 internal and 49 external) and 570 reactions (521 intracellular and 49 transport) across 52 distinct pathways (ESI,† file 1). It accounts for 518 ORFs out of the 1722 presently known for protein-coding genes (approx. 30% coverage). 498 ORFs out of the 518 have annotation in different databases while 20 of them were assigned functions by us based on sequence similarity searches. Basic features of the *M. maripaludis* S2 genome and the *in silico* reconstructed model are summarized in Table 1. The reactions cover central metabolism, energy metabolism, amino acid metabolism, nucleotide metabolism, vitamin and cofactor metabolism, and production of other secondary metabolites (Table 2). The average confidence score for the 570 reactions is 4.67, which indicates that the majority of the reactions added in the model are built on strong evidence.

Table 1 Basic features of the *M. maripaludis* S2 genome and its *in silico* reconstructed model

Feature	Number
Genome features	
Genome size (bp)	1 661 137
% GC content	33.1
Open Reading Frames (ORFs) (coverage %)	1722 (88.9%)
In silico reconstruction	
Total reactions	570
Protein coding genes (ORFs)	518
% ORF coverage	30
Reaction associated with genes	464
Reaction not associated with genes	106
Intracellular/extracellular metabolites	556/49
Transport reactions	49

The transport reactions were based on the available genome annotations and known physiological information for various metabolites such as formate, acetate, cobalt, molybdenum, iron, bicarbonates, and sulfates.[16,35–37] Acetate was included in this list based on the work by Shieh *et al.*,[38] who demonstrated a pathway for acetate assimilation in *M. maripaludis*.

Our initial network (draft model) included 397 unique reactions from KEGG.[36] It lacked reactions essential for the synthesis of several biomass precursors (*e.g.* proline, methionine, glycine, histidine, and cysteine), cofactors (coenzyme B, coenzyme M, flavin adenine dinucleotide, tetrahydromethanopterin (THMPT)), and vitamins (folic acid, riboflavin, cobalamin) required for enzymatic activities and growth. Therefore, 124 more reactions were added from METACYC[37] and the published literature[10,38–47] to fill these gaps for achieving *in silico* growth (ESI,† file 1). This gap-filling process was based on a thorough biological understanding of the various intracellular functions in methanogenic species. For instance, the genome annotation studies by Hendrickson *et al.*[16] report that only one ORF (*metE*, MMP0401) is present in *M. maripaludis* for methionine biosynthesis, but the other literature[46] mentions that cystathionine β-lyase (*metC*, 4.4.1.8) also plays a key role. Our BLASTp similarity search (*e*-value = 10^{-17}) showed that an ORF (MMP1072) in *M. maripaludis* S2 matches the sequence of cystathionine β-lyase present in *Methanobacterium* sp. AL-21. Thus, the reactions corresponding to MMP1072 were also added to complete the synthesis of methionine. Similarly, the synthesis of proline from ornithine was completed by adding the pathway reactions from a closely related methanogen.[40] No ortholog is known to code for histidinol-phosphatase (*hisJ*, 3.1.3.15) required for histidine synthesis. However, the reactions catalyzed by histidinol-phosphatase were also added to the model, because all other genes (*hisA*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*) involved in the pathway are already reported[36,37] and hence the presence of *hisJ* locus is also expected. All such reactions based on assumptions were assigned low confidence scores in the model.

M. maripaludis S2 (mmp) uses the Wood–Ljungdahl pathway for synthesizing acetyl coenzyme A (acetyl-CoA) from two molecules of CO_2.[24] As shown in Fig. 1, the methyl carbon of acetyl-CoA comes from methylTHMPT, an intermediate produced by

Table 2 Functional classification of the metabolic reactions in the *M. maripaludis* S2 genome-scale model

Central metabolism	*40*	*Nucleotide metabolism*	*82*
Glycolysis/gluconeogenesis	24	Purine	45
Citrate cycle (TCA cycle)	5	Pyrimidine	37
Pentose phosphate pathway	11	*Metabolism of cofactors and vitamins*	*127*
Energy metabolism	*38*	*Others*	*200*
Methane metabolism	28	Selenocompound and glutathione metabolism	9
Nitrogen metabolism	6	Metabolism of terpenoids and polyketides	9
Oxidative phosphorylation	4	Xenobiotics biodegradation and metabolism	6
Amino acid metabolism	*131*	Pyruvate metabolism	6
Valine, leucine and isoleucine	26	Other carbohydrates metabolism	34
Alanine, aspartate and glutamate	11	Lipid metabolism	7
Glycine, serine and threonine	11	tRNA charging	26
Cysteine and methionine	17	Transport reactions	48
Arginine and proline	14	Others (biomass formulation)	7
Phenylalanine, tyrosine and tryptophan	28	*Total*	*570*
Histidine	10		
Lysine	12		
Asparagine and glutamine	2		

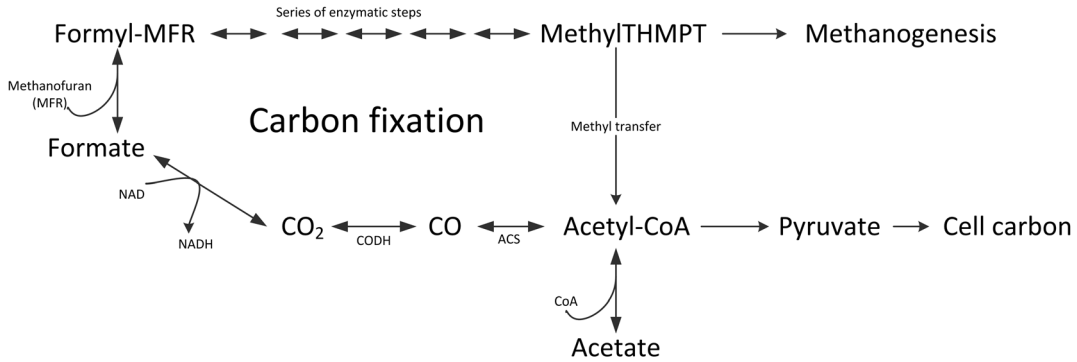

Fig. 1 Wood–Ljungdahl pathway for CO_2 fixation occurring in *M. maripaludis*. One CO_2 molecule is reduced to CO and the other CO_2 molecule is reduced to a methyl group bound to carrier methylTHMPT; subsequent methyl transfer to CO in the presence of CoA leads to acetyl-CoA synthesis. Formyl-MFR, formyl-methanofuran; methylTHMPT, methyltetrahydromethanopterin, acetyl-CoA, acetyl-coenzyme A; CODH, carbon monoxide dehydrogenase; ACS, acetyl-CoA synthase.

formaldehyde and THMPT. The carboxyl carbon of acetyl-CoA comes from the reduction of CO_2 to carbon monoxide (CO) *via* carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS), a key enzyme in methanogenesis. CODH catalyzes the reduction of CO_2 to CO and ACS helps in the condensation of CO, the methyl group, and coenzyme A to acetyl-CoA. Based on the work by Shieh *et al.*,[38] acetyl-CoA can also be synthesized from acetate using acetyl-CoA synthetase, which then participates in the synthesis of biomass precursors for growth and methanogenesis.

Fig. 2 shows the major metabolic pathways in our model along with their GPR associations. They include glycolysis, citrate cycle, amino acid biosynthesis, pentose phosphate pathway, and methane metabolism. The map also contains useful information about the roles of various pathways in substrate utilization and product formation. For instance, it confirms that CO_2 can indeed be the sole carbon source with H_2 as the reducing agent. In addition to H_2, formate may also act as an electron donor because of the following formate-dependent lyase activity common in *methanococci*.[48]

$$HCO_2^- + H_2O \rightarrow HCO_3^- + H_2 \quad \Delta G = +1.3\ \text{kJ} \qquad (1)$$

In addition to eqn (1), two potential pathways for formate-dependent hydrogen production in *M. maripaludis* S2 reported by Lupa *et al.*[49] were also included. Both pathways include the formation of reduced coenzyme F_{420} ($F_{420}H_2$). However, no biochemical or genomic evidence exists for flavin oxidoreductase, the enzyme responsible for formation of $F_{420}H_2$. This, in our judgment, can be explained by the presence of two formate dehydrogenases in *M. maripaludis* S2, namely *fdhA* and *fdhB*. In the process of oxidizing formate to bicarbonates by these hydrogenses, one electron is donated to NAD or coenzyme F_{420} to generate NADH or $F_{420}H_2$.

An alternative way of supplying hydrogen to methanogens is *via* water splitting, which can generate protons and electrons for the reduction of CO_2 to methane as follows.

$$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O$$

Cheng *et al.*[50] reported the above in *Methanobacterium palustre*, but no evidence exists so far to suggest that *M. maripaludis* can use water as a H_2 source. We are currently doing experiments to observe if *M. maripaludis* can perform electrosynthesis. Meanwhile, we did not include this reaction in our metabolic model. Once some evidence exists, the above reaction can be added to our model.

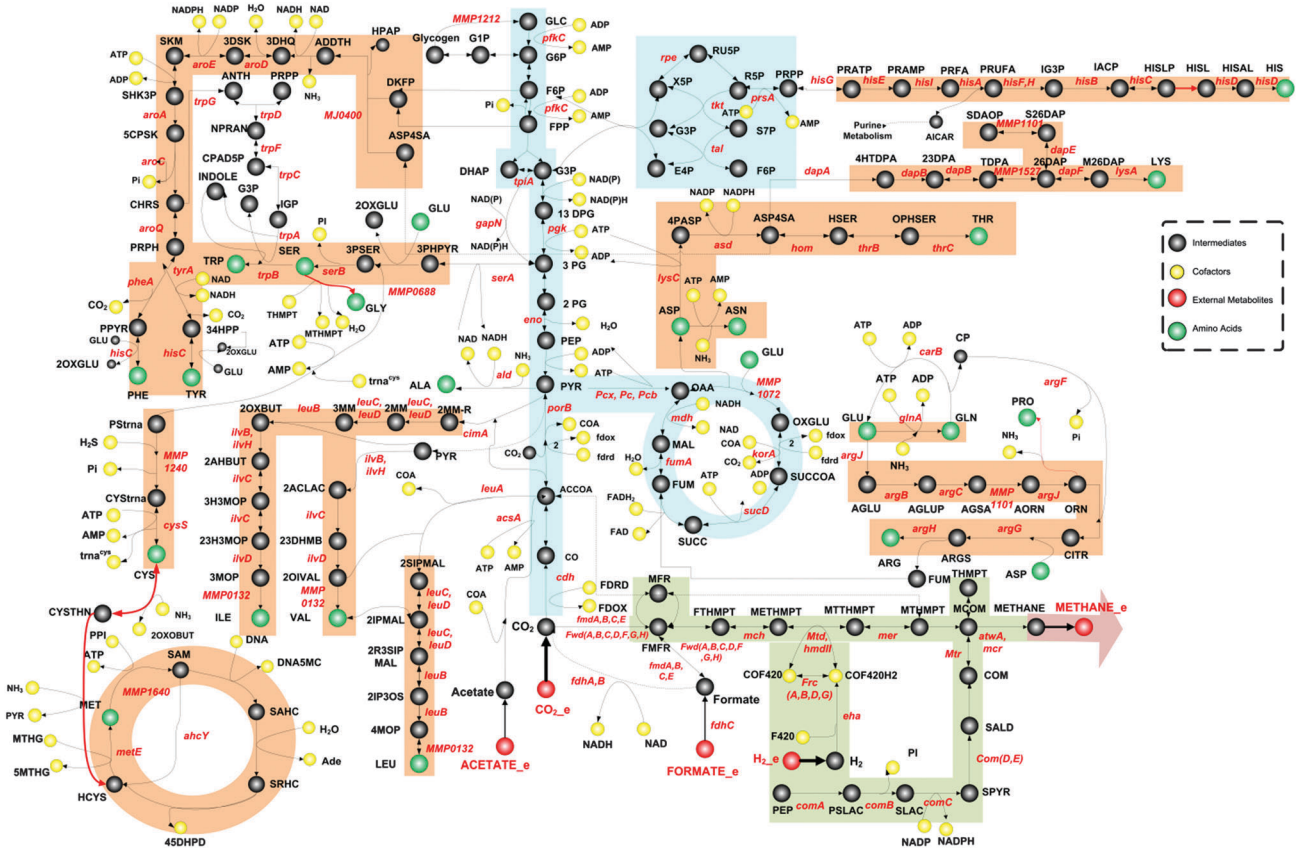

Fig. 2 Central metabolism, amino acid metabolism, and methane metabolism shown in different shades along with GPR associations for *M. maripaludis* S2. The figure shows how CO_2 can act as the sole carbon substrate for growth and energy production (see ESI† for abbreviations and reaction details).

Table 3 ORFs with suggested annotations for *M. maripaludis* S2

Open reading frame (ORF)	Suggested annotation	*E*-value (cut-off 10^{-5})	NCBI accession number
MMP0281	ATPase	0	NP_987401
MMP0413	Methyl-accepting chemotaxis sensory transducer	0	NP_987533.1
MMP0487	Methyl-accepting chemotaxis sensory transducer	0	NP_987607.1
MMP0788	Methyl-accepting chemotaxis protein (MCP) signalling domain	0	NP_987908.1
MMP1716	5,10-MethenylTHMPT hydrogenase (hmdII)	$0.00 \times 10^{+00}$	NP_988836.1
MMP1493	S-layer protein	0	NP_988613.1
MMP1122	Translation-associated GTPase	2.00×10^{-116}	NP_988242.1
MMP0057	FO synthase subunit 2	5.00×10^{-92}	NP_987177.1
MMP0814	APHP domain-containing protein	8×10^{-74}	NP_987934.1
MMP0883	HEAT domain-containing protein	3×10^{-64}	NP_988003.1
MMP1653	HEAT domain containing protein	1×10^{-46}	NP_988773.1
MMP1302	Selenium binding protein	2×10^{-40}	NP_988422.1
MMP1428	Preprotein translocase subunit SecG	3×10^{-24}	NP_988548.1
MMP1300	Nucleic acid binding, OB-fold, tRNA/helicase-type	1×10^{-15}	NP_988420.1
MMP0978	Predicted RNA-binding protein	1.6×10^{-14}	NP_988098.1
MMP1077	Phosphoglucomutase/phosphomannomutase	2.00×10^{-12}	NP_988197.1
MMP1072	Aminotransferase	1.00×10^{-17}	NP_988192.1
MMP0542	Phosphoserine phosphatase	3.00×10^{-9}	NP_987662.1
MMP1358	Pyruvate formate-lyase	6.00×10^{-07}	NP_988478.1
MMP1237	Acetolactate decarboxylase	9.00×10^{-05}	NP_988357.1
MMP0339	NUDIX hydrolase	5.00×10^{-06}	NP_987459.1

In total, 20 new ORFs with suggested annotations in Table 3 were identified in *M. maripaludis* S2 using BLASTp similarity searches (*e*-value cutoff 10^{-5}). These include the loci that have been designated as "unique protein with unknown function" by Hendrickson *et al.*[16] to allow for the possibility of unknown pathway links. In spite of our best efforts in exhausting the known information and updating the model continuously, the model still has 163 dead-end metabolites. However, it has the highest ORF coverage of the existing models for other organisms.

Model validation

Experimental conditions of Timothy *et al.*[51] and Lupa *et al.*[49] were simulated using our model as follows. Unless dictated otherwise by the experimental conditions, fluxes for all metabolites were unrestricted. Growth Associated Maintenance (GAM) was set as 30 mmol ATP g_{DCW}^{-1} h^{-1}, where DCW stands for dry cell weight. For simulating anaerobic experiments, O_2 uptake flux was set as zero. Using the hydrogen uptake rates of *M. maripaludis* reported by Timothy *et al.*[51] as lower bounds in our model, specific growth rates (h^{-1}) and MERs with and without organics were predicted. For each of these two predictions, a growth yield (gDCW $mol_{CH_4}^{-1}$) was computed. These were then compared with average growth yields computed from the data of Lupa *et al.*[49] for the two scenarios. As shown in Fig. 3, the predictions from our model match closely with the experimentally observed growth yields.

Lupa *et al.*[49] measured MERs at various times during the growth of *M. maripaludis* S2. These rates were used as upper bounds in our model along with an assumed Non-Growth Associated Maintenance (NGAM) value to estimate specific growth rates. The NGAM value (0.4 mmol ATP g_{DCW}^{-1} h^{-1}) that gave the closest agreement for the growth values was fixed in the model. The resulting match between the experimental and predicted specific growth rates is shown in Fig. 4. Experimental values of NGAM vary from 0.2 to 7 mmol g_{DCW}^{-1} h^{-1} [10,52,53] for a microbial cell, thus our fitted estimate of 0.4 mmol ATP g_{DCW}^{-1} h^{-1} is quite acceptable. As seen in Fig. 5, NGAM has a significant effect on growth and methanogenesis, thus fitting the best value was necessary.

The predicted yields of methane from CO_2 are 83–85% mol mol^{-1} from our model for the MERs reported by Lupa *et al.*[49] While Stolyar *et al.*[54] reported a yield of 95% based on one single experimental observation, the details regarding their experiment are insufficient for us to simulate the same in our model.

For further verification of our model, gene/reaction knockout experiments performed by Haydock *et al.*,[55] Lin *et al.*,[56] and Lie *et al.*[57] were mimicked. Haydock *et al.*[55] created a leucine auxotroph (S52) by deleting the MMP1063 (*leuA*) gene that encodes 2-isopropylmalate synthase (IPMS), the first enzyme in the biosynthesis of leucine from 2-ketoisovalerate. To emulate S52, the reactions associated with IPMS were inactivated in our model. No growth was predicted by the model for S52. However, when leucine was supplied directly to the medium by adding a transport reaction, full growth was predicted. This validates our model concerning the essentiality of the *leuA* gene in leucine production.

Lin *et al.*[56] reported that pyruvate serves as an electron donor in the absence of H_2 and formate, and the deletions of *porE and porF* affect growth and oxidation of pyruvate. To show this, hydrogen and formate uptake was set to zero and pyruvate (10 mmol g_{DCW}^{-1} h^{-1}) was supplied in the medium. Our model could successfully predict the growth rate (0.09321 h^{-1}) and MER (10.166 mmol g_{DCW}^{-1} h^{-1}). Furthermore, the reaction catalyzed by pyruvate synthase (EC number: 1.2.7.1), also known as pyruvate:ferredoxin oxidoreductase (PFOR), associated with

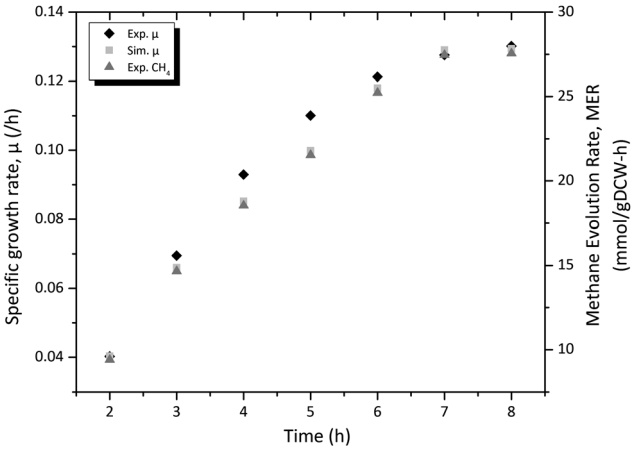

Fig. 4 Comparison of the specific growth rates observed *in silico* and *in vivo* at different time points during the growth of *M. maripaludis* S2 (experimental data from Lupa *et al.*[49]).

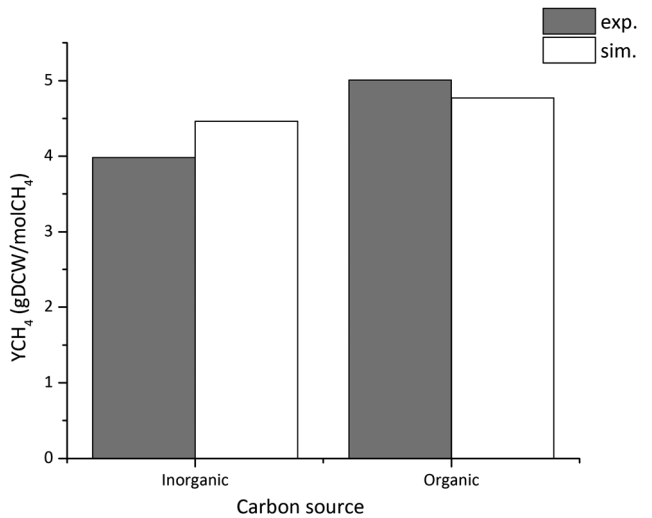

Fig. 3 Comparison of experimentally observed and model predicted growth yields of *M. maripaludis* under hydrogen limited conditions using data from Timothy *et al.*[51]

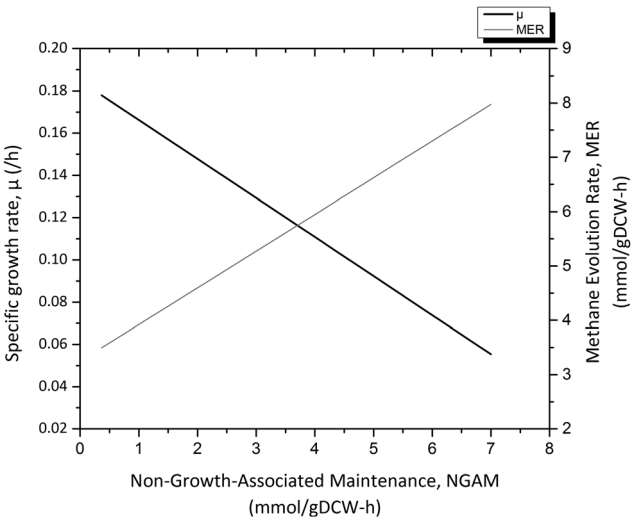

Fig. 5 Influence of NGAM on specific growth rates and the MER.

porE and porF was inactivated in our model. Significant reduction in the growth (0.0188 h^{-1}) and MER (9.77 mmol g_{DCW}^{-1} h^{-1}) was predicted, which validates our model concerning the essential roles of *porE* and *porF* in maintaining wild-type growth and methanogenesis.

Lie *et al.*[57] mentioned that *nif* and *glnA* expressions depend on nitrogen availability. To simulate these diazotrophic conditions, nitrogen was allowed in our model *via* a transport reaction. Our model correctly predicted the expressions of *nif* and *glnA* and demonstrated the ability of *M. maripaludis* to fix nitrogen, which is unique to hydrogenotrophic methanogens. In the absence of nitrogen and the presence of ammonia (non-diazotrophic condition), *nif* did not express, but *glnA* expression was predicted to be high, which is also consistent with the observations of Lie *et al.*[57] When both ammonia and free nitrogen were supplied, our model showed zero nitrogen uptake. Thus, ammonia seems to be the preferred nitrogen source. This is likely due to the fact that a significant amount of energy is required for nitrogen fixation as compared to ammonia. This is consistent with the lower growth rate predicted by our model for free nitrogen *versus* ammonia, as discussed later.

Lastly, our model showed a nonzero flux of "hydroxypyruvaldehyde phosphate" during growth, which is the side product of the 6-deoxy-5-ketofructose-1-phosphate (DKFP) pathway[21] and responsible for the degradation of *methanococcus* cells in the presence of oxygen due to the production of hydrogen peroxide (H_2O_2).

Minimal media

Essential media components were identified by predicting the cell growth for zero uptake of a given extracellular metabolite. Sources of carbon, hydrogen, nitrogen, sulfur, iron salts, sodium ions, cobalt, nickel, and phosphate were found to be essential for growth (see ESI,† file 1 for full details). None of the vitamins or amino acids were essential, thus *M. maripaludis* has the ability to synthesize them. However, the presence of some vitamins and/or amino acids enhanced growth in varying extents. While pyruvate, formate, and alanine can be alternative primary carbon sources, acetate, leucine, and isoleucine can only stimulate growth. Among CO_2, pyruvate, and formate, pyruvate seems to be the most efficient carbon source for growth. Free nitrogen, ammonia, and alanine are alternative nitrogen sources. Cysteine can be an alternative sulfur source, as it reduces H_2S intake significantly.

Gene essentiality and flux variability analyses

The results of these analyses are represented as a gene or locus essentiality matrix indicating the essential proteins or enzymes in the ESI,† file 1. Of the 518 genes in our model, 278 proved essential for cell growth, and 241 non-essential. 282 of the 570 reactions proved essential for growth.

FVA showed 476 of 570 reactions (86%) to have zero variability for maximum growth. Alternative solutions with nonzero fluxes are possible for 67 of the 356 zero fluxes. Most of these belong to the citrate cycle and energy metabolism. However, only 9 of the 214 nonzero reaction fluxes can be zero in alternative solutions. This gives an indication of the robustness of our flux predictions.

Formate as an alternative carbon and hydrogen substrate

Literature[58,59] on *M. maripaludis* S2 mentions that this methanogen is exceptionally equipped with the enzymes for H_2 metabolism and contains genes (*Eha*, *Ehb*, *Fru*, *Frc*, *Vhu*, *Vhc*, and *Hmd*) for seven different hydrogenases. Hydrogenases catalyze the reversible reaction $H_2 \leftrightarrow 2H^+ + 2e^-$ to generate electrons for redox reactions. In another study,[19] two sets of formate dehydrogenases were observed in *M. maripaludis* S2 encoded by *fdhA1B1* and *fdhA2B2*. They claimed that *M. maripaludis* S2 growth was superior on CO_2:H_2 than formate. Our model could successfully predict higher growth in the presence of CO_2 than formate irrespective of the nitrogen source used (Fig. 6).

Hydrogen-independent growth of hydrogenotrophic methanogen *M. maripaludis* S2 in the presence of formate has been demonstrated by Costa *et al.*[60] However, our model could not predict the growth in absence of hydrogen. This might be due to the existing gaps in the model or the unbalanced hydrogen

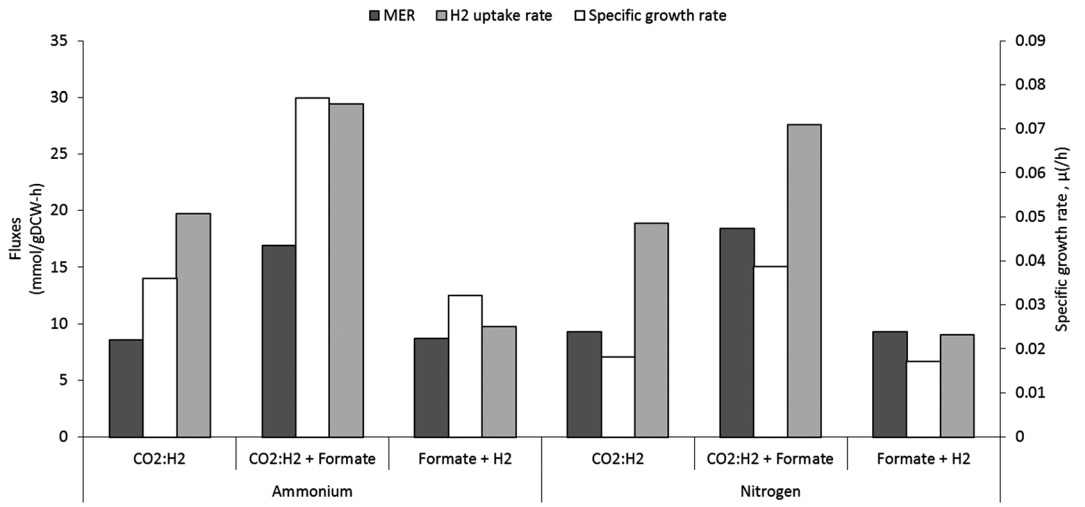

Fig. 6 Influence of different nitrogen sources on MERs, hydrogen uptake rates, and specific growth rates at a fixed carbon uptake of 10 mmol g_{DCW}^{-1} h^{-1}.

in our model. However, significant reduction in the hydrogen uptake rate was observed in the presence of formate (Fig. 6), which clearly shows the presence of formate-hydrogen lyase formate activity and formate-dependent H_2 production in *M. mariapludis* S2 as demonstrated by Lupa *et al.*[49]

Effect of nitrogen sources

Nitrogen is essential for the biosynthesis of amino acids, purines, pyrimidines, and polyamines. Nitrogen, ammonium, and alanine are known alternative sole nitrogen sources[61,62] for *M. maripaludis*. The presence of the *nif* cluster was demonstrated[63] earlier for nitrogen fixation. Ammonium transporter proteins (*Amt*) encoded by MMP0065 and MMP0068 are responsible for the uptake of NH_3 or NH_4^+ from extracellular medium. Pathways also exist for ATP-dependent glutamine synthesis from NH_4^+ and glutamate *via* glutamine synthetase. The presence of the sodium alanine symporter ensures the uptake of alanine in *M. maripaludis.*

The model could assimilate alanine as a sole nitrogen source for growth and methanogenesis. The presence of alanine dehydrogenase encoded by locus MMP1513, uniquely present in archaea, explains the ability of *M. maripaludis* to convert alanine to ammonia and pyruvate. Ammonia then acts as a nitrogen source and pyruvate as an intermediate carbon source. Thus, alanine can act as both the carbon and nitrogen source. Our model predicts that the higher the supply of alanine, the higher the growth and MER. However, alanine is not a preferred nitrogen source, as it reduces CO_2 uptake significantly.

The specific growth rates in the presence of nitrogen or ammonium as predicted from our model are shown in Fig. 6. Ammonium seems to be significantly better than free nitrogen for growth. This could be due to the additional ATP required by nitrogenases for fixing nitrogen to form ammonia. This prediction also agrees with the experimental observations of Belay *et al.*[64] and Fardeau *et al.*[65] for other archaebacteria. Interestingly, methanogenesis is higher in the presence of nitrogen than ammonium as shown in Fig. 6. The possible reason could be changes in flux distribution due to the activation of *nif* genes, which requires the presence of an electron donor (reduced ferredoxin or fd(rd)) for reducing free nitrogen to ammonia. In addition, lower flux was observed in our model towards the CO_2 to acetyl-CoA pathway indicating that reduced supply of the electron donor affects growth significantly.

In the presence of ammonium, formate reduces hydrogen uptake significantly from 19.720 mmol g_{DCW}^{-1} h^{-1} to 9.734 mmol g_{DCW}^{-1} h^{-1} compared to CO_2:H_2. Replacing ammonium with free nitrogen further decreases the uptake to 8.99 mmol g_{DCW}^{-1} h^{-1} as shown in Fig. 6. A possible explanation for this is the evolution of molecular hydrogen during nitrogen fixation:

$$8\text{fd(rd)} + 8H^+ + N_2 + 16\text{ATP} + 16H_2O \rightarrow 8\text{fd(ox)} + H_2 + 2NH_3 + 16\text{ADP} + 16\text{pi} \quad (2)$$

where fd(rd) and fd(ox) indicate reduced and oxidized ferredoxin, respectively, and pi is orthophosphate generated during ATP hydrolysis.

Novel strains for enhanced methanogenesis

Several approaches can be explored for enhancing methanogenesis: best media, enzyme activity regulation, and gene deletions or additions. As seen earlier, free nitrogen is the best source for enhancing the MER. Thus, for studying MER enhancements, the carbon uptake was fixed at 10 mmol g_{DCW}^{-1} h^{-1} in our model with unlimited supply of free nitrogen.

Fig. 7 shows the relation between the MER and growth, which was obtained for the wild type strain by maximizing the MER for varying demands on biomass growth. Clearly, the MER and growth compete with each other for the available carbon, and the MER can only be increased at the expense of growth. With these limits in mind, we now study how the cell metabolism could be manipulated to enhance the MER from its base value for the wild type strain.

Enzymes control reaction fluxes, and their activities can be altered *via* inhibitors, activators, coenzymes, or cofactors. Reaction fluxes were varied in our model to study the effects of enzyme activities on the MER. Table 4 shows the top 10 reactions with the most impact on the MER. All of them are essential reactions and most of them belong to methane metabolism. Two criteria can be used to judge their effectiveness for enhancing the MER. One is the ratio of ΔMER to ΔFlux, and the other is the ratio of ΔMER to ΔGrowth. The former weighs the benefits of the MER against efforts required to change a reaction flux, while the latter weighs the benefits of the MER against the loss of growth. Using the latter as the more desirable metric, R72 catalyzed by nitrogenase seems to be the best reaction for enhancing the MER, with R53 catalyzed by 5,10-methyleneTHMPT reductase as the second best. When the flux of R72 increases, the cell uptakes more free nitrogen. It then requires more energy for growth. Since methanogenesis is the energy-producing pathway in this organism, the MER increases to supply the additional energy demand for growth. R53 causes reduction of 5,10-methyleneTHMPT to 5-methylTHMPT with the help of $F_{420}H_2$. Increased flux through R53 requires more

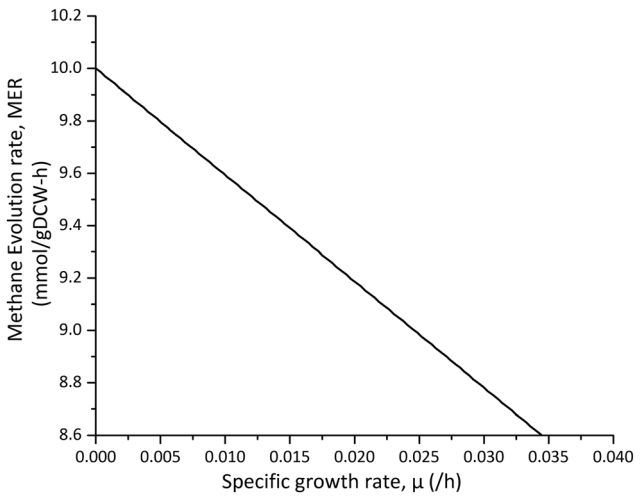

Fig. 7 Effect of the MER on specific growth rates at a carbon uptake of 10 mmol g_{DCW}^{-1} h^{-1}.

Table 4 Top 10 reactions for which MER changes with reaction flux. ΔMER, ΔFlux, and ΔGrowth are the changes in the methane evolution rate, reaction flux, and the specific growth rate, respectively, from their base values for the wild type strain. The base values for the wild type strain are an MER of 9.2619 mmol g_{DCW}^{-1} h^{-1} and a growth rate of 0.01817 h^{-1} for a CO_2 uptake of 10 mmol g_{DCW}^{-1} h^{-1}. ΔFlux is the 10% of the range over which a reaction flux can change to effect methanogenesis

Reaction	Enzyme	E.C. number	Flux range	ΔFlux ($\times10^{-2}$)	ΔMER ($\times10^{-2}$)	ΔGrowth ($\times10^{-3}$)	ΔMER/ ΔFlux	ΔMER/ ΔGrowth
R72	Nitrogenase	1.18.6.1	0.095–0.182	0.87	7.24	−1.78	8.34	−40.69
R53	5,10 MethyleneTHMPT reductase	1.5.99.11	9.537–9.921	3.84	6.12	−1.51	1.59	−40.53
R360	F420-non-reducing hydrogenase	1.12.98.3	9.262–9.988	7.26	7.25	−1.79	0.99	−40.53
R55	THMPT *S*-methyltransferase	2.1.1.86	9.262–9.988	7.26	7.25	−1.79	0.99	−40.53
R56	Coenzyme-B sulfoethylthiotransferase	2.8.4.1	9.262–9.988	7.26	7.25	−1.79	0.99	−40.53
R57	CoB CoM heterodisulfide reductase	1.8.98.1	9.262–9.988	7.26	7.25	−1.79	0.99	−40.53
R46	F420-dependent methylene THMPT dehydrogenase	1.5.99.9	9.921–9.561	3.99	6.15	−1.52	1.54	−40.51
R45	MethenylTHMPT cyclohydrolase	3.5.4.27	9.921–9.561	3.99	6.15	−1.52	1.54	−40.51
R43	FormylmethanofuranTHMPT *N*-formyltransferase	1.2.99.5	9.921–9.561	3.99	6.15	−1.52	1.54	−40.51
R76	H + -transporting two-sector ATPase	3.6.3.14	2.962–3.327	3.65	6.14	−1.68	1.68	−36.58

F_{420}, which in turn increases the activity of coenzyme F_{420} hydrogenases coupled to the reduction of electron acceptors such as CO_2. Thus, the MER increases.

The third approach is to increase the MER by deleting the genes for one or more nonessential reactions. Table 5 shows five genes whose individual or combined deletions enhance the MER. For combinations of up to three genes, the deletion of *adkA* seems to be the best choice based on the ΔMER to ΔGrowth ratio. *adkA* is involved in the biosynthesis of ATP. Its deletion reduces the supply of ATP, hence the MER increases to make up for the shortfall in energy. The deletion of *acd* and *mdh* together seems to be the next best option. Acetate CoA ligase encoded by *acd* plays a key role in propanoate metabolism. Since propanoate metabolism generates energy by phosphorylation of AMP, the deletion of *acd* creates a shortfall in energy supply. Thus, the MER increases to supply the shortage of energy. Malate dehydrogenase encoded by *mdh* catalyzes the conversion of malate and oxaloacetate with simultaneous reduction of NAD. In the absence of *mdh*, NAD specific malic enzymes decarboxylate to form pyruvate. Pyruvate oxidation requires more energy and hence the MER goes up. Table 5 also shows that the deletion of *acd* and *adkA* together may lead to zero growth. Compared to flux variations, the impact of gene deletions seems to be low. But, these can also be used in tandem to further increase methanogenesis.

Table 5 Effects of *in silico* gene knockouts on the MER and cell growth

Gene deletions	Reactions	ΔMER ($\times10^{-2}$)	ΔGrowth ($\times10^{-3}$)	ΔMER/ ΔGrowth
adkA	R243	0.59	−0.14	−59.0
acd, mdh	R495, R26	1.74	−0.43	−43.5
adkA, mdh	R243, R26	0.84	−0.21	−42.0
acd	R495	1.50	−0.37	−37.5
acd, cimA, mdh	R495, R103, R26	1.84	−0.63	−30.7
acd, cimA	R495, R103	1.78	−0.61	−29.7
acd, leuB, cimA	R495, R95, R103	1.78	−0.61	−29.7
mdh	R26	0.24	−0.06	−24.0
adkA, cimA, mdh	R243, R103, R26	0.95	−0.41	−23.8
adkA, leuB	R243, R95	0.89	−0.39	−22.3
adkA, cimA	R243, R103	0.89	−0.39	−22.3
adkA, leuB, cimA	R243, R95, R103	0.89	−0.04	−22.3
leuB, mdh	R95, R26	0.37	−0.27	−12.3
cimA, mdh	R103, R26	0.37	−0.27	−12.3
leuB, cimA, mdh	R95, R103, R26	0.37	−0.27	−12.3
leuB	R95	0.31	−0.25	−10.3
cimA	R103	0.31	−0.25	−10.3
leuB, cimA	R95, R103	0.31	−0.25	−10.3
acd, adkA	R495, R243	N.A.	0.00	N.A.
acd, adkA, leuB	R495, R243, R95	N.A.	0.00	N.A.
acd, mdh, adkA	R495, R26, R243	N.A.	0.00	N.A.

Comparison with other methanogens

Table 6 shows a comparison of our genome-scale metabolic model for *M. maripaludis* S2 with the existing genome-scale models for three methanogens, *viz. M. barkeri*,[10] *M. acetivorans*,[32] and *S. solfataricus.*[34] Our model has the highest ORF coverage among the four. Clearly, large amount of annotated information is available for *M. maripaludis* genome. Among the hydrogenotrophic methanogens, methane metabolism is better studied and characterized in *M. maripaludis*[16] compared to the others. The information helped us to compare reactions/genes involved in the methanogenesis pathways of several methanogens.

The proven higher specific growth rate is a clear advantage for *M. maripaludis* compared to the other three methanogens. The model (*i*AF692) of *M. barkeri* by Fiest *et al.*[10] reported a specific growth of 0.071 h^{-1} and a MER of 8.824 mmol g_{DCW}^{-1} h^{-1}, when hydrogen uptake was limited to 41 mmol g_{DCW}^{-1} h^{-1} with CO_2 as the sole carbon source and cysteine as the sulfur source. Under the same conditions, our model (*i*MM518) predicted a growth of 0.0847 h^{-1} and a MER of 18.40 mmol g_{DCW}^{-1} h^{-1}. This highlights the extraordinary capability of *M. maripaludis* for methane production. The reason could be the presence of H(2)-dependent methyleneTHMPT dehydrogenase associated with *hmd*, which the other three methanogens do not have. This enzyme helps in the reversible oxidation of molecular hydrogen to hydron. It contains neither Ni nor the Fe–S cluster, but a Fe-containing cofactor, as recently characterized by Shima *et al.*[66] *M. acetivorans* is incapable of growing under H_2 and CO_2. Also, no MER data are available for *M. acetivorans*, and *S. solfataricus* lacks most of the enzymes required for methanogenesis. *S. solfataricus* can utilize and assimilate C1 compounds *via* the 3-hydroxypropionate/4-hydroxybutyrate cycle only in the presence of hydrogen and oxygen.[67]

Table 6 Comparative analysis of the *M. maripaludis* S2 model with previously modeled methanogenic archaebacteria

Features	*Methanococcus maripaludis* S2 "mmp"	Methanosarcina barkeri "mba"	*Methanosarcina acetivorans* "mac"	*Sulfolobus solfataricus* "sso"
Reference	*i*MM518, this study 2014	*i*AF692, Feist *et al.* 2006	*i*MB745, Benedict *et al.* 2011	*i*TU515, Ulas *et al.* 2012
Genome size (bp)	1 661 137	4 873 766	5 751 492	2 992 245
ORFs in the organism/ORFs in the model	1722/518	5072/692	4540/745	2978/515
ORF coverage in the model (%)	30.1	13.6	16.4	17.3
Number of reactions/metabolites in the model	570/558	619/558	756/715	718/706
Number of reactions/genes associated with methanogenesis	28/141	30/138	26/145	7/52
Known growth rate (μ h^{-1})	0.301	0.0231	0.0288	0.0577

Materials and methods

Reconstructing metabolic network

The process (Fig. 8) of metabolic network building has been well reviewed in the literature.[68–70] Genome annotations of *M. maripaludis* were obtained from the sequencing studies performed by Hendrickson *et al.*[16] Various biochemical, genetic, and physiological information about *M. maripaludis* was mined from the published literature and public databases such as KEGG,[36] METACYC,[37] and BRENDA.[71] Reactions from multiple databases were integrated with the help of their EC numbers in KEGG. Spontaneous as well as non-gene associated reactions were included, if their existence was proven by physiological or experimental data. Transport reactions were also added to allow the exchange of metabolites between the cell and its environment. The reaction stoichiometries were checked to verify elemental balances for C, S, P, N, and O. Several ORFs were annotated in the process based on known physiological information and employing comparative genomics and local sequence similarity searches[72] (BLASTp) of the Swiss-Prot database[73] with an *e*-value of 10^{-5}. Once the reaction set was completed, it was checked for the production of biomass precursors using the FBA in MetaFluxNet.[74]

A score of 1–5 reflecting our confidence in the available experimental information was assigned to each reaction. The statistical properties of these scores were then used to evaluate model reliability. All the reactions in our metabolic model along with their GPR associations, EC numbers, and metabolites with their chemical formulas are available as an Excel® file (ESI†). An SBML (Systems Biology Markup Language) file of the model has also been deposited in the BioModels database[63] with MODEL1304120000 as its identifier.

Genome-scale metabolic model

The reaction fluxes in a genome-scale model are governed by the following equation. This assumes a pseudo-steady state (no accumulation) for all metabolites:[69]

$$S \cdot v = b \quad (3)$$

where S is the matrix of stoichiometric coefficients of the reactants and products, v is the vector of reaction fluxes (mmol g_{DCW}^{-1} h^{-1}) at any given time, and b is the vector of net fluxes into the cell. An external metabolite is treated distinct from its

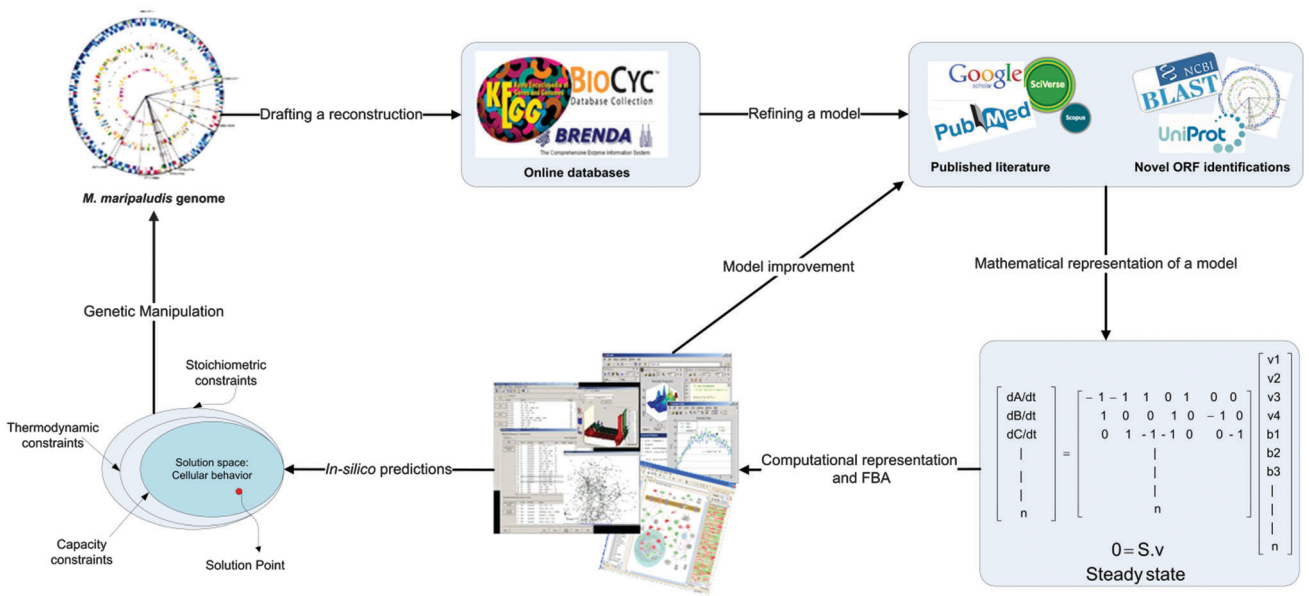

Fig. 8 Schematic representation of the procedure used to reconstruct the constraint-based genome-scale metabolic model of *M. maripaludis* S2.

internal form, and both forms are connected *via* a transport reaction. The *b* values are zero for the internal metabolites.

Constraint-based FBA models were obtained by adding a suitable cellular objective (such as maximize biomass growth) to eqn (3). In these models, suitable lower and upper bounds were assigned to each reaction flux. For instance, the limits on the flux of an irreversible reaction was set as [0, 1000], while that for a reversible reaction was set as [−1000, 1000]. These optimization models (which are linear programs) were solved using MetaFluxNet version 1.8[74] to obtain the flux distributions corresponding to various cellular phenotypes.

The specific growth rate (h^{-1}) of biomass is a commonly used cellular objective for FBA. It quantifies the assembly of various macromolecular precursors into cell biomass. A biomass reaction was formulated to describe this assimilation. Due to the unavailability of an experimental biomass composition for *M. maripaludis*, it was assumed to consist of DNA, RNA, proteins, phospholipids, glycogen, and ATP. DNA and RNA were expressed in terms of five nucleotides (adenine, thymine, guanine, uracil and cytosine) with their compositions estimated following the procedure of Lee *et al.*[75] The compositions of proteins and phospholipids for *M. maripaludis* are available in the literature.[54,76,77] Glycogen synthesis was modeled by the following polymerization reaction of UDP-glucose by glycogen synthase as reported by Yu *et al.*[39] For this reaction, glycogen was assumed to consist of UDP-glucose as monomeric entities. ATP quantifies the energy consumption required for cell growth. It is also known as GAM, which is not available for this organism. Therefore, the procedure reported by Verduyn *et al.*[78] was used to estimate a GAM of 30 mmol ATP g_{DCW}^{-1} h^{-1}. In addition to GAM, a cell also needs energy for the maintenance of ion gradients, regulatory metabolism, RNA turnover, *etc.* This is known as NGAM. The NGAM energy usage was modeled as $ATP + H_2O \rightarrow ADP + Pi$. The NGAM value that best described the experimental data on cell growth was determined during model validation. Detailed description of the biomass composition can be found in ESI,† file 2. Atypical of most bacteria, *M. maripaludis* possesses a fragile proteinaceous S-layer (surface layer)[14,79] instead of a cell wall. Therefore, the major components of an archaebacterial cell wall such as muramic acid and peptidoglycan were not included in our biomass composition.

During our metabolic pathway reconstructions, we identified several gaps. These included known reactions in *M. maripaludis* S2 with unknown ORFs, and known ORFs with unknown reactions. For the former, we added reactions from closely related methanogens and performed BLASTp searches in the protein database of *M. maripaludis* S2 to identify new ORFs using translated protein queries. For the latter, we performed BLASTx searches in the protein database using translated nucleotide queries. Several other genes exist in *M. maripaludis* S2 that are classified as ''conserved hypothetical'', ''hypothetical'', ''genes of unknown function'', *etc.* These genes/proteins can in principle be annotated to enhance the ORF coverage of our model using various methods based on comparative genomics, protein–protein interactions, *etc.* presented in the literature.[80,81]

Experimental data

M. maripaludis is known to grow on CO_2 as the sole carbon source and hydrogen as the electron donor. However, quantitative data on CO_2 or H_2 utilization profiles *vs.* growth are missing in the literature. Our genome-scale model was validated using the experimental data from Timothy *et al.*[51] and Lupa *et al.*[49]

Timothy *et al.*[51] grew a *M. maripaludis* culture in MSH media (MS medium with NaCl, $MgCl_2$, and KCl) with and without organics. A total H_2–CO_2 (75 : 25 v/v) pressure of 300 kPa and a temperature of 25 °C were maintained in anaerobic pressure tubes. A hydrogen utilization rate of 1.6 ng μg_{DCP}^{-1} min^{-1} = 28.8 mmol g_{DCW}^{-1} h^{-1} was measured during cell growth without the organics, and 45 mmolg_{DCW}^{-1} h^{-1} with the organics. DCP here stands for dry cell protein and is assumed to be 60% of DCW. The experimental value of hydrogen exchange was used as the lower bound in our model and FBA was used to compute other reaction fluxes.

Lupa *et al.*[49] performed two experiments with wild-type *M. maripaludis* S2. In the first experiment, cell growth (*via* absorbance at 600 nm) and MERs (*via* gas chromatography) at 37 °C were measured over a period of 25 h. The cultures for this experiment were grown at 37 °C in McNA (minimal medium supplemented with acetate) medium[82] under 276 kPa of the H_2–CO_2 (80 : 20 v/v) atmosphere. In the second experiment, sodium formate was used as a source for H_2 under the O_2-free N_2 atmosphere. The cultures for this experiment were grown under 138 kPa of the N_2 : CO_2 (80 : 20 v/v) atmosphere. To simulate this experiment in our model, the MER values were converted to mmol g_{DCW}^{-1} h^{-1} and set as upper bounds on methane exchange fluxes. FBA predicted the growth rates corresponding to these MERs.

After validation, model predictions were verified using the phenotypical observations from Haydock *et al.*,[55] Lin *et al.*,[56] and Lie *et al.*[57] Haydock *et al.*[55] constructed a leucine-auxotrophic mutant for demonstrating the essential role of the *leuA* gene in the growth of *M. maripaludis* S2. Lin *et al.*[56] showed the importance of *porE* and *porF* genes on growth and oxidation of pyruvate. Lie *et al.*[57] studied the effect of nitrogen on the regulation of *nif* (nitrogen fixation) and *glnA* (glutamine synthetase) operons. Both *nif* and *glnA* expressions were high under diazotrophic (nitrogen-rich) conditions. Also, both operon expressions were observed in the presence of alanine as a nitrogen source.[62]

Gene essentiality/flux variability analyses

Gene essentiality and flux variability analyses were performed using the COnstraint-Based Reconstruction and Analysis (COBRA) toolbox in the MATLAB version 7.12.0.635 (R2011a).

Gene/reaction essentiality analyses are useful for strain improvement strategies, as the deletion of an essential gene/reaction may result in a lethal phenotype. To this end, the minimal media identified in our study were used in a series of simulations with CO_2 and H_2 as the sole carbon and hydrogen sources. 10^{-4} h^{-1} was taken as the threshold for zero cell

growth. For gene essentiality, genes were deleted one at a time by inactivating all the reactions associated with each individual gene. If the deletion of a gene resulted in near-zero biomass, then the gene was classified as essential for growth. Since a reaction catalyzed by multiple genes cannot be inactivated *via* single gene deletions, single reaction deletion was also performed by inactivating reactions one at a time and solving for growth.

Flux variability analysis (FVA) helps to establish theoretical limits on various fluxes for a given phenotype. While FBA gives one set of flux values, FVA shows the extent by which they may vary for the same phenotype prediction. We performed FVA for cell growth and identified the essential reactions.

Conclusions

Biochemical conversion of carbon dioxide using methanogens can be a useful approach for reducing CO_2 emissions. The first genome-scale metabolic model of *M. maripaludis* S2, a methanogen capable of consuming CO_2 as the sole carbon source for biomass and methane production using the Wood–Ljunghdahl pathway, was constructed, validated, and analyzed in this work. The model allows us to study methanogenesis, nitrogen fixation, and other metabolic pathways in *M. maripaludis* S2. The developed model is robust, and successfully predicted various observed phenotypes such as substrate uptake rates, MERs, and the specific growth rate under different experimental conditions. The model enabled the identification of essential genes/reactions for growth and mutant strains capable of enhancing methanogenesis. A model-based comparison with three other methanogens showed the superiority of *M. maripaludis* in growth and methanogenesis. Experimental studies are underway to further validate and enhance the proposed model.

Acknowledgements

The authors would like to acknowledge National University of Singapore for funding this research through its President Graduate Fellowship (PGF) program. This paper is an extended version of the six-page manuscript published in 23rd European Symposium on Computer Aided Process Engineering.[83]

References

1 J. Hansen, M. Sato, R. Ruedy, K. Lo, D. W. Lea and M. Medina-Elizade, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 14288–14293.
2 E. Rubin and H. de Coninck, *TNO (2004): Cost Curves for CO2 Storage, Part*, Cambridge University Press, UK, 2005, vol 2.
3 S. Park, D. Bézier and M. Brookhart, *J. Am. Chem. Soc.*, 2012, **134**, 11404–11407.
4 S.-I. In, D. D. Vaughn and R. E. Schaak, *Angew. Chem.*, 2012, **124**, 3981–3984.
5 W. Li, *ChemInform*, 2012, **43**, 55–76.
6 N. C. M. Earl Boysen, *Nanotechnology For Dummies*, 2nd edn 2011.
7 R. Saini, R. Kapoor, R. Kumar, T. O. Siddiqi and A. Kumar, *Biotechnol. Adv.*, 2011, **29**, 949–960.
8 J. G. Ferry, *Methanogenesis – Ecology, Physiology, Biochemistry & Genetics*, 1994.
9 S. E. Hook, A. D. Wright and B. W. McBride, *Archaea*, 2010, **2010**, 945785.
10 A. M. Feist, B. Ø. Palsson, F. J. Brockman and T. Ideker, *Mol. Syst. Biol.*, 2006, **2**, DOI: 10.1038/msb4100046.
11 J. H. Tsao, S. M. Kaneshiro, S. S. Yu and D. S. Clark, *Biotechnol. Bioeng.*, 1994, **43**, 258–261.
12 W. Whitman, E. Ankwanda and R. Wolfe, *J. Bacteriol.*, 1982, **149**, 852–863.
13 J. Zeikus, G. Fuchs, W. Kenealy and R. Thauer, *J. Bacteriol.*, 1977, **132**, 604–613.
14 W. J. Jones, M. J. B. Paynter and R. Gupta, *Arch. Microbiol.*, 1983, **135**, 91–97.
15 T. J. Lie, G. E. Wood and J. A. Leigh, *J. Biol. Chem.*, 2005, **280**, 5236–5241.
16 E. L. Hendrickson, R. Kaul, Y. Zhou, D. Bovee, P. Chapman, J. Chung, E. Conway de Macario, J. A. Dodsworth, W. Gillett, D. E. Graham, M. Hackett, A. K. Haydock, A. Kang, M. L. Land, R. Levy, T. J. Lie, T. A. Major, B. C. Moore, I. Porat, A. Palmeiri, G. Rouse, C. Saenphimmachak, D. Soll, S. Van Dien, T. Wang, W. B. Whitman, Q. Xia, Y. Zhang, F. W. Larimer, M. V. Olson and J. A. Leigh, *J. Bacteriol.*, 2004, **186**, 6956–6969.
17 B. C. Moore and J. A. Leigh, *J. Bacteriol.*, 2005, **187**, 972–979.
18 I. Porat, W. Kim, E. L. Hendrickson, Q. Xia, Y. Zhang, T. Wang, F. Taub, B. C. Moore, I. J. Anderson, M. Hackett, J. A. Leigh and W. B. Whitman, *J. Bacteriol.*, 2006, **188**, 1373–1380.
19 G. E. Wood, A. K. Haydock and J. A. Leigh, *J. Bacteriol.*, 2003, **185**, 2548–2554.
20 J. L. Argyle, D. L. Tumbula and J. A. Leigh, *Appl. Environ. Microbiol.*, 1996, **62**, 4233–4237.
21 D. L. Tumbula, R. A. Makula and W. B. Whitman, *FEMS Microbiol. Lett.*, 1994, **121**, 309–314.
22 G. M. Jones, J. Wu, Y. Ding, K. Uchida, S. Aizawa, A. Robotham, S. M. Logan, J. Kelly and K. F. Jarrell, *J. Bacteriol.*, 2012, **194**, 2693–2702.
23 A. D. Walters, S. E. Smith and J. P. Chong, *Appl. Environ. Microbiol.*, 2011, **77**, 2549–2551.
24 J. Ladapo and W. B. Whitman, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 5598–5602.
25 C. E. Blank, P. S. Kessler and J. A. Leigh, *J. Bacteriol.*, 1995, **177**, 5773–5777.
26 J. L. Reed and B. O. Palsson, *J. Bacteriol.*, 2003, **185**, 2692–2699.
27 N. D. Price, J. L. Reed and B. O. Palsson, *Nat. Rev. Microbiol.*, 2004, **2**, 886–897.
28 R. Fleischmann, M. Adams, O. White, R. Clayton, R. Tatusov, A. Mushegian, P. Bork, N. Brown, W. Hayes and O. White, *Science*, 1995, **269**, 496–512.
29 J. Pramanik and J. D. Keasling, *Biotechnol. Bioeng.*, 1997, **56**, 398–421.

30 S. Aggarwal, I. A. Karimi and D. Y. Lee, *Mol. BioSyst.*, 2011, **7**, 3122–3131.
31 H. Widiastuti, J. Y. Kim, S. Selvarasu, I. A. Karimi, H. Kim, J. S. Seo and D. Y. Lee, *Biotechnol. Bioeng.*, 2011, **108**, 655–665.
32 V. Satish Kumar, J. G. Ferry and C. D. Maranas, *BMC Syst. Biol.*, 2011, **5**, 28.
33 S. Selvarasu, I. A. Karimi, G.-H. Ghim and D.-Y. Lee, *Mol. BioSyst.*, 2010, **6**, 152–161.
34 T. Ulas, S. A. Riemer, M. Zaparty, B. Siebers and D. Schomburg, *PLoS One*, 2012, **7**, e43401.
35 K. C. Costa, S. H. Yoon, M. Pan, J. A. Burn, N. S. Baliga and J. A. Leigh, *J. Bacteriol.*, 2013, **195**, 1456–1462.
36 M. Kanehisa, S. Goto, Y. Sato, M. Furumichi and M. Tanabe, *Nucleic Acids Res.*, 2012, **40**, D109–D114.
37 R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, T. C. Walk, P. Zhang and P. D. Karp, *Nucleic Acids Res.*, 2008, **36**, D623–D631.
38 J. Shieh and W. B. Whitman, *J. Bacteriol.*, 1987, **169**, 5327–5329.
39 J.-P. Yu, J. Ladapo and W. B. Whitman, *J. Bacteriol.*, 1993, **176**, 325–332.
40 M. Graupner and R. H. White, *J. Bacteriol.*, 2001, **183**, 5203–5205.
41 R. P. Hausinger, W. H. Orme-Johnson and C. Walsh, *Biochemistry*, 1985, **24**, 1629–1633.
42 S. C. Namboori and D. E. Graham, *J. Bacteriol.*, 2008, **190**, 2987–2996.
43 R. H. White, *Biochemistry*, 2004, **43**, 7618–7627.
44 V. Worrell and D. Nagle Jr, *J. Bacteriol.*, 1988, **170**, 4420–4423.
45 S. W. Ragsdale, *Encyclopedia of Catalysis*, John Wiley & Sons, Inc., 2002.
46 Y. Liu, Doctor of philosophy, The University of Georgia, 2010.
47 M. Strassman and L. N. Ceci, *Biochem. Biophys. Res. Commun.*, 1964, **14**, 262.
48 N. Belay, R. Sparling and L. Daniels, *Appl. Environ. Microbiol.*, 1986, **52**, 1080–1085.
49 B. Lupa, E. L. Hendrickson, J. A. Leigh and W. B. Whitman, *Appl. Environ. Microbiol.*, 2008, **74**, 6584–6590.
50 S. Cheng, D. Xing, D. F. Call and B. E. Logan, *Environ. Sci. Technol.*, 2009, **43**, 3953–3958.
51 T. A. Kral, K. M. Brink, S. L. Miller and C. P. McKay, *Origins Life Evol. Biospheres*, 1998, **28**, 311–319.
52 A. M. Feist, B. Ø. Palsson, F. J. Brockman and T. Ideker, *Mol. Syst. Biol.*, 2006, **2**, DOI: 10.1038/msb4100046.
53 R. Mahadevan, D. Bond, J. Butler, A. Esteve-Nunez, M. Coppi, B. Palsson, C. Schilling and D. Lovley, *Appl. Environ. Microbiol.*, 2006, **72**, 1558–1568.
54 S. Stolyar, S. Van Dien, K. L. Hillesland, N. Pinel, T. J. Lie, J. A. Leigh and D. A. Stahl, *Mol. Syst. Biol.*, 2007, **3**, 92.
55 A. K. Haydock, I. Porat, W. B. Whitman and J. A. Leigh, *FEMS Microbiol. Lett.*, 2004, **238**, 85–91.
56 W. Lin and W. B. Whitman, *Arch. Microbiol.*, 2004, **181**, 68–73.
57 T. J. Lie and J. A. Leigh, *Mol. Microbiol.*, 2003, **47**, 235–246.
58 T. J. Lie, K. C. Costa, B. Lupa, S. Korpole, W. B. Whitman and J. A. Leigh, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 15473–15478.
59 T. A. Major, Y. Liu and W. B. Whitman, *J. Bacteriol.*, 2010, **192**, 4022–4030.
60 K. C. Costa, T. J. Lie, M. A. Jacobs and J. A. Leigh, *mBio*, 2013, **4.2**, e00062-13.
61 P. S. Kessler, C. Daniel and J. A. Leigh, *J. Bacteriol.*, 2001, **183**, 882–889.
62 T. J. Lie and J. A. Leigh, *J. Bacteriol.*, 2002, **184**, 5301–5306.
63 T. J. Lie and J. A. Leigh, *Mol. Microbiol.*, 2002, **47**, 235–246.
64 N. Belay, R. Sparling, B.-S. Choi, M. Roberts, J. Roberts and L. Daniels, *Biochim. Biophys. Acta, Bioenerg.*, 1988, **971**, 233–245.
65 M.-L. Fardeau, J.-P. Peillex and J.-P. Belaich, *Arch. Microbiol.*, 1987, **148**, 128–131.
66 S. Shima, O. Pilak, S. Vogt, M. Schick, M. S. Stagni, W. Meyer-Klaucke, E. Warkentin, R. K. Thauer and U. Ermler, *Science*, 2008, **321**, 572–575.
67 I. A. Berg, D. Kockelkorn, W. Buckel and G. Fuchs, *Science*, 2007, **318**, 1782–1786.
68 R. S. Senger and E. T. Papoutsakis, *Biotechnol. Bioeng.*, 2008, **101**, 1036–1052.
69 M. Terzer, N. D. Maynard, M. W. Covert and J. Stelling, *Wiley Interdiscip. Rev.: Syst. Biol. Med.*, 2009, **3**, 285–297.
70 W. Zou, L. Liu, J. Zhang, H. Yang, M. Zhou, Q. Hua and J. Chen, *J. Biotechnol.*, 2012, **161**, 42–48.
71 M. Scheer, A. Grote, A. Chang, I. Schomburg, C. Munaretto, M. Rother, C. Sohngen, M. Stelzer, J. Thiele and D. Schomburg, *Nucleic Acids Res.*, 2011, **39**, D670–D676.
72 S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403–410.
73 A. Bairoch and B. Boeckmann, *Nucleic Acids Res.*, 1992, **20**, 2019.
74 D.-Y. Lee, H. Yun, S. Park and S. Y. Lee, *Bioinformatics*, 2003, **19**, 2144–2146.
75 J. Lee, H. Yun, A. M. Feist, B. O. Palsson and S. Y. Lee, *Appl. Microbiol. Biotechnol.*, 2008, **80**, 849–862.
76 J. Lombard, P. Lopez-Garcia and D. Moreira, *Archaea*, 2012, **2012**, 630910.
77 B. Tenchov, E. M. Vescio, G. D. Sprott, M. L. Zeidel and J. C. Mathai, *J. Biol. Chem.*, 2006, **281**, 10016–10023.
78 C. Verduyn, A. H. Stouthamer, W. A. Scheffers and J. P. Dijken, *Antonie van Leeuwenhoek*, 1991, **59**, 49–63.
79 Q. Xia, T. Wang, E. L. Hendrickson, T. J. Lie, M. Hackett and J. A. Leigh, *BMC Microbiol.*, 2009, **9**, 149.
80 J. L. Reed, I. Famili, I. Thiele and B. O. Palsson, *Nat. Rev. Genet.*, 2006, **7**, 130–141.
81 S. Sivashankari and P. Shanmughavel, *Bioinformation*, 2006, **1**, 335.
82 W. B. Whitman, J. Shieh, S. Sohn, D. S. Caras and U. Premachandran, *Syst. Appl. Microbiol.*, 1986, **7**, 235–240.
83 N. Goyal, H. Widiastuti, I. A. Karimi and G. Zhou Zhi, in *Computer Aided Chemical Engineering*, ed. K. Andrzej and T. Ilkka, Elsevier, 2013, vol. 32, pp. 181–186.