Biodegradation (2008) 19:99–127
DOI 10.1007/s10532-007-9119-5

ORIGINAL PAPER

Expanded thermodynamic true yield prediction model: adjustments and limitations

Jinghua Xiao · Jeanne M. VanBriesen

Received: 5 September 2006 / Accepted: 10 April 2007 / Published online: 12 June 2007

Abstract Bacterial yield prediction is critical for bioprocess optimization and modeling of natural biological systems. In previous work, an expanded thermodynamic true yield prediction model was developed through incorporating carbon balance and nitrogen balance along with electron balance and energy balance. In the present work, the application of the expanded model is demonstrated in multiple growth situations (aerobic heterotrophs, anoxic, anaerobic heterotrophs, and autolithotrophs). Two adjustments are presented that enable improved prediction when additional information regarding the environmental conditions (pH) or degradation pathway (requirement for oxygenase- or oxidase-catalyzed reactions) is known. A large data set of reported yields is presented and considered for suitability in model validation. Significant uncertainties of literature-reported yield values are described. Evaluation of the model with experimental yield values shows good predictive ability. However, the wide range in reported yields and the variability introduced into the prediction by uncertainty in model parameters, limits comprehensive validation. Our results suggest that the uncertainty of the experimental data used for validation limits further improvement of thermodynamic prediction models.

Keywords Bacterial thermodynamics · Bacterial yield · Yield prediction · Modeling bacterial growth · Microbial yield

List of symbols

CS	carbon source
D	dilution rate in continuous culture, 1/h
ED	electron donor
EA	electron acceptor
E_{syn} (1)	the Gibbs energy change from carbon source to acetate, kJ/mol-C
E_{syn} (2)	the Gibbs energy change from acetate and ammonia to biomass, kJ/mol-C
f_{cell}:	the fraction of carbon from carbon source to new biomass
f_{CO_2}	the fraction of carbon from carbon source to carbon dioxide;
$f_{CS}(i)$	the fraction of carbon in carbon source going to product (i)
$f_{ED}(i)$	the fraction of electron donor element in electron donor going to its product (i)
$f_{N\text{-}C}$	the fraction of nitrogen in nitrogen source going to cell synthesis
$f_N(i)$:	the fraction of nitrogen in nitrogen source going to product (i)
$g(j)$	the electrons accepted by electron acceptor (j), e^- eq

J. Xiao · J. M. VanBriesen (✉)
Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA
e-mail: jeanne@cmu.edu

[H] the concentration of hydrogen ion, mol/l
DW dry weight, g/mol
K the energy utilizing coefficient
MW molecular weight, g/mol
NS nitrogen source
t_{oxy} the times of oxygenase reaction happened during 1 mol ED degraded, times/mol-ED
Y^{m}_{exp} the experimental maximum yield (or true yield) of bacteria, mol-C cell/mol-C or mol-C cell/mol-ED
Y^{m}_{est} the estimated maximum yield (or true yield) of bacteria, mol-C cell/mol-C or mol-C cell/mol-ED
γ the reductance degree, e^- eq/mol-C or e^- eq/mol-ED
γ_X the reductance degree of biomass, e^- eq/mol-C
ΔG Gibbs energy change, kJ/mol
ΔG_{ED} the Gibbs energy change of ED to its products, kJ/mol-ED
ΔG_{eD}: the Gibbs energy change of ED to its products, kJ/e^- eq.
ΔG_{EA} the Gibbs energy change of EA to its products, kJ/mol-EA
$\Delta G_{e\text{-}EA}$ the energy change of EA to its products, kJ/e^- eq.

Introduction

Bacterial yield is a critical parameter in understanding the behavior of microorganisms in the environment and is a key factor to consider in biological process design. Prediction of biomass allows optimization of biological treatment processes, prevention of biofouling, and understanding the biodegradation of anthropogenic compounds and the transport of biomass in natural systems (Argaman and Brenner 1986; Corseuil and Weber 1994; Heijnen 1984, 1994; Henze et al. 1987, 1995; Kappler et al. 1997; Liebi et al. 2001; Muller et al. 2003; Woo and Rittmann 2000; Yuan and VanBriesen 2002). With so many applications that rely on an estimate of bacterial yield, different thermodynamic methods have been developed to predict yield (Heijnen 1999; McCarty 1965). However, yield predictions from thermodynamic models are sometimes far from experimental results. Many factors cause this, including: (1) experimental error associated with yield determination, (2) reliance on estimates of bacterial maintenance energy to extrapolate from observed to theoretical maximum yield, (3) choices of reference state in thermodynamic models, and (4) simplifying assumptions in predictive models.

Xiao and VanBriesen (2006) presented a thermodynamic model, based on the original work of McCarty (1965) and incorporating carbon and nitrogen balance explicitly to overcome the limitation introduced by the assumption that electron and energy balances alone controlled the maximum bacterial growth. The conceptual model was compared with other models and showed good predictive ability for aerobic heterotrophs. In the present work, we improve on the foundation laid by Xiao and VanBriesen (2006) by (1) including consideration of the effect of environmental conditions (pH) and pathway information (oxygenase and oxidase reactions), (2) evaluating the effect of uncertainties in assumed model parameters (for biomass formula and efficiency terms) on the yield prediction, (3) evaluating variability in reported experimental values and discussing causes of this variability, (4) considering how variability in reported values for yields affects model validation.

Model refinement

The expanded true yield prediction model can be summarized in four balance equations: carbon balance, electron balance, nitrogen balance, and energy balance with an energy utilizing coefficient, K, where K was estimated from theoretical considerations as 0.41 (Xiao and VanBriesen 2006). The mathematical expressions of those equations and their explanation are shown in Table 1. For a detailed explanation of the development of these balances and their validation, the reader is referred to our earlier work (Xiao and VanBriesen 2006). Implementation of the model requires the specification of the carbon source (CS), nitrogen source (NS), electron donor (ED), electron acceptor (EA), and their products in the system. The energy balance requires assuming an appropriate reference state to write the half reactions and requires analyzing the energy released from some particular

Table 1 Mathematical expressions of the expanded true yield prediction model[a,b]

		Mathematic expression
Elemental balances	Carbon	$f_{\text{cell}} + f_{\text{CS}}\ (i) = 1$
	Nitrogen	$f_{\text{N-C}} + \sum f_{\text{N}}(i) = 1$
	Electron	$\sum f_{\text{ED}}(i)^{i} \times (\gamma_{\text{ED}_2} - \gamma(i)) = \sum g(j)$
Energy balance		$K^{i} \times \sum_{i} G(i) + \sum_{i=1}^{j} K(i) \times E_{\text{syn}}^{j}(i) = 0$

[a] All the symbols are shown in the list of symbols

[b] K(i) = K if $E_{\text{syn}}(\text{i}) < 0$, $K(i) = K^{-1}$ if $E_{\text{syn}}(\text{i}) > 0$

reactions like oxygenase-catalyzed transformations. Therefore, in order to improve true yield predictions, two systematic refinements are made to consider these two factors that bear directly on the predictive capability of thermodynamic models—selection of the reference state and assumptions regarding the energy associated with specialized enzymatic reactions.

Reference state

Critical for thermodynamic yield prediction models is determination of the energy available from the reaction of the ED and EA. Commonly, the energy is calculated using Gibbs energy changes for a pair of half reactions (one for the ED and one for the EA). These half reactions are generated for a specific frame of reference. For instance, Rittmann and McCarty (2001) select H_2CO_3*and* HCO_3^- as the frame of reference for inorganic carbon in the system. VanBriesen (2002) recommended a frame of reference consisting of H_2CO_3, H_2O(aq), N_2(g), O_2(g), and H^+ at pH 7. Heijnen and van Dijken (1993) specifically question these selected frames of reference, suggesting that efficiency values (denoted as K in the models of McCarty and colleagues) are dependent upon the thermodynamic frame of reference. VanBriesen (2002) showed an equal dependence of the dissipation-based prediction methodology (Heijnen and vanDijken 1992) on frame of reference. McCarty (2007) has suggested these effects are significant when low values of Gibbs energy are associated with the reactions such as for anaerobic processes. Our work confirms this and develops a method to create a set of system-dependent half reactions by specification of the system pH.

Different frames of reference in thermodynamic prediction models are based predominantly on selecting different forms of inorganic carbonate to express the carbon balance in the overall reaction. The actual distribution of inorganic carbon in an aqueous system depends on environmental pH rather than some preselected reference carbon compound. Thus, we suggest using all the three species as the reference for inorganic carbon and calculating the stoichiometry of those species based on the system pH value.

In the expanded model, the energy released during the oxidation process of substrate ($C_aH_bO_c^{n-}$) which is symbolized as $\Delta G_{C_aH_bO_c^{n-}}$ (in the unit of kJ/mol-C) is calculated based on the half reaction shown in Eq. (1).

$$\begin{aligned}\frac{1}{a}C_aH_bO_c^{n-} + (3-\tfrac{c}{a})H_2O \\ = \alpha 1 H_2CO_3 + \alpha 2 HCO_3^- + \alpha 3 CO_3^{2-} \\ +(\tfrac{b}{a} + 2\times(3-\tfrac{c}{a}) - 2\alpha 1 - \alpha 2)H^+ + \gamma_s e^-\end{aligned} \quad (1)$$

where $\alpha 1 = 1/(\frac{k_{a1}k_{a2}}{[H]^2} + \frac{k_{a1}}{[H]} + 1)$, $\alpha 2 = 1/(\frac{k_{a2}}{[H]} + \frac{[H]}{k_{a1}} + 1)$, and $\alpha 3 = 1-\alpha 1-\alpha 2$. Here, k_{a1} and k_{a2} are the ionization constants of carbonic acid and reported as $10^{-6.35}$ and $10^{-10.33}$, respectively (Benjamin 2002); [H] is the activity of H^+(aq); at pH 7, $\alpha 1 = 0.18$, $\alpha 2 = 0.82$, and $\alpha 3 = 0$.

Gibbs energy changes for reaction (1) (ΔG_{eD}) of different compounds at pH 4 and 7 were calculated with and without pH adjustment. The results are shown in Table 2. At neutral pH, low error is associated with assuming a single form of carbonate as reference state. However, at low pH the error can be considerable; the average difference in predicted energy associated with the electron donor half reaction is 21%. This change is because when the reference state is allowed to vary with pH value, the

Table 2 Gibbs energy of the electron donor half reaction (ΔG_{eD}) calculated based on different reference states at different pH values

	pH 7			pH 4		
	Without pH adjustment	With pH adjustment	Difference	Without pH adjustment	With pH adjustment	Difference[a]
Acetate	−26.32 [b]	−26.58	−0.98%	−9.21	−10.92	−17.0%
Citrate	−31.63	−31.98	−1.10%	−14.53	−16.80	−14.5%
Formate	−39.10	−38.79	0.80%	−21.99	−12.64	−27.0%
Malonate	−29.48	−29.67	−0.64%	−12.38	−21.73	−24.3%
Oxalic acid	−52.30	−51.68	1.18%	−35.19	−17.63	−32.7%
Tartrate	−38.52	−38.93	−1.06%	−21.41	−24.14	−12.0%
Average			−0.30%			−21.2%

[a] The difference of ΔG_{eD} estimated by two methods at the same pH is calculated as $Difference = \frac{\Delta G_{eD}(\text{with pH adjust}) - \Delta G_{eD}(\text{without pH adjust})}{\text{Average } of\ \Delta G_{eD}}$

[b] All values in kJ/e-eq. By convention a negative sign represents energy released while a positive sign refers to energy consumed

distribution of carbonate is significantly different at pH 4 (e.g., 99.6% H_2CO_3 and 0.4% HCO_3^-) compared with pH 7 (18.3% H_2CO_3 and 81.7% HCO_3^-). This change to the acetate electron donor half reaction may lead to a small change in predicted yield, such as when oxygen is the electron acceptor (less than 1% change). Or, it can lead to significant yield prediction errors, such as when sulfate is the electron acceptor (8% change).

In addition to the effect of pH on the distribution of CO_2 in aqueous systems, pH can also affect the speciation of the primary substrate in the system since many organic compounds are weak acids or bases. If the dominant aqueous species of the substrate can be determined, an improved yield estimation can be made. This effect is easily demonstrated with the well-known chelating agent, ethylenediaminetetraacetic acid (EDTA). Table 3 shows the difference of ΔG_{eD} for $EDTA^{4-}$ and H_6EDTA^{2+} is as high as 37% and causes a 17% difference in true yield estimation (range 0.300–0.355 according to the degradation pathway with 2 monooxygenase reactions and 2 oxidase reactions) (Egli 2001). The distribution of EDTA species in aqueous system (0.01 mol/l $EDTA^{-4}$ at pH 7) can be obtained using chemical speciation models (e.g., Visual Minteq). The results show the dominant species is $HEDTA^{-3}$ at neutral pH, suggesting the model predicted yield of 0.311 mol-C/mol-C is most representative for organisms grown near neutral pH. The effect of speciation on bacterial biodegradation is an active area of research, and the concentration of the dominant aqueous form affects rates as well as thermodynamic predictions (Kari and Giger 1996; VanBriesen and Rittmann 2000).

It is worth noting that the proposed change herein, to consider system pH and its effect on energetics, is a step away from a black-box model based on thermodynamic constants at standard state (e.g., ΔG_f^{01}). Generally, this type of model makes no attempt to incorporate the concentrations of species in order to use system-specific Gibbs energy values. The potential importance of including this type of detail in the model is discussed by Rittmann and McCarty (2001) as well as in the work of Noguera et al (1998), where the concentration of hydrogen as a reactant was specifically considered. Additional research to explore the effect of environmental conditions, including substrate concentrations, and speciation on biological processes is needed.

Enzymatic reactions with lost electrons

When the energy released during catabolism can not be captured by bacteria and used for synthesizing ATP, this ‘‘lost’’ energy must be deducted from the energy balance. Woo and Rittmann (2000), VanBriesen (2001), and Yuan and VanBriesen (2002) all discuss overprediction using the thermodynamic yield method when substrate transformation depends upon oxygenase-catalyzed reactions. These authors hypothesized that oxygenase-catalyzed reactions involved in biodegradation of these substrates divert electrons from the electron acceptor, thus reducing

Table 3 The estimated true yield on different EDTA species at pH 7

Species	Distribution[a] (%)	ΔG_f (kJ/mol)[b]	ΔG_{eD} (kJ(e^- eq.)$^{-1}$)	Estimated true yield (mol-C)(mol-C)$^{-1}$
$EDTA^{4-}$	0.22	−1209.15	−32.52	0.300
$HEDTA^{3-}$	93.76	−1144.33	−35.60	0.311
H_2EDTA^{2-}	6.02	−1079.51	−38.68	0.323
H_3EDTA^{1-}	0.00	−1014.69	−41.76	0.334
H_4EDTA	0.00	−949.87	−44.84	0.345
H_5EDTA^{1+}	0.00	−941.09	−46.27	0.350
H_6EDTA^{2+}	0.00	−932.31	−47.71	0.355
Difference[c]			36.99%	16.61%
Range			−47.71 *to* −32.52	0.300–0.355

[a] The distribution of different EDTA species is calculated by Visual Minteq

[b] All the ΔG_f values are estimated based on Mavrovouniotis (1991)

[c] The difference is defined as Difference $= \frac{\text{biggest}-\text{smallest}}{\text{average}}$

energy available for synthesis. McCarty (2007) has suggested a similar approach to consideration of oxygenase-catalyzed reactions by considering the energy loss for each oxygenation reaction as equivalent to the energy associated with NADH oxidation. In this section, we incorporate the core ideas presented in VanBriesen (2002) into the expanded model of Xiao and VanBriesen (2006). In addition, we also take into account the energy incorporated into the substrate during oxygenase reactions, and we add consideration of the energy loss involved in certain oxidase reactions. Thus the complete model contains the four critical balances and consideration of electrons diverted and energy dissipated in oxygenase and oxidase reactions.

In general, if oxygenase reactions are involved in the degradation pathway, an additional EA, EA(i), the corresponding electrons gained by it, $g(i)$, and the energy released per electron equivalent, $\Delta G_{\text{e-EA}}(i)$ must be considered (see Example in Appendix A for details). Based on the enzyme participating in the reaction, oxygenase reactions can be divided into two groups: monooxygenase-catalyzed reactions and dioxygenase-catalyzed reactions. Different conceptual models and energy calculation are needed for these reactions (see Appendix A for computation details). For monooxygenase reactions, the electrons gained are, $g(i) = 4 \times t_{\text{monooxy}}$, where t_{monooxy} is the number of the monooxygenase reactions during transformation of 1 mol-C substrate. The available energy released from EA(i) is estimated as $\Delta G_{\text{e-EA}}(i) = \Delta G_{\text{e-monooxy}} = -2.263$ kJ/e^- eq for the monooxygenase reaction reacting on the group of $-CH_3$. For a dioxygenase reaction, $g(i) = 4 \times t_{\text{dioxy}}$. The available released energy is estimated as $\Delta G_{\text{e-EA}}(i) = \Delta G_{\text{e-dioxy}} = -12.23$ kJ/e^- eq for those reactions on the benzene ring. The energy inserted into the substrate varies a little with the structure of the substrate. However, comparing with the energy released from O_2 reduction as a common electron acceptor ($\Delta G_{\text{e-EA}}(O_2) = -78.685$ kJ/e^- eq), the stored energy during oxygenase reaction is negligible. Therefore, this small energy modification, while providing explicit consideration of the effect of the inserted oxygen, generally does not significantly alter the yield estimation for heterotrophs. Thus, it is feasible to ignore it and only consider the importance of the 4 electron equivalents lost per oxygenase reaction. For autotrophic systems, however, the stored energy due to the oxygenase reaction has to be considered since the overall energy generation is very low.

Certain oxidase-catalyzed reactions can also affect the net electrons released from biodegradation of a substrate although more experimental evidence is needed is this area. For example, flavin-continaing oxidase enzymes remove two electrons from the substrate with reduction of the flavin (FAD to $FADH_2$). The re-oxidation by molecular oxygen in the cytoplasm is not energy generating, therefore, the removed electrons must be taken out of the electron balance just as the electrons involved in oxygenation reactions. This idea can be expressed numerically in the expanded thermodynamic model by considering

oxygen in oxidase reaction as an independent electron acceptor, EA(i), and the corresponding electrons and energy as $g(i) = 2 \times t_{\text{oxidase}}$ and $\Delta G_{\text{e-EA}}(i) = 0$, where t_{oxidase} is the times of oxidase reaction during 1 mole ED degradation.

In order to show details of model calculation and adjustment application, an example of methanotrophic growth is attached as Appendix B. In the next section, we consider the limitations of the model that remain even after the critical refinements just discussed.

Model uncertainty and limitations

The thermodynamic yield prediction model is based on a series of assumptions related to the uniformity of bacterial cells and the suitability of average metabolic parameters as representative of most growth processes. However use of these assumptions and averages introduces some uncertainty into the model and limits the application of the model to predict the yield for a specific organism on a specific substrate under specific experimental conditions unless the assumptions and averages are updated based on those very specific conditions.

Cell synthesis cost and bacterial biomass composition

As a black box model, the expanded thermodynamic model ignores differences among bacterial species and growth stages. It assumes the same energy need for cell synthesis and the same reductance degree of the carbon in biomass. However, research has shown biomass composition varies with the species, the growth condition, and the available substrate(s) (Atkinson and Mavituna 1991; Roels 1983). Rittmann and McCarty (2001) listed 19 empirical formulas for prokaryotic cells taken from literature source. The molecular weight varies from 19 to 26.8 g/mol-C and the reductance degree of carbon varies from 3.9 to 4.7 (average = 4.2). Therefore, the assumption of a single formula to represent cells must introduce some error to the estimated true yields. Cell synthesis cost is estimated in the model as

$$\Delta G_{\text{cell}} = \frac{\Delta G_{\text{ATP}} \times MW_{\text{cell}}}{Y_{\text{ATP}} \times \text{fraction organic}} \quad (2)$$

where ΔG_{ATP} is assumed to be 30.53 kJ/mol ATP (Burton 1958) and Y_{ATP} is assumed to be 10.5 gDW/mol ATP (Bauchop 1958). The organic fraction of the dry weight (often assumed to be 90%) and the molecular weight of the cell biomass depend upon the organism and its cellular composition. Thus, using this equation, different cellular costs would be computed for different cellular compositions. This is shown in Table 4; the computed synthesis cost varies from 61.57 kJ/mol to 86.58 kJ/mol-C (average is 77.7 kJ/mol-C).

Efficiency terms

Variability is also expected in the energy utilizing coefficients required in the model. In our previous work (Xiao and VanBriesen 2006), two energy utilizing coefficients, *K1* and *K2* were estimated as 0.393 ± 0.003 and 0.41 ± 0.06 with 95% confidence level, respectively. Therefore, the energy coefficient has the range from 0.35 to 0.47. But only one coefficient ($K = 0.41$) was proposed to be used in the model for simplicity. Using one coefficient eases the computational burden for the model; however, it adds uncertainty into the prediction. In addition, the *K2* was estimated based on the overall model structure with an assumed biomass formula of $CH_2O_{0.6}N_{0.2}$ so that the model uncertainty caused by diverse biomass formation discussed above is then transferred to *K2*. The uncertainty in the expanded model can be represented as $K = 0.41 \pm 0.06$. The uncertainty in yield prediction caused by the uncertainty of the energy coefficient is related to the specific growth system. For aerobic systems, the variability in K leads to a relative estimation error of ±16% or ±0.08 mol-C/mol-C of absolute estimation difference. For systems with lower energy consumption, the introduced absolute difference should be smaller but the relative estimation error will be higher due to the low yields in these systems. For example, assuming the energy released from the EA is 0 (to simulate a low energy situation), the absolute difference based on the range of K values is 0.05 mol-C/mol-C but the relative estimation error is 24%. It is critical to note here that the K value proposed by Xiao and VanBriesen (2006) was computed from *a priori* values for critical constants and is NOT fit from a set of experimental data. McCarty (2007) recently updated the suggested

Table 4 The effect of assumed biomass formula on the estimated energy need for cell synthesis

Formula	MW (g)(mol-C)$^{-1}$	γ_X	ΔG_{cell} (kJ)(mol-C)$^{-1}$
$CH_2O_{0.6}N_{0.2}$[a]	26.40	4.20	85.29
$C_5H_7O_2N$	22.60	4.00	73.01
$C_7H_{12}O_4N$	24.86	4.14	80.31
$C_9H_{15}O_5N$	24.11	4.22	77.90
$C_9H_{16}O_5N$	24.22	4.33	78.25
$C_{4.9}H_{9.4}O_{2.9}N$	26.24	4.12	84.79
$C_{4.7}H_{7.7}O_{2.1}N$	23.77	4.11	76.78
$C_{4.9}H_9O_3N$	26.49	4.00	85.58
$C_5H_{8.8}O_{3.2}N$	26.80	3.88	86.58
$C_{4.1}H_{6.8}O_{2.2}N$	25.66	3.85	82.89
$C_{5.1}H_{8.5}O_{2.5}N$	24.25	4.10	78.36
$C_{5.3}H_{9.1}O_{2.5}N$	23.91	4.21	77.23
$C_5H_8O_2N$	22.80	4.20	73.66
$C_5H_{8.33}O_{0.81}N$	19.06	4.74	61.57
$C_5H_8O_2N$	22.80	4.20	73.66
$C_{4.17}H_{7.42}O_{1.38}N$	22.43	4.40	72.47
$C_{4.54}H_{7.91}O_{1.95}N$	23.70	4.22	76.56
$C_{4.17}H_{7.21}O_{1.79}N$	23.95	4.15	77.39
$C_{4.16}H_8O_{1.25}N$	22.10	4.60	71.39
$C_{3.85}H_{6.69}O_{1.78}N$	24.77	4.03	80.03
Range	19.06–26.80	3.85–4.74	61.57–86.58

[a] $CH_2O_{0.6}N_{0.2}$ is the biomass formula suggested by VanBriesen (2002); other formulas are reported in Rittmann and McCarty (2001)

value of K for his thermodynamically based model (TEEM) by fitting a larger and more reliable set of experimental data; a suggested value of this fitted parameter was 0.37.

Experimental yield data

In order to evaluate any true yield prediction model, reported experimental true yield data (observed yields adjusted by maintenance energy) are needed. Table 5 shows a collection of aerobic heterotropic bacterial yield data. The 134 yield values are from seventeen sources (Andrews 1993; Atkinson and Mavituna 1991; Bally et al. 1994; Birou et al. 1987; Boogerd et al. 1984; De Vries et al. 1980; Heijnen and Roels 1981; Heijnen and vanDijken 1992; Henneken et al. 1998; Linton and Stephenson 1978; Pronk et al. 1990; Rutgers 1990; Reardon et al. 2000; van Verseveld 1979; Verduyn 1991; vonStockar and Liu 1999; Yuan 2004). Aerobic heterotrophic yield data previously compiled in Heijnen and van Dijken (1992) are included. About 57 different biological scenarios and 46 organic compounds are included in this table. There are many other organic compounds that can be utilized by bacteria for which experimental yields are not readily available, and thus they are not considered for validation. In addition to the reported yield value, this table includes significant details from the original citation, including yield reported in original units and any error terms reported. While several yield compilations exist in the literature (Heijnen and Roels 1981; Linton and Stephenson 1978; Heijnen and vanDijken 1992), none include this critical information. The importance of this information will be discussed below.

For fermentors, methanogens, and autotrophic microorganisms, data from Heijnen and vanDijken (1992) will be used for analysis and are not re-tabulated here. Rather they are presented in comparative tables in the results section below. Compared with the aerobic microorganisms, fewer anaerobic data and autotrophic data are reported, and pure culture studies are limited in the literature. Moreover, incomplete degradation is common in anaerobic systems. The products of fermentation are varied

Table 5 Experimental data for heterotrophic aerobic microorganisms

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Oxalate	1	O_2/H_2O	0	NH_3	0.086	Rutgers 1990; the extrapolated yield at dilution rate of 0.2/h	Standard deviation (stdev) = 0.003
Oxalic acid	1	O_2/H_2O	0	NH_3	0.07	Heijnen and vanDijken 1992[e]	
Oxalic acid	1	O_2/H_2O	0	NH_3	0.061	Linton and Stephenson 1978[f]	
Formate	2	O_2/H_2O	0	NH_3	0.162	Rutgers 1990; the extrapolated yield at dilution rate of 0.2/h	Stdev = 0.05.
Formate	2	O_2/H_2O	0	NH_3	0.18	Heijnen and vanDijken 1992[e]	
Formate	2	O_2/H_2O	0	NH_3	0.116	van Verseveld 1979	The original true yield was reported as 2.9 (95% CI of 2.6–3.3) gbiomass/mol; 0.12 mol-C/mol-C was reported in Heijnen and vanDijken 1992 (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Formate	2	O_2/H_2O	0	NH_3	0.17	Andrews 1993;	
Formate	2	O_2/H_2O	0	NH_3	0.155	Linton and Stephenson 1978[f]	
Formate	2	O_2/H_2O	0	NH_3	0.155	Linton and Stephenson 1978[f]	
Formate	2	NO_3^-/N_2	0	NO_3^-	0.166	De Vries et al. 1980	The original reported net yields were 3.9 and 4.3 g/mol at dilution rate, D = 0.12 and 0.24/h. 0.166 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e].
Glyoxylate	2	O_2/H_2O	0	NH_3	0.220	Rutgers 1990;	Stdev = 0.016. The extrapolated yield at dilution rate of 0.2/h
Tartrate	2.5	O_2/H_2O	0	NH_3	0.280	Rutgers 1990;	Stdev = 0.001. The extrapolated yield at dilution rate of 0.2/h
Malonate	2.67	O_2/H_2O	0	NH_3	0.238	Rutgers 1990;	Stdev = 0.041. The extrapolated yield at dilution rate of 0.2/h
Nitrilotriacetate (NTA)	3	O_2/H_2O	1	NH_3	0.272	Yuan 2004;	Decay values reported by others were used to estimate true yields
Nitrilotriacetate (NTA)	3	O_2/H_2O	1	NH_3	0.271	Bally et al. 1994;	
Iminodiacetate (IDA)	3	O_2/H_2O	1	NH_3	0.240	Yuan 2004;	Decay values reported by others were used to estimate true yields
Iminodiacetate (IDA)	3	O_2/H_2O	0	NH_3	0.333	Yuan 2004;	Decay values reported by others were used to estimate true yields

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Citrate	3	O_2/H_2O	0	NH_3	0.39	Rutgers 1990;	Stdev = 0.011. The extrapolated yield at dilution rate of 0.2/h
Citrate	3	O_2/H_2O	0	NH_3	0.424	Verduyn 1991;	0.411 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Citrate	3	O_2/H_2O	0	NH_3	0.365	Heijnen and vanDijken 1992[e];	
Citrate	3	O_2/H_2O	0	NH_3	0.37	Andrews 1993;	
Citrate	3	O_2/H_2O	0	NH_3	0.338	Linton and Stephenson 1978[f];	
Citrate	3	O_2/H_2O	0	NH_3	0.359	Birou et al. 1987;	
Malate	3	O_2/H_2O	0	NH_3	0.416	van Verseveld 1979 (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$)	The original reported yield was 41.7 (95% CI of 38.8–45.1) gbiomass/mol; 0.42 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e].
Malate	3	O_2/H_2O	0	NH_3	0.25	Pronk et al. 1990 (the yields reported were not true yield but net yield; assuming the biomass formula of $C_5H_7O_2N$ and 90% biomass content in dry weight; the same mol-C yields were reported in Heijnen and vanDijken 1992[e])	The original report was 25 gDW/mol.
Malate	3	O_2/H_2O	0	NH_3	0.375	Heijnen and vanDijken 1992[e];	
Malate	3	O_2/H_2O	0	NH_3	0.327	Linton and Stephenson 1978[f];	
Glycine	3	O_2/H_2O	0	NH_3	0.248	Yuan 2004;	Decay values reported by others were used to estimate true yields
Fumaric acid	3	O_2/H_2O	0	NH_3	0.323	Linton and Stephenson 1978[f];	
Pyruvate	3.33	O_2/H_2O	0	NH_3	0.446	Verduyn 1991;	0.434 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Pyruvate	3.33	O_2/H_2O	0	NH_3	0.32	Pronk et al. 1990	The original report was 24 gDW/mol. The yields reported were not true yield but net yield; assuming the biomass formula of $C_5H_7O_2N$ and 90% biomass content in dry weight; the same mol-C yields were reported in Heijnen and vanDijken 1992[e]

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Ethylenedianime-tetraacetate (EDTA)	3.4	O_2/H_2O	4	NH_3	0.272	Yuan 2004;	Decay values reported by others were used to estimate true yields
EDTA	3.4	O_2/H_2O	4	NH_3	0.270	Henneken et al. 1998;	
Succinate	3.5	O_2/H_2O	0	NH_3	0.385	Rutgers 1990;	Stdev = 0.014 The extrapolated yield at dilution rate of 0.2/h
Succinate	3.5	O_2/H_2O	0	NH_3	0.449	Verduyn 1991;	0.448 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Succinate	3.5	O_2/H_2O	0	NH_3	0.395	van Verseveld 1979;	The original reported yield was 39.6 (95% CI of 36.3–43.6) gbiomass/mol; 0.48 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Succinate	3.5	O_2/H_2O	0	NH_3	0.4	Heijnen and vanDijken 1992[e];	
Succinate	3.5	O_2/H_2O	0	NH_3	0.41	Andrews 1993;	
Succinate	3.5	O_2/H_2O	0	NH_3	0.344	Linton and Stephenson 1978[f];	
Succinate	3.5	O_2/H_2O	0	NH_3	0.360	Linton and Stephenson 1978[f];	
Succinate	3.5	O_2/H_2O	0	NH_3	0.392	Birou et al. 1987;	
Succinate	3.5	NO_3^-/N_2	0	NH_3	0.380	Boogerd et al. 1984;	The original reported yield was 38.1 g biomass/mol; 0.387 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (the yield reported were not true yield but net yield at dilution rate of 0.15/h; yield calculation was based on the biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Succinate[c]	3.5	NO_3^-/N_2	0	NH_3	0.343	Atkinson and Mavituna 1991;	
Succinate[d]	3.5	NO_3^-/N_2	0	NH_3	0.323	Atkinson and Mavituna 1991;	
Gluconate	3.67	O_2/H_2O	0	NH_3	0.561	Verduyn 1991;	0.559 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Gluconate	3.67	O_2/H_2O	0	NH_3	0.518	van Verseveld 1979;	The original yield was reported as 77.8 (95% CI of 73.8–82.3) gbiomass/mol; 0.51 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e](yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Gluconate	3.67	O_2/H_2O	0	NH_3	0.51	Heijnen and vanDijken 1992[e];	
Gluconate	3.67	NO_3^-/N_2	0	NH_3	0.461	van Verseveld 1979;	The original yield was reported as 69.2 (95% CI of 63.8–75.5) gbiomass/mol; 0.505 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Gluconate[c]	3.67	NO_3^-/N_2	0	NH_3	0.442	Atkinson and Mavituna 1991;	
Gluconate[d]	3.67	NO_3^-/N_2	0	NH_3	0.373	Atkinson and Mavituna 1991;	
Acetate	4	O_2/H_2O	0	NH_3	0.406	Rutgers 1990;	Stdev = 0.006 The extrapolated yield at dilution rate of 0.2/h
Acetate	4	O_2/H_2O	0	NH_3	0.456	Verduyn 1991;	0.455 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Acetate	4	O_2/H_2O	0	NH_3	0.41	Heijnen and vanDijken 1992[e];	
Acetate	4	O_2/H_2O	0	NH_3	0.44	Andrews 1993;	
Acetate	4	O_2/H_2O	0	NH_3	0.389	Linton and Stephenson 1978[f];	
Acetate	4	O_2/H_2O	0	NH_3	0.471	Birou et al. 1987;	
Acetate	4	O_2/H_2O	0	NH_3	0.368	vonStockar and Liu (1999)	
Glucose	4	O_2/H_2O	0	NH_3	0.597	Verduyn 1991;	0.595 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Glucose	4	O_2/H_2O	0	NH_3	0.458	Pronk et al. 1990;	The original true yield was reported as 69 gDW/mol; net yield was reported as 60 gDW/mol; 0.4 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (the yields reported were not true yield but net yield; assuming the biomass formula of $C_5H_7O_2N$ and 90% biomass content in dry weight; the same mol-C yields were reported in Heijnen and vanDijken 1992[e]).

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Glucose	4	O_2/H_2O	0	NH_3	0.61	Heijnen and vanDijken 1992[e];	
Glucose	4	O_2/H_2O	0	NH_3	0.55	Andrews 1993;	
Glucose	4	O_2/H_2O	0	NH_3	0.540	Linton and Stephenson 1978[f];	
Glucose	4	O_2/H_2O	0	NH_3	0.540	Linton and Stephenson 1978[f];	
Glucose	4	O_2/H_2O	0	NH_3	0.458	Linton and Stephenson 1978[f];	
Glucose	4	O_2/H_2O	0	NH_3	0.515	Linton and Stephenson 1978[f];	
Glucose	4	O_2/H_2O	0	NH_3	0.605	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.451	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.482	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.399	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.558	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.539	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.54	Birou et al. 1987;	
Glucose	4	O_2/H_2O	0	NH_3	0.595	vonStockar and Liu 1999	
Glucose	4	O_2/H_2O	0	NH_3	0.557	vonStockar and Liu 1999	
Fructose	4	O_2/H_2O	0	NH_3	0.505	Rutgers 1990;	Stdev = 0.049 The extrapolated yield at dilution rate of 0.2/h
Galactose	4	O_2/H_2O	0	NH_3	0.557	Birou et al. 1987;	
Lactate	4	O_2/H_2O	0	NH_3	0.510	Heijnen and vanDijken 1992[e];	
Lactate	4	O_2/H_2O	0	NH_3	0.48	Andrews 1993;	
Lactate	4	O_2/H_2O	0	NH_3	0.492	Verduyn 1991;	Yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight.
Lactate	4	NO_3^-/N_2	0	NO_3^-	0.274	De Vries et al. 1980	The original reported net yield is 20.2 g/mol at D = 0.24/h; 0.274 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e].
Lactose	4	O_2/H_2O	0	NH_3	0.51	Birou et al. 1987;	
Xylose	4	O_2/H_2O	0	NH_3	0.492	Verduyn 1991;	0.490 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Formaldehyde	4	O_2/H_2O	0	NH_3	0.470	Heijnen and vanDijken 1992[e];	
Benzoate	4.29	O_2/H_2O	2 dioxy	NH_3	0.421	Linton and Stephenson 1978[f];	

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Mannitol	4.33	O_2/H_2O	0	NH_3	0.542	van Verseveld 1979;	Original yield was reported as 81.4 (95% CI of 74.8–89.2) gbiomass/mol; 0.62 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e](yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Mannitol	4.33	O_2/H_2O	0	NH_3	0.56	Heijnen and vanDijken 1992[e];	
Mannitol	4.33	O_2/H_2O	0	NH_3	0.544	Linton and Stephenson 1978[f];	
Mannitol	4.33	NO_3^-/N_2	0	NH_3	0.496	van Verseveld 1979;	Original true yield was reported as 74.5 (95% CI of 71.9–77.3) gbiomass/mol; 0.506 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Sorbitol	4.33	O_2/H_2O	0	NH_3	0.552	Linton and Stephenson 1978[f];	
Phenylacetic acid	4.5	O_2/H_2O	2 dioxy	NH_3	0.475	Linton and Stephenson 1978[f];	
Phenol	4.67	O_2/H_2O	2 dioxy	NH_3	0.358	Reardon et al. 2000;	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.569	Rutgers 1990;	Stdev = 0.019. The extrapolated yield at dilution rate of 0.2/h
Glycerol	4.67	O_2/H_2O	0	NH_3	0.694	Verduyn 1991;	0.692 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Glycerol	4.67	O_2/H_2O	0	NH_3	0.55	Pronk et al. 1990;	Original report was 41 gDW/mol. The yields reported were not true yield but net yield; assuming the biomass formula of $C_5H_7O_2N$ and 90% biomass content in dry weight; the same mol-C yields were reported in Heijnen and vanDijken 1992[e]
Glycerol	4.67	O_2/H_2O	0	NH_3	0.67	Heijnen and vanDijken 1992[e];	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.69	Andrews 1993;	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.579	Linton and Stephenson 1978[f];	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.581	Linton and Stephenson 1978[f];	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.572	Linton and Stephenson 1978[f];	
Glycerol	4.67	O_2/H_2O	0	NH_3	0.666	Birou et al. 1987;	
Propionate	4.67	O_2/H_2O	0	NH_3	0.480	Heijnen and vanDijken 1992[e];	
Benzene	5	O_2/H_2O	2 dioxy	NH_3	0.427	Reardon et al. 2000;	

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Butyric acid	5	O_2/H_2O	0	NH_3	0.523	Verduyn 1991;	Yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight
Ethylenediamine (ED)	5	O_2/H_2O	0	NH_3	0.658	Yuan 2004;	Decay values reported by others were used to estimate true yields
Acetoin	5	O_2/H_2O	0	NH_3	0.421	Verduyn 1991;	0.424 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Toluene	5.14	O_2/H_2O	2 dioxy	NH_3	0.525	Reardon et al. 2000;	
Acetone	5.33	O_2/H_2O	0	NH_3	0.445	Heijnen and vanDijken 1992[e];	
2,3-butanediol	5.5	O_2/H_2O	0	NH_3	0.456	Verduyn 1991;	0.446 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Methanol	6	O_2/H_2O	0	NH_3	0.535	van Verseveld 1979	Original yield was reported as 15 (95% CI of 12.1–15) gbiomass/mol; 0.54 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yield calculation was based on the reported biomass formula of $C_6H_{10.8}O_{2.9}N_{1.5}$).
Methanol	6	O_2/H_2O	0	NH_3	0.54	Heijnen and vanDijken 1992[e];	
Methanol	6	O_2/H_2O	0	NH_3	0.548	Linton and Stephenson 1978[f];	
Methanol	6	O_2/H_2O	0	NH_3	0.589	Linton and Stephenson 1978[f];	
Methanol	6	O_2/H_2O	0	NH_3	0.599	Birou et al. 1987;	
Ethanol	6	O_2/H_2O	0	NH_3	0.558	Rutgers 1990;	Stdev = 0.013. The extrapolated yield at dilution rate of 0.2/h
Ethanol	6	O_2/H_2O	0	NH_3	0.619	Verduyn 1991	0.617 mol-C/mol-C was reported in Heijnen and vanDijken 1992[e] (yields were calculated based on the reported formula of $C_{3.9}H_{6.5}O_{1.93}N_{0.7}$ and 94% biomass content in dry weight).
Ethanol	6	O_2/H_2O	0	NH_3	0.53	Heijnen and vanDijken 1992[e];	
Ethanol	6	O_2/H_2O	0	NH_3	0. 64	Andrews 1993;	
Ethanol	6	O_2/H_2O	0	NH_3	0.532	Linton and Stephenson 1978[f];	
Ethanol	6	O_2/H_2O	0	NH_3	0.703	Birou et al. 1987;	

Table 5 continued

ED & CS	γ_s	EA	t_{oxy} [a]	NS	Y_{exp}^{mb}	Reference	Note
Ethanol	6	O_2/H_2O	0	NH_3	0.496	Birou et al. 1987;	
Ethanol	6	O_2/H_2O	0	NH_3	0.597	vonStockar and Liu 1999	
Propanol	6	O_2/H_2O	0	NH_3	0.575	Heijnen and vanDijken 1992[e];	
n-alkane	6.13	O_2/H_2O	1*	NH_3	0.570	Heijnen and vanDijken 1992[e];	
n-alkane	6.13	O_2/H_2O	1*	NH_3	0.56	Andrews 1993;	
Hexadecane	6.13	O_2/H_2O	1*	NH_3	0.538	Birou et al. 1987;	
Butane	6.5	O_2/H_2O	1	NH_3	0.445	Heijnen and vanDijken 1992[e];	
Propane	6.67	O_2/H_2O	1	NH_3	0.626	Linton and Stephenson 1978[f];	
Ethane	7	O_2/H_2O	1	NH_3	0.622	Linton and Stephenson 1978[f];	
Methane	8	O_2/H_2O	1	NH_3	0.550	Heijnen and vanDijken 1992[e];	
Methane	8	O_2/H_2O	1	NH_3	0.54	Andrews 1993;	
Methane	8	O_2/H_2O	1	NH_3	0.574	Linton and Stephenson 1978[f];	
Methane	8	O_2/H_2O	1	NH_3	0.552	Linton and Stephenson 1978[f];	
Methane	8	O_2/H_2O	1	NH_3	0.597	Linton and Stephenson 1978[f];	
Methane	8	O_2/H_2O	1	NO_3^-	0.506	Heijnen and Roels 1981;	

[a] t_{oxy} refers to the number of oxygenase reactions involved the degradation process of 1 mole ED. Dioxy refers to and dioxygenase reaction; *indicates the number of oxygenation reactions involved in the pathway is not certain

[b] Y_{exp}^{m} is in the unit of (mol-C cell) (mol-C ED)$^{-1}$

[c] Sulfate participates the degradation process but ED is still the limit factor

[d] Sulfate participates the degradation process and limits the growth

[e] Data were compiled from multiple sources given in Table 6 VE

[f] Data were cited from multiple sources

and are often NOT identified in the literature. For example, in pure culture, *P. carbinolicus* can convert organic substrates to acetate, ethanol, and biomass. In mixed culture, more complex product mixtures are observed (Schink 1984). It is necessary to know the carbon products in order to predict bacterial yield with thermodynamics. Therefore, when fermentation products were not well specified in literature reports of yields, these values were not used for subsequent analyses.

Variability in reported true yield data

Reported true or theoretical yields (even on the same substrates (ED, EA, and nitrogen source)) show significant variability. This can be due to routine measurement error, different methods of yield determination, assumptions inherent in calculations of theoretical yield from observed yield, or due to undetermined differences among experimental systems. For example, Rutgers (1990) measured the yields on several substrates and reported the yields with their standard deviation (data included in Table 5 and shown in Fig. 1). The interval of mean ± one standard deviation is 7.50% of the mean value in average. But for formate, this value is 30.86%. Considering the yields on formate measured by others, Table 5 included six experimental yield values from different references. They have a mean of 0.156 with a standard deviation of 0.022 which is 14% of the mean. Four of the reported aerobic values are lower than the reported yields on formate with NO_3^- as EA and NS (0.166 mol-C/mol-C shown in Table 5). This result is not supported by our understanding of electron-donor limited growth and thermodynamics of oxygen and nitrate. Thus, for some compounds like formate, the true yield shows significant variability in the reported literature. The thermodynamic yield prediction model is not expected to capture this variability.

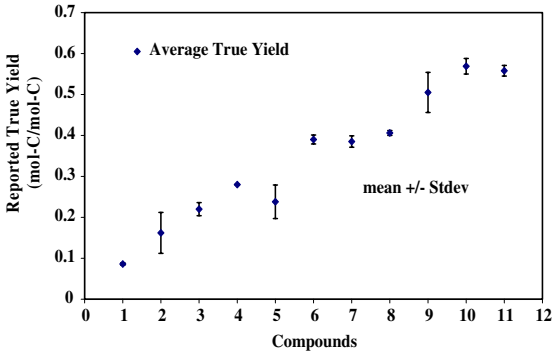

Fig. 1 The experimental yield data reported by Rutgers (1990) with one standard deviation interval. The compounds shown are (1) oxalate, (2) formate, (3) glyoxylate, (4) tartrate, (5) malonate, (6) citrate, (7) succinate, (8) acetate, (9) fructose, (10) glycerol, and (11) ethanol. The intervals of mean ± standard deviation for formate, malonate, and fructose are very large

Uncertainty caused by the method of true yield calculation

True yield (Y^m) can not be measured directly. It is usually calculated based on the measurement of observed yield (Y) and maintenance energy (or bacterial decay) through Roels' equation, $\frac{1}{Y} = \frac{1}{Y^m} + \frac{m}{\mu}$, where, m is maintenance energy with the dimension of [substrate]/[cell].[time] and μ is specific growth rate of bacteria with the dimension of 1/[time]. Usually, a series of continuous culture experiments are run under different growth rates (μ) and the observed yields (Y) is determined under steady state at each growth rate. Then, Y^m and maintenance energy (m) are computed by a linear regression of $\frac{1}{Y}$ and $\frac{1}{\mu}$ (Roels 1983). Therefore, the calculation of Y^m rests on the validity of Roel's equation, which has been called into question (Bally et al. 1994; Henneken et al. 1998). Further, maintenance energy sometimes is not measured during yield determination experiments, especially when batch culture is used for determining the bacterial growth. Instead, the value reported by others is often used to estimate the true yield under this circumstance. For example, Yuan (2004) used the decay value reported by Henneken et al. (1998) to calculate the true yield of *BNC1/BNC2* on EDTA. However, maintenance energy is strongly related to experimental conditions. The reported maintenance energies shown in Heijnen and Roels (1981) vary even when two systems have the same substrates and species. Thus, using literature values for maintenance energy to calculate true yields from observed yields can lead to inaccurate reports of true yield in the literature.

Uncertainty caused by unclear reporting

True yield is needed to validate thermodynamic models because net yield is not predicted by these models. However, observed yields are sometimes used for model generation and demonstration without being adjusted by maintenance energy or decay. For example, the yields measured by Verduyn (1991) were observed yields but were reported as true yields in Heijnen and van Dijken (1992). In addition, yield data are occasionally misreported or reported in secondary sources without details necessary to interpret the values.[1] To avoid these errors, all data presented in Table 5 were verified in the primary source unless noted in the Table footnotes.

Uncertainty caused by experimental measurement and unit conversion

Experimental measurement of yield values shows inherent sample variability. Various methods are used to monitor bacterial growth, such as carbon measurement, protein measurement, dry weight measurement, and optical density (Atkinson and Mavituna 1991). Then, bacterial yield is calculated on the basis of certain assumed molecular formula or protein content. The variety of the methods and assumptions introduces uncertainty into the reported yield values. In addition, often 100% of the carbon cannot be recovered during measurement. For example, the recovered carbon as cells and carbon dioxide in Hernandez and Johnson (1967) varied from 67.6% to 93.8%. The low carbon percentage in cell and CO_2 might suggest incomplete degradation or generation of soluble microbial products. The carbon recovery variability reported in Schink's experiments (Schink 1984) (from 98.3% to 108.1%) could result in up to 0.08 mol-C/mol-C difference for yield values. This recovery rate error might not significantly affect the yields of aerobic microorganisms as they are usually quite high on a carbon basis. However, the reported yields in Schink's paper were mostly lower than 0.1 mol-C/mol-C. Consequently, for bacteria with relatively low yields, measurement error is not negligible and thus seriously affects the evaluation of the prediction accuracy of any model. This is one of the reasons that the estimation error of thermodynamic models is usually larger for anaerobic and autotrophic organisms than aerobic organisms.

The mol-C unit is often used for biomass content. However, bacterial yields are sometimes reported as gram dry weight (gDW)/mol-C substrate and the chemical formula of the bacteria is not specified. Then, the molecular weight (e.g., 26.4 g/mol biomass based on the assumed formula of $CH_2O_{0.6}N_{0.2}$) is used to convert the unit to mol-C/mol-C. However, as mentioned above, the biomass composition varies with species and growth conditions and the corresponding molecular weight is reported from 19 g/mol-C to 26.8 g/mol-C. For example, the yields of *C. utilis* and *S. cerevisiae* on glucose were both reported as 0.51 gDW/g glucose (Verduyn et al. 1991). The biomass formulas and the content of C, H, O, N in the biomass were measured respectively, shown in Table 6. Converting units for the observed yields results in 0.597 mol-C/mol-C and 0.611 mol-C/mol-C, respectively. However, only one value, 0.595 mol-C/mol-C is commonly reported (see for example, Heijnen and van Dijken 1992 Table VB). Further, if the formula of $CH_2O_{0.6}N_{0.2}$ and 90% organic content are assumed, the yield is converted to 0.522 mol-C/mol-C. Thus, a 12.6–14.6% difference in the theoretical yield that might be reported arises from conversion of the units on the reported yield.

Results and discussion

For comparison, Tables 7, 8, and 9 present reported experimental yields, predicted yield, and an error based on comparing the reported yield and the average of the predicted yield range. The average value is used for the reported yield when more than one yield is reported for the same system (defined as the systems with the same ED, EA, NS, and limiting factor). The predicted yield shows a range because of the variability of the K value.

Overall, the average estimation error of the aerobic dataset in Table 7 is 10.3 ± 4.4% (95% CI) and the average of the absolute error is 13.1 ± 3.8% (95% CI). Bold text is used to highlight yield prediction values

[1] For example, 0.1 mol-C/mol-C was shown as the yield on formate in the Table VD in Heijnen and van Dijken (1992). However, the original source (Pronk et al. 1990) reported no growth on formate during their experiments.

Table 6 The effect of biomass formula on reported bacterial yield

Formula	Content	Yield	Unit	Note
		0.51	$(gDW)(g)^{-1}$	Yields of *C. utilis* and *S. cerevisiae* on glucose original reported by Verduyn et al (1991)
$C_{3.9}H_{6.5}O_{1.93}N^{a}_{0.7}$	94%	0.597	$(mol\text{-}C)(mol\text{-}C)^{-1}$	*C. utilis* formula reported in Verduyn et al (1991)
$C_4H_{6.5}O_{1.94}N^{a}_{0.61}$	94%	0.611	$(mol\text{-}C)(mol\text{-}C)^{-1}$	*S. cerevisiae* formula reported in Verduyn et al (1991)
$CH_2O_{0.6}N_{0.2}$	90%	0.522	$(mol\text{-}C)(mol\text{-}C)^{-1}$	Formula suggested by VanBriesen (2000)
$C_5H_7O_2N$	90%	0.609	$(mol\text{-}C)(mol\text{-}C)^{-1}$	Formula suggested by Rittmann and McCarty (2001)
unknown	unknown	0.595	$(mol\text{-}C)(mol\text{-}C)^{-1}$	Value converted by Heinjen and van Dijken (1992)

[a] The formulae were obtained through elemental analysis. The content of oxygen was based on the assumption that C, H, O, and N make up 94% of the dry weight

that appear to be significantly different from observed yields. Reasons for these discrepancies are discussed below.

Incorporating the adjustment of oxygenase reaction, the expanded model gives very accurate true yield prediction for substrates requiring reactions catalyzed by oxygenations and oxidases (e.g., methane, EDTA, and NTA). However, the incorporation of this refinement relaxes the requirement of a black box model. The information about the numbers and types of specialized reactions is needed to predict yields. Unfortunately, this information is not always available. For example, for n-alkane, although we expect monooxygenase reactions (Hamamura et al. 1999), there have no reports of the specific pathway. Assuming one oxygenation per mol substrate during estimation of yield on *n*-alkane still leads to predictions higher than reported values. This may suggest more than one oxygenase reaction or other special reaction in the pathway; however, experimental validation is needed.

Table 8 shows yield predictions for anaerobic systems where products were specified. The average estimation error is −17% and the average of absolute error is 22%. Bold text is used to highlight the substrates where the observed yield is not within the 95% confidence interval of the predicted yield. Table 9 includes prediction for a selection of autotrophic microorganisms; reported yields were only considered if sufficient detail was provided in the original source to ascertain the nitrogen source and the likely bacterial formula. There were many reported yields for autotrophs that are not shown here as we did not consider them suitable for model evaluation. The average estimation error is 8.3% and the absolute estimation error for the autotrophic dataset is 8.3%. The high estimation error of anaerobic systems partly results from the various products during anaerobic processes. For autotrophic systems, in order to get electrons from the inorganic electron donors, some special reactions might be required. For example, ammonia monooxygenase is the enzyme involving in the first step of ammonia oxidation (Arp et al. 2002). Kappler et al. (2001) reported that monooxygenase reaction was involved during the degradation of thiosulfate. The tetrathionate has a similar structure to thiosulfate. So, 2 monooxygenase reactions and 1.25 times of monooxygenase are assumed for the substrates of tetrathionate and thiosulfate, respectively, in model predictions. Further, some autotrophs can utilize insoluble gases as substrates (e.g., hydrogen and carbon monoxide). The partial pressure of those gases is not commonly reported and this limits the application of the thermodynamic model in these specific systems.

The estimated yields, Y^{m}_{est}, and the experimental yields, Y^{m}_{exp}, reported in Tables 7–9 are compared in Fig. 2. The regression line is expressed as $Y^{m}_{est} = 1.05Y^{m}_{exp}$ (the intercept is set to 0) with R^2 of 0.88. The slope is very close to 1 which represents the perfect estimation. This suggests the expanded model has strong prediction ability for different types of microorganisms under different conditions.

While the results in Tables 7–9 and Fig. 2 are encouraging, the significant variability in reported yields argues for caution and another approach to evaluation of the model results. Thus, we consider whether the *predicted* yield values could reasonably be assumed to come from the distribution of values assumed to represent the true yield. We visually represent this with Fig. 3, which shows the average

Table 7 Comparison between the estimation results of the expanded thermodynamic model and the average experimental data for heterotrophic aerobic microorganisms (pH 7 is assumed for estimation)

ED & CS	EA	t_{oxy}	Average Y^{m}_{exp}	Y^{m}_{est}	Error based on average
Oxalate	**O_2/H_2O**	**0**	**0.072**	**0.107 ± 0.021**	**47.93%**
Formate	**O_2/H_2O**	**0**	**0.156**	**0.216 ± 0.031**	**38.17%**
Formate	NO_3^-/N_2	0	0.166	0.197 ± 0.039	18.67%
Glyoxylate	O_2/H_2O	0	0.220	0.247 ± 0.042	12.27%
Tartrate	O_2/H_2O	0	0.280	0.297 ± 0.049	6.07%
Malonate	O_2/H_2O	0	0.238	0.268 ± 0.048	12.61%
Nitrilotriacetate (NTA)	O_2/H_2O	1	0.272	0.287 ± 0.048	5.51%
Iminodiacetate (IDA) through oxygenation	O_2/H_2O	1	0.240	0.258 ± 0.043	7.50%
Iminodiacetate (IDA)	O_2/H_2O	0	0.333	0.337 ± 0.056	1.20%
Citrate	O_2/H_2O	0	0.374	0.334 ± 0.056	−10.77%
Malate	O_2/H_2O	0	0.342	0.342 ± 0.056	−0.00%
Glycine	**O_2/H_2O**	**0**	**0.248**	**0.354 ± 0.057**	**42.74%**
Fumaric acid	O_2/H_2O	0	0.323	0.345 ± 0.056	6.81%
Pyruvate	O_2/H_2O	0	0.377	0.397 ± 0.062	5.31%
Ethylenedianime-tetraacetate (EDTA)	O_2/H_2O	2 mono and 2 oxidase	0.271	0.300 ± 0.048	10.70%
Succinate	O_2/H_2O	0	0.392	0.383 ± 0.063	−2.26%
Succinate	NO_3^-/N_2	0	0.380	0.37 ± 0.063	−2.63%
Succinate	NO_3^-/N_2	0	0.343	0.37 ± 0.063	7.87%
Succinate	NO_3^-/N_2	0	0.323	0.37 ± 0.063	14.55%
Gluconate	O_2/H_2O	0	0.530	0.464 ± 0.073	−12.40%
Gluconate	NO_3^-/N_2	0	0.461	0.451 ± 0.073	−2.17%
Gluconate	NO_3^-/N_2	0	0.442	0.451 ± 0.073	2.04%
Gluconate	NO_3^-/N_2	0	0.373	0.451 ± 0.073	20.91%
Acetate	O_2/H_2O	0	0.420	0.446 ± 0.07	6.19%
Glucose	O_2/H_2O	0	0.526	0.501 ± 0.079	−4.69%
Fructose	O_2/H_2O	0	0.505	0.501 ± 0.079	−0.79%
Galactose	O_2/H_2O	0	0.557	0.500 ± 0.079	−10.23%
Lactate	O_2/H_2O	0	0.494	0.48 ± 0.075	−2.83%
Lactate	**NO_3^-/N_2**	**0**	**0.274**	**0.436 ± 0.076**	**59.12%**
Lactose	O_2/H_2O	0	0.510	0.508 ± 0.080	−0.39%
Xylose	O_2/H_2O	0	0.492	0.515 ± 0.081	4.67%
Formaldehyde	O_2/H_2O	0	0.470	0.524 ± 0.082	11.49%
Benzoate	O_2/H_2O	2 dioxy	0.421	0.399 ± 0.063	−5.23%
Mannitol	O_2/H_2O	0	0.549	0.538 ± 0.084	−1.94%
Mannitol	NO_3^-/N_2	0	0.496	0.523 ± 0.084	5.44%
Sorbitol	O_2/H_2O	0	0.552	0.538 ± 0.085	−2.54%
Phenylacetic acid	O_2/H_2O	2 dioxy	0.475	0.489 ± 0.067	2.95%
Phenol	O_2/H_2O	2 dioxy	0.358	0.425 ± 0.069	18.72%
Glycerol	O_2/H_2O	0	0.619	0.578 ± 0.091	−6.62%
Propionate	O_2/H_2O	0	0.480	0.521 ± 0.082	8.54%
Benzene	O_2/H_2O	2 dioxy	0.427	0.465 ± 0.074	8.90%

Table 7 continued

ED & CS	EA	t_{oxy}	Average Y^{m}_{exp}	Y^{m}_{est}	Error based on average
Butyric acid	O_2/H_2O	0	0.523	0.569 ± 0.089	8.80%
Ethylenediamine (ED)	O_2/H_2O	0	0.658	0.616 ± 0.097	−6.38%
Acetoin	**O_2/H_2O**	**0**	**0.421**	**0.584 ± 0.092**	**38.72%**
Toluene	O_2/H_2O	2 dioxy	0.525	0.492 ± 0.078	−6.29%
Acetone	**O_2/H_2O**	**0**	**0.445**	**0.606 ± 0.095**	**36.18%**
2,3-butanediol	**O_2/H_2O**	**0**	**0.456**	**0.65 ± 0.102**	**42.54%**
Methanol	**O_2/H_2O**	**0**	**0.562**	**0.728 ± 0.114**	**29.49%**
Ethanol	**O_2/H_2O**	**0**	**0.576**	**0.692 ± 0.109**	**20.05%**
Propanol	O_2/H_2O	0	0.575	0.682 ± 0.107	18.61%
n-alkane	O_2/H_2O	1	0.565	0.657 ± 0.104	16.28%
Hexadecane	**O_2/H_2O**	**1**	**0.538**	**0.661 ± 0.104**	**22.86%**
Butane	O_2/H_2O	1	0.445	0.638 ± 0.102	43.37%
Propane	O_2/H_2O	1	0.626	0.634 ± 0.100	1.28%
Ethane	O_2/H_2O	1	0.622	0.615 ± 0.097	−1.13%
Methane	O_2/H_2O	1	0.563	0.552 ± 0.087	−1.88%
Methane	O_2/H_2O	1	0.506	0.51 ± 0.076	0.79%
The average estimation error	10.3 ± 4.4% (95% CI)				
The average of absolute estimation error	13.1 ± 3.8% (95% CI)				

reported yield value as well as the interval of mean ± one standard deviation. Only eleven compounds with at least 4 reported yield values are shown. For example, the formate data have an average of 0.156 and a standard deviation of 0.022; thus the measurement interval is 0.156 ± 0.022. The estimated yield by the expanded thermodynamic model (0.216 mol-C/mol-C) is *not* within this range. For malate, the estimate (0.342) is equal to the average of the reported values (0.342 ± 0.071) but the standard deviation is high, 21% of the average. From Fig. 3, it is obvious that 7 out of 11 predicted yield values are within the one standard deviation interval of the observed yields. This suggests that the estimates on formate, citrate, methanol, and ethanol are distinctly different from their reported yields. It also suggests that the estimation error is not an ideal determinant for predictive power of the model due to the significant variability in the observed values. However, the method described here cannot be used more broadly to evaluate predictive yield models due to the paucity of experimental data for substrates.

When sufficient experimental data are not available to create a range (average ± one standard deviation) for the observed yield, it is difficult to assess the prediction ability of the model. However, one way to evaluate this is to calculate a predicted yield range (based on 95% confidence interval of K) and determine if the observed yield falls in that range. This is how the prediction range shown in Table 7 was computed. As mentioned earlier, the bold text identifies compounds with observed yield outside the prediction range. If we presume the thermodynamic model framework is accurate, the significant over estimation of the yields of these compounds suggests either the experimental data are not accurate or model adjustment is needed. For some volatile compounds such as methanol and ethanol, the substrates are easily lost to the atmosphere, especially under strong aeration. This might be a reason causing the higher estimation of our model for those compounds. Alternatively, some special reactions with unusual energy consumption (like oxygenase reaction) might be involved in the degradation pathways for these

Table 8 The estimation results for some heterotrophic anaerobic and aerobic product-forming microorganisms[a]

ED	EA/product	γ_{ED}	Y^m_{exp}	Y^m_{est}	Estimation error
Formate	Formate/CH_4	2	0.053	0.047 ± 0.015	−11.13%
Acetate	Acetate/CH_4	4	0.024	0.029 ± 0.008	+20.83%
Glucose	**Glucose/CH_4**	**4**	**0.25**	**0.13 ± 0.037**	**−47.88%**
Methanol	Methanol/CH_4	6	0.13	0.152 ± 0.043	+16.92%
Methanol	**Methanol/CH_4**	**6**	**0.3**	**0.152 ± 0.043**	**−49.33%**
Lactate[b]	Lactate/acetate and propionate	4	0.085	0.08 ± 0.012	−6.24%
Acetoin[b]	**Acetoin/acetate and propionate**	**5**	**0.08**	**0.06 ± 0.009**	**−25.43%**
2,3-butanediol[b]	**2,3- butanediol/acetate and propionate**	**5.5**	**0.063**	**0.04 ± 0.006**	**−36.26%**
Acetoin[a]	**Acetoin/acetate and ethanol**	**6**	**0.07**	**0.059 ± 0.009**	**−16.07%**
Glucose	O_2/H_2O and glucose/0 mol-C ethanol	4	0.57	0.501 ± 0.079	−12.11%
Glucose	**O_2/H_2O and glucose/0.082 mol-C ethanol**	**4**	**0.52**	**0.444 ± 0.07**	**−14.57%**
Glucose	**O_2/H_2O and glucose/0.228 mol-C ethanol**	**4**	**0.4**	**0.343 ± 0.054**	**−14.19%**
Glucose	O_2/H_2O and glucose/0.44 mol-C ethanol	4	0.23	0.200 ± 0.031	−14.56%
Glucose	**O_2/H_2O and glucose/0.512 mol-C ethanol**	**4**	**0.19**	**0.147 ± 0.023**	**−22.79%**
Glucose	**O_2/H_2O and glucose/0.566 mol-C ethanol**	**4**	**0.14**	**0.109 ± 0.017**	**−21.91%**

[a] pH is assumed as 7.2 for methanogens and as 7 for others. All the experimental yield data are cited from Heijnen & vanDijken 1992 The bold data refer to the predicted yield ranges that do not encompass the observed yield

[b] The information about products is from Schink 1984

compounds and a similar adjustment as oxygenase reaction might be necessary for improving the model prediction. As shown in Table 7, oxalate, formate, glycine, lactate with NO_3^-/N_2 as EA, Acetoin, Acetone, 2,3-butanediol, methanol, and ethanol have experimental yields out of the prediction ranges. Ensign et al. (1998) found CO_2 plays a very important role during microbial metabolism of aliphatic ketones including acetone. The step of inserting CO_2 needs energy which is provided by the hydrolysis process of adenosine triphosphate (ATP) to adenosine monophosphate (AMP). With the consideration of the energy loss during this step, the yield estimation on acetone would be very close to the

Table 9 The yield of some autotrophic microorganisms[a]

ED (partial pressure, atm)	EA[b]	γ_{ED}	Y^m_{exp} [c]	Y^m_{exp} [c]	$\Delta G_{e\text{-}ED}$	$\Delta G_{e\text{-}EA}$ (1)	t_{oxy}	Y^m_{est}	Estimation error
NH_4^+/NO_2^-	O_2/H_2O, CO_2/cell	6	0.06	0.06	32.94	−78.685	1	0.062 ± 0.014	+3.33%
$S_2O_3^{2-}/SO_4^{2-}$	O_2/H_2O, CO_2/cell	8	0.16	0.193	−23.50	−78.685	1.25	0.209 ± 0.047	+8.57%[e]
$S_2O_3^{2-}/SO_4^{2-}$	O_2/H_2O, CO_2/cell	8	0.16						
$S_2O_3^{2-}/SO_4^{2-}$	O_2/H_2O, CO_2/cell	8	0.22						
$S_2O_3^{2-}/SO_4^{2-}$	O_2/H_2O, CO_2/cell	8	0.23						
$S_4O_6^{2-}/SO_4^{2-}$	O_2/H_2O, CO_2/cell	14	0.410	0.41	−27.18	−78.685	2	0.432 ± 0.098	+5.37%
CO (1)	CO/CH_4, CO/cell	2	0.11	0.11	−55.66[d]	−78.685	0	0.121 ± 0.035	+10.00%
CO (1)	O_2/H_2O, CO_2/cell	2	0.16	0.16	−55.66	−78.685	0	0.183 ± 0.041	+14.38%

[a] All the data are cited from Heijnen and vanDijken (1992)

[b] ED and EA are written as the form of ED/ED product or EA/EA product

[c] Y_{exp}^m and Y_{est}^m are in the unit of (mol-C)(mol-ED)$^{-1}$

[d] The Gibbs energy of formation of CO (aq), ΔG_f (CO_{aq}) is estimated based on ΔG_f (CO_g) and its Henry's constant, 1,050 atm.l/mol (Benjamin 2002)

[e] The estimation error was computed according to the average of the four experimental yield data on $S_2O_3^{2-}$

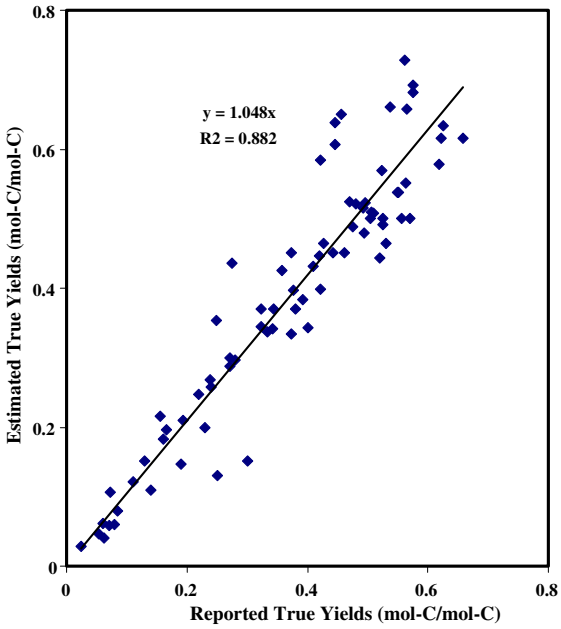

Fig. 2 The comparison between the reported yields (Y^{m}_{exp}) and the estimated yields (Y^{m}_{est}) The horizontal axis represents Y^{m}_{exp} and the vertical axis represents Y^{m}_{est}. The regression line of the data is also shown in this figure and the regression equation is $Y^{m}_{est} = 1.048 Y^{m}_{exp}$ with R square of 0.882 by setting intercept as 0

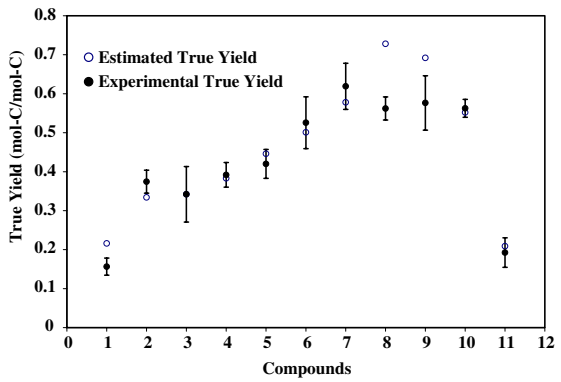

Fig. 3 The experimental yields on 11 different compounds with one standard deviation interval. The open circle represents the average of the estimated yields, the solid dot represents the experimental true yield and the standard deviation is calculated based on the reported data. The compounds 1–11 are (1) formate, (2) citrate, (3) malate, (4) succinate, (5) acetate, (6) glucose, (7) glycerol, (8) methanol, (9) ethanol, (10) methane, and (11) thiosulfate. Compounds with less than 4 reported yields are not shown in this figure

reported value. For acetoin, no report on its degradation pathway was found. However, the first degradation step of methyl ethyl ketone, which also contains 2-ketone but no 3-hydroxy, is an oxidase reaction (Eubanks et al. 1974). As discussed above, oxidase reactions can divert electrons away from energy-generating pathways. Based on the structures, we hypothesize the degradation pathway of acetoin and 2,3-butanediol might also involve an oxidase reaction but experimental evidence is needed to support these suppositions.

Two possible pathways are indicated in Kyoto Encyclopedia of Genes and Genomes (KEGG) database for glycine: one includes glycine dehydrogenase, which acts on the CH–NH_2 group of glycine with a cytochrome as acceptor (Sanders et al. 1972); the other one includes D-amino-acid oxidase which also acts on the CH–NH_2 group but with oxygen as acceptor (Dixon and Kleppe 1965). Although the estimate is much closer to the experimental yield of glycine presented in Table 5 if we consider the oxidase pathway, the enzyme was not specified in the experiment that determined the yield value.

From Table 7, we also notice that the expanded model always overestimates the yields on one or two carbon substrates, such as formate, methanol, oxalate, and ethanol. The explanation might be that more energy is needed to lengthen the carbon chain during anabolism because those compounds are simpler than cell components. For example, CO_2 is always used as carboxyl donor for single carbon substrate like methanol. Bainotti and Nishio (2000) found that *Acetobacterium sp*. has higher energy yield if formate replaces CO_2 as carboxyl donor. This suggests extra energy cost for use of some simple carbon compounds.

This discussion highlights one of the critical limitations of all thermodynamically based yield prediction methods. The objective is to develop a model that allows yield prediction from the thermodynamics of the substrates alone, without knowledge of their biodegradation pathways. However, as shown by VanBriesen (2001) and Yuan and VanBriesen (2002), and confirmed in the present work, yields for compounds that involve oxygenase-catalyzed reactions cannot be predicted without consideration of the energetics of these reactions. Further, McCarty (2007) provides an alternative framework to the carbon-balance construction of Xiao and VanBriesen (2006) that is based on knowledge of the mechanism of carbon conversion in the cell. These modifications clearly require knowledge of the catabolic reactions for the target substrate. While for some substrates

these are well known, for most anthropogenic compounds whose yield we wish to predict, these aspects will be more difficult to ascertain. Thus, thermodynamically-based yield predictions for novel compounds must always be used with acknowledgment of their limitations.

Conclusions

This work includes refinement of the thermodynamic model described by Xiao and VanBriesen (2006) to account for additional information of environmental conditions (specifically pH) and the utilization of investment reactions in the pathway of degradation. Comparison with reported yields demonstrates that the expanded thermodynamic true yield prediction model has strong ability to predict the bacterial yields under different growth conditions (aerobic heterotrophs, anaerobic and anoxic heterotrophs, and lithotrophs). We also analyzed the uncertainty of the experimental data for reported true yield and the estimation data used in the model. The uncertainty of the current true yield data suggests models based on fitting of parameters to experimental data will introduce significant uncertainty and hence confirms the advantage of the expanded true yield prediction model.

The uncertainty in the reported true yield values is a significant hindrance to further improvement of the thermodynamic true yield prediction models. Some uncertainty cannot be overcome, for example, the systematic error of measurement. However, the uncertainty due to the conversion from observed yield to true yield can be reduced through a more complete understanding of bacterial decay or maintenance energy. Since the decay or maintenance constant is needed to interpret observed yield values or to modify thermodynamically-predicted true yields for comparison with reported yields, methods to predict this constant are urgently needed. Ultimately, a thermodynamic model that enables *a priori* prediction of OBSERVED yield rather than theoretical yield is required.

The uncertainty due to unit conversion in reported yields can be reduced by increasing the transparency of data analysis in papers reporting yields. It is common practice to report a yield in mass units; however, within the thermodynamic literature secondary citations of these values often state them in molar carbon units without clear descriptions of the assumptions inherent in these conversions. Improved reporting or evaluations based only on primary literature data is recommended.

We believe the greatest concern within the thermodynamic prediction model structure is the dependence on values for energetic efficiency. As discussed by McCarty (2007) these values are usually estimated from literature reported yield values that are limited by the errors we discuss in this paper. McCarty (2007) shows improved predictions over his original model structure (McCarty, 1965) predominately because of a re-evaluation of the thermodynamic efficiency term. Even the *a priori* method of Xiao and VanBriesen (2006), which avoids the problem of estimating K from reported yield values, has limitations due to the uncertainty in the measured values used to derive K. Further, it has been widely hypothesized that bacterial energy capture efficiency is NOT a constant, but rather is controlled by environmental conditions. If this is true, the thermodynamic method for bacterial yield prediction will remain limited to cases when optimal conditions ensure predictable, high energy capture efficiencies. The next great challenge for bacterial yield prediction with thermodynamics is to understand how environmental conditions affect energy capture efficiency.

Acknowledgement The authors acknowledge the support of the National Science Foundation, under grant number BES-0219330.

Appendix A: Estimation of the usable energy released during oxygenase reaction

Oxygenase reactions are utilized by organisms to create more biodegradable forms of substrates. Some examples are when alkanes are hydroxylated, alkenes are converted into the corresponding epoxides (Peters and Witholt 1994), CO is oxidized to CO_2 (Timkovich and Thrasher 1988), ammonia is oxidized to

hydroxylamine (Arp et al. 2002), and some aromatic compounds and cyclic alkanes are hydroxylated (Enroth et al. 1998; Kauppi et al. 1998). Based on the enzyme participating in the reaction, oxygenase reactions can be divided into two groups: monooxygenase-catalyzed reactions and dioxygenase-catalyzed reactions. Different conceptual models and energy calculation are needed for them individually.

Monooxygenase reaction

Monooxygenase enzymes catalyze the NADPH- (or NADH-) and oxygen- dependent oxidation of a wide range of chemicals. (Fosdike et al. 2005; Demirdogen and Adali, 2005) When 1 mole of molecular oxygen is reduced by a monooxygenase enzyme, 4 electron equivalents of electrons are needed: 2 e^- equivalents are from the oxidation of substrate and the other 2 e^- equivalents are invested by NADPH (or NADH). In order to keep NADH at a constant level within the cell, 2 e^- equivalents of electrons released during the further oxidation of substrate are used to regenerate the NADH from NAD^+. Thus, in total, 4 moles of electrons released from the substrate are utilized for each mole of substrate transformed by the monooxygenation reaction. The process of oxygenase reaction can be simplified as Fig. 4.

Because the oxidation half reactions (of NADH and substrate) and the reduction half reaction (of oxygen as co-substrate) are coupled directly, the electrons involved do not pass through other coenzymes such as Flavoprotein or Cytochrome c. Thus, the electron flow can not enhance the hydrogen gradient between the

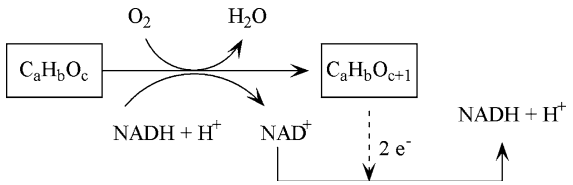

Fig. 4 Conceptual model of monooxygenase reaction (During one monooxygenase reaction, the reduction of 1 mole of molecular oxygen, O_2, is directly coupled with the oxidation of substrate and NADH but only 1 mol oxygen atom, [O], is inserted into the substrate. In total, 4 e^- equivalents of electrons are required where 2 e^- eq. are from substrate and 2 e^- eq. are from NADH. But 2 e^- eq. released from the further oxidation of substrate have to recoup to NAD^+ in order to keep the concentration of NADH in microorganisms constant.)

inner and outer membrane that is the mechanism of ATP formation. Energy produced during monooxygenase reactions is dissipated rather than stored; thus it is unavailable for cell synthesis. However, a small amount of energy is stored in the product, $C_aH_bO_{c+1}$ (the energy carried per electron in the substrate is increased after inserting oxygen). This energy is symbolized as $\Delta G_{\text{monooxy}}$ in the unit of kJ/mol or $\Delta G_{\text{e-monooxy}}$ in kJ/e^- eq. Therefore, during 1 mole of monooxygenase reactions, the electrons accepted by O_2 should be $g(\text{i}) = 4$ e^- equivalents and the corresponding energy released is $\Delta G(i) = \Delta G_{\text{monooxy}}$ or $g(i) \times \Delta G_{\text{e-monooxy}}$ kJ.

If several monooxygenase reactions occur in the degradation pathway and the direct products of each time are known, we can calculate $\Delta G_{e-C_aH_bO_{c+1}}$ to replace $\Delta G_{e-C_aH_bO_c}$ directly (e.g., we use ΔG_{e-CH_4O} to replace ΔG_{e-CH_4} in the methane example). Further, we modify the degree of reductance for the substrate to $\gamma_{C_aH_bO_c} - \frac{4}{a}$ where a is the carbon number in the substrate molecule. If the $\Delta G_{\text{monooxy}}$ (in kJ/mol monooxygen reaction) is not known, it can be estimated based on the group Gibbs formation energy estimation (Mavrovouniotis 1990, 1991). Appendix A provides details and an example of this type of estimation. When applying the expanded model, the oxygen involved in the monooxygenase reaction is considered a sink for electrons (an electron acceptor). We term this as EA(O_2− monooxy) and the corresponding energy released is $\Delta G_{\text{e-EA}}(O_2-\text{monooxy}) = \Delta G_{\text{monooxy}}/4$ kJ/e^- eq, where 4 is the amount of electrons obtained by 1 mol monooxygen reaction. The electrons flowing into this EA are $t_{\text{monooxy}} \times 4$ e^- eq/mol-C, where t_{monooxy} refers to the number of oxygeanse reactions during the degradation of 1 mol-C substrate.

The most common monooxygenase reaction happens on functional group of –CH_3. Here we use this common case as an example to explain the estimation method based on Mavrovouniotis (1991).

Generally, 1 mol oxygen atoms are added into the substrate and group –CH_3 becomes group –CH_2OH after one mol times of monooxygenase reaction. Correspondingly, the substrate (denoted as $C_aH_{bOc}N_d^{n+}$) changes to $C_aH_bO_{c+1}N_d^{n+}$. The half reactions of $C_aH_bO_cN_d^{n+}$ and $C_aH_bO_{c+1}N_d^{n+}$ to CO_2 are shown as Eqs. (3) and (4), respectively. The reductance degree of carbon in the substrate reduces from γ_s to (γ_s − 2/

a), but the coefficients of carbonate species, α1, α2 and α3, do not change as long as pH does not change.

$$\begin{aligned}&C_aH_bO_cN_d^{n+} + (3a - c)H_2O = \\ &(a \times \alpha 1)H_2CO_3 + (a \times \alpha 2)HCO_3^- + (a \times \alpha 3)CO_3^{2-} \\ &+ dNH_4^+ + (a \times (6 - 2\alpha 1 - \alpha 2) \\ &+ b - 4d - 2c)H^+ + (\gamma_s \times a)e^-\end{aligned} \tag{3}$$

$$\begin{aligned}&C_aH_bO_{c+1}N_d^{n+} + (3a - c - 1)H_2O = \\ &(a \times \alpha 1)H_2CO_3 + (a \times \alpha 2)HCO_3^- + (a \times \alpha 3)CO_3^{2-} \\ &+ dNH_4^+ + (a \times (6 - 2\alpha 1 - \alpha 2) + b - 4d \\ &- 2(c + 1))H^+ + (\gamma_s \times a)e^-\end{aligned} \tag{4}$$

Therefore, $\Delta G_{monooxy}$ can be calculated as Eq. (5), where $\Delta G_f(H_2O)$ is −237.18 kJ/mol (Benjamin 2002) and the change of standard Gibbs energy of formation between $C_aH_bO_cN_d^{n+}$ and $C_aH_bO_cN_d^{n+}$ can be estimated based on their molecular structure, $\Delta G_f(C_aH_bO_cN_d^{n+} - C_aH_bO_{c+1}N_d^{n+}) = 7.9 - (-29.3) - 1.7 = 35.5$ kcal/mol , i.e., 148.39 kJ/mol, where 7.9, −29.3 and 1.7 kcal/mol is the estimated Gibbs energy of formation of group –CH_3, –OH and –CH_2, respectively. (Mavrovouniotis 1991)

$$\begin{aligned}\Delta G_{monooxy} &= \Delta G_{C_aH_bO_{c+1}N_d^{n+}} - \Delta G_{C_aH_bO_cN_d^{n+}} \\ &= \Delta G_f(C_aH_bO_cN_d^{n+} - C_aH_bO_{c+1}N_d^{n+}) \\ &\quad + \Delta G_f(H_2O) - 2RT\ln[H^+]\end{aligned} \tag{5}$$

Thus, $\Delta G_{monooxy} = -88.79 - 2 \times (-39.87) = -9.05$ kJ/mol monooxygenase reaction at pH 7 or $\Delta G_{e\text{-}EA}(O_2\text{−monooxy}) = \Delta G_{monooxy}/4 = -2.263$ kJ/e^- eq. Then, the energy of $C_aH_bO_{c+1}N_d^{n+}$ carried per electron can be estimated as Eq. (6):

$$\Delta G_e = \frac{\Delta G_e^0 \times \gamma_s \times a - 9.05}{\gamma_s \times a - 2} \tag{6}$$

where, $\Delta\ Ge^0$ is the energy carried per electron in $C_aH_bO_cN_d^{n+}$. Some estimation examples are shown in Table 10 and the results suggest this estimation method won't introduce distinct error to $\Delta G_e(C_aH_bO_{c+1}N_d^{n+})$.

Dioxygenase reaction

Dioxygenase-catalyzed reactions always happen on the double bond either in aromatic compounds or in alkenes. It is very similar to a monooxygenase reaction. The major difference is both oxygen atoms of O_2 are inserted into the substrate after reaction, shown in Fig. 5. After one mole of substrate is converted via the dioxygenase reaction, the substrate loses 4 moles of electrons and two –H bonds in the substrate are replaced by two –OH bonds. The gained energy, termed as ΔG_{dioxy} in the unit of kJ/mol dioxygenase reaction, is that stored in $C_aH_bO_{c+2}$. For example, a dioxygenase reaction happens on the benzene ring. Same as monooxygenase reaction discussed above, the energy stored is estimated as Eu. (7).

$$\begin{aligned}\Delta G_{dioxy} &= \Delta G(C_aH_bO_{c+2}) - \Delta G(C_aH_bO_c) \\ &= (\Delta G_f(C_aH_bO_{c+2}) - \Delta G_f(C_aH_bO_c)) \\ &\quad + 2\Delta G_f(H_2O) - 4\Delta G_f(H^+) \\ &= 265.98 - 474.36 - 4\Delta G_f(H^+) \\ &= -208.38 - \Delta G_f(H^+)\end{aligned} \tag{7}$$

At pH = 7, $\Delta G_{dioxy} = -48.9$ kJ/mol dioxygenase reaction or $\Delta G_{e\text{-}EA}(O_2\text{−dioxy}) = -48.9/4 = -12.23$ kJ/e^- eq.

Table 10 The estimation of ΔG_e of $C_aH_bO_{c+1}$ (at pH 7)

$C_aH_bO_c$	ΔG_e^0	γ_s	$C_aH_bO_{c+1}$	ΔG_e	Estimated ΔG_e	Estimation error
CH_4	−24.498	8	CH_4O	−35.421	−34.1723	−3.53%
C_2H_6	−25.961	7	C_2H_6O	−29.723	−31.042	4.44%
C_3H_8	−26.204	6.67	C_3H_8O	−28.507	−29.6183	3.90%
C_4H_{10}	−26.220	6.5	$C_4H_{10}O$	−27.817	−28.781	3.47%

Note: ΔG_e^0 represents the energy released from the oxidation half reaction of $C_aH_bO_c$ to CO_2 per e^- eq. electron and it is calculated based on the reported Gibbs energy of formation and Eq. (3); ΔG_e represents the energy released from the oxidation half reaction of $C_aH_bO_{c+1}$ to CO_2 per e^- eq. electron and it is calculated based on Eq. (4); The estimated ΔG_e is estimated according to Eq. (6); Estimation error is defined as (estimated $\Delta G_e - \Delta G_e)/\Delta G_e$; The average estimation error of the dataset is 2.07% and the average of absolute estimation error is 3.84%)

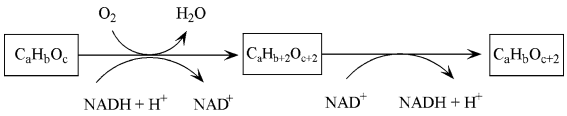

Fig. 5 Simplified process model of dioxygenase reaction (During one dioxygenase reaction, the reduction of one mole molecular oxygen, O_2, and the oxidation of substrate are coupled together. In total, the 4 e^- equivalents of electrons needed for O_2 reduction are from the oxidation of substrate. But in contrast to monooxygeanse reactions, two moles of oxygen atoms, [O], are inserted into the substrate.)

Generally, the stored energy during oxygenase reaction is very small compared with the energy released from the reduction of the common electron acceptor, oxygen ($\Delta G_{e\text{-}EA}(O_2)$ = −78.685 kJ/e^- eq). Therefore, this mini energy modification, while providing explicit consideration of the effect of the inserted oxygen, generally does not alter the yield estimation for heterotrophs significantly. Thus, it is feasible to ignore it and only consider the importance of the 4 electron equivalents lost per oxygenase reaction. However, for autotrophic systems, the stored energy due to the oxygenase reaction has to be considered since the overall energy generation is also very low.

Appendix B: Application of the expanded thermodynamic true yield prediction model—methanotrophic growth example

In order to show the calculation process step by step, an example is considered. We consider methanotrophic bacteria growing on the media containing methane and nitrate under aerobic condition at pH 7 and temperature 35°C. Based on this information, the different roles during degradation are identified in Table 11. In this example, methane acts as sole CS and sole ED; oxygen acts as the terminal EA associated with energy generation, EA(1), and nitrate acts as the nitrogen source. Nitrate is also considered a secondary EA, EA(2), because it is an additional sink for electrons. Correspondingly, the products of the CS and ED are CO_2 and new biomass; the products of EA(1) and EA(2) are O(−2) and N(−3), respectively; and the product of the NS is N(−3). Therefore, the carbon balance (Eq. (1)) requires that the fraction of carbon in CO_2 (f_{CO_2}) and in new biomass (f_{cell}) sum to 1.

Table 11 Compounds and relevant roles in methanotrophic example system

	Substrate	Product(s)
Carbon source	Methane	CO_2, cell
Nitrogen source	Nitrate	Cell-N (−3)
Electron donor	Methane	CO_2, cell
Electron acceptor (1)	O_2	O(−2)
Electron acceptor (2)	Nitrate	Cell-N (−3)
Electron acceptor (3)	O_2 in oxygenase reaction	O(−2)

$$f_{CO_2} + f_{Cell} = 1 \tag{8}$$

The electron balance (Eq. (9)) requires the electrons donated during oxidation of methane to CO_2 and biomass equal the electrons accepted by O_2 and nitrate.

$$f_{CO_2} \times (\gamma_{ED} - \gamma_{CO_2}) + f_{cell} \times (\gamma_{ED} - \gamma_X) = g(1) + g(2) + g(3) \tag{9}$$

The electrons associated with the oxidation of methane (to CO_2 and biomass) are calculated by considering the fraction of carbon going to CO_2(f_{CO_2}) and the difference in reductance degree between methane and CO_2($\gamma_{ED} - \gamma_{CO_2}$) and separately considering the fraction of carbon going to cells (f_{cell}) and the difference in reductance degree between methane and cells ($\gamma_{ED} - \gamma_X$). For the right side of Eq. (2), although the reduction product of O_2 can be expressed as O(−2) generally, evidence shows that the first step of the aerobic degradation of methane is catalyzed by a monooxygenase enzyme. (Colby et al. 1977) The monooxygenase reaction has special energy consumption which is discussed below. Consequently, we split the electrons accepted by oxygen into two parts, the electrons accepted by oxygen during the oxygenase reaction, $g(3)$, and the electrons accepted by oxygen as a terminal electron acceptor for energy generation, $g(1)$. The electrons accepted by nitrate are considered $g(2)$.

Generally, the nitrogen balance can be ignored as long as nitrogen is not the limiting element of bacterial growth. For example, in this case, ignoring the nitrogen balance does not affect the whole model since the nitrogen balance is expressed as nitrogen in NS is equal to the nitrogen incorporated into new biomass. But when nitrogen limits bacteria growth, it

has to be considered more carefully. For example, the nitrogen balance is very important for nitrifiers (where ammonia is the electron donor) and denitrifiers (where nitrate is the electron acceptor) where nitrogen is used in catabolism and anabolism. Nitrogen balance is less critical for aerobic heterotrophs using dissolved ammonia or nitrate only as a nitrogen source for anabolism.

After defining the carbon balance, nitrogen balance and electron balance for the example, the next step is to analyze the energy transfer during metabolism. Metabolism is a complex, multi-faceted process involving numerous reaction steps. However, for thermodynamic modeling of yield prediction, Fig. 6 provides sufficient detail. The catabolic process in our example includes four half reactions: (1) from methane to CO_2 (see Fig. 6a); (2) from O_2 to O(−2) during energy generation; (3) from O_2 to O(−2) during the oxygenase reaction (Fig. 6b); (4) from nitrate to ammonia (Fig. 6c). The anabolic process is simplified as two half reactions by assuming acetate is the critical intermediate: (1) from methane to acetate; (2) from acetate to new cell.

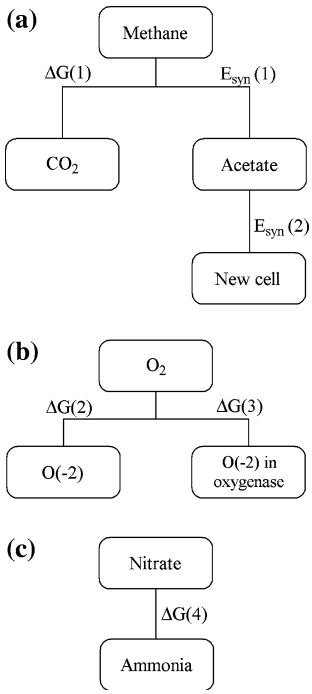

Fig. 6 Simplified metabolism process and the energy consumption of each step in the methanotrophic system with O_2 as electron acceptor and nitrate as nitrogen source (During catabolism, the methane is oxidized into carbon dioxide, oxygen is reduced to the valence of −2 and nitrate is reduced to ammonia; during anabolism, methane is transformed to acetate and then new cells are synthesized based on acetate and ammonia.)

With the two systematic adjustments that we propose to improve the predictive ability of the expanded thermodynamic model, the energy computation of the target system can be calcluated. Based on the model formulation, the energy balance is written as Eq. (10).

$$K \times \sum_{i=1}^{4} \Delta G(i) + \frac{E_{\text{syn}}(1)}{K^{\text{m}}} + \frac{E_{\text{syn}}(2)}{K} = 0 \qquad (10)$$

The sum of the $\Delta G(i)$ includes $\Delta G(1)$: the energy released during the oxidation process of CH_4 to CO_2. $\Delta G(1) = f_{CO_2} \times \Delta G_{CH_4}$, where ΔG_{CH_4} is the Gibbs energy change of $CH_4 + 3H_2O = 0.183H_2CO_3 + 0.817HCO_3^- + 8.818H^+ + 8e^-$, $\Delta G_{CH_4} = -198.827$ kJ/mol-C and f_{CO_2} is the fraction of the carbon in methane that is oxidized to CO_2; $\Delta G(2)$: the energy released during the reduction process of O_2 to O(−2) excluding oxygenase reaction. $\Delta G(2) = g(1) \times \Delta G_{e-O_2}$, where, ΔG_{e-O_2} is the Gibbs energy change of $\frac{1}{4}O_2 + H^+ = \frac{1}{2}H_2O$, $\Delta G_{e-O_2} = -78.685$ kJ/e$^-$ eq. and $g(1)$ is the equivalents of electrons accepted by oxygen during all reductive reactions except the oxygenation reaction; $\Delta G(3)$: the energy released during the reduction process of O_2 to O(−2) in oxygenase reaction only. $\Delta G(3) = g(3) \times \Delta G_{\text{e-monooxy}}$, where $g(3)$ = 4 e$^-$ eq./mol substrate and $\Delta G_{\text{e-monooxy}}$ = −2.263 kJ/e$^-$ eq.; and $\Delta G(4)$: the energy released during the reduction process of NO_3^- to NH_3 during cell synthesis. $\Delta G(4) = g(2) \times \Delta G_{\text{e-nitrate}}$, where $\Delta G_{\text{e-nitrate}}$ is the Gibbs energy change of the half reaction of $\frac{1}{8}NO_3^- + \frac{9}{8}H^+ + e^- = \frac{1}{8}NH_3 + \frac{3}{8}H_2O$, $\Delta G_{\text{e-nitrate}}$ = −33.54 kJ/e$^-$ eq. and $g(2)$ is again the equivalents of electrons gained by nitrate, $g(2) = f_{\text{cell}} \times (\gamma_{NH_3} - \gamma_{NO_3^-}) = 8 \times f_{\text{cell}}$.

The energy associated with synthesis in the model is described with two terms. E_{syn} (1): the energy released during the process of transferring f_{cell} mol methane to acetate. $E_{\text{syn}}(1) = f_{\text{cell}} \times (\Delta G_{CH_4} - \Delta G_{\text{acetate}})$. $\Delta G_{\text{acetate}}$ is the Gibbs energy change of the half reaction of acetate, $\frac{1}{2}C_2H_3O_2^- + 2H_2O = 0.183H_2CO_3 + 0.817HCO_3^- + 4.317H^+ + 4e^-$, $\Delta G_{\text{acetate}}$ = −106.302 kJ/mol-C. E_{syn} (2): the energy needed for cell synthesis based on acetate and ammonia. $E_{\text{syn}}(2) = f_{\text{cell}} \times \frac{\Delta G_{\text{ATP}} \times \text{MW}_{\text{cell}}}{Y_{\text{ATP}} \times 0.9} = f_{\text{cell}} \times \frac{30.53 \times 26.4}{10.5 \times 0.9} = 85.29 f_{\text{cell}}$.

The energy efficiency term, K, is taken as 0.41 (see Xiao and VanBriesen (2006) for details of K estimation). $m = +1$ if the sign of E_{syn} (1) is positive (the corresponding half reaction consumes energy), otherwise, $m = -1$ (energy is released).

Consequently, the model becomes the three equations, Eqs. (8), (9) and (10). Then the estimation of the true yield in this case can be obtained by solving these equations simultaneously, $f_{cell} = 0.51$ mol-C cell/mol-C CH_4. Comparing with the experimental true yield reported by Heijnen and Roels (1981), 0.506 mol-C cell/mol-C CH_4 (calculated from observed yield, maintenance and growth rate), the estimation error is low (+ 1%).

References

Andrews G (1993) The yield equations in the modeling and control of bioprocesses. Biotechnol Bioeng 42:549–556

Argaman Y, Brenner A (1986) Single-sludge nitrogen removal—modeling and experimental results. J Water Pollut Control Federat 58:853–860

Arp DJ, Sayavedra-Soto LA, Hommes NG (2002) Molecular biology and biochemistry of ammonia oxidation by *Nitrosomonas europaea*. Arch Microbiol 178:250–255

Atkinson B, Mavituna F (1991) Biochemical engineering and biotechnology handbook. Stockton Press, New York

Bainottie AE, Nishio N (2000) Growth kinetics of *Acetobacterium* sp. on methanol-formate in continuous culture. J Appl Microiol 88:191–201

Bally M, Wilberg E, Kuhni M, Egli T (1994) Growth and regulation of enzyme synthesis in the nitrilotriacetic acid (NTA)-degrading bacterium *Chelatobacter heintzii* ATCC 29600. Microbiology 140:1927–1936

Bauchop T (1958) Observations on some molar growth yields of bacteria and yeasts. J Gen Microbiol 18:vii

Benjamin MM (2002) Water chemistry. McGraw-Hill Companies, Inc., New York

Birou B, Marison IW, vonStockar U (1987) Calorimetric investigation of aerobic fermentations. Biotechnol Bioeng 30:650–660

Boogerd FC, van Verseveld HW, Torenvliet D, Braster M, Stouthamer AH (1984) Reconsideration of the efficiency of energy transduction in *Paracoccus denitrificans* during growth under a variety of culture conditions: a new approach for the calculation of $P/2e^-$ ratios. Arch Microbiol 139:344–350

Burton K (1958) Energy of adenoising triphosphate. Nature 181:1594–1595

Colby J, Stirling DI, Dalton H (1977) The soluble methane mono-oxygenase of *Methylococcus capsulatus* (Bath). Its ability to oxygenate *n*-alknaes, *n*-alkenes, ethers, and alicyclic, aromatic and heterocyclic compounds. Biochem J 165:395–402

Corseuil HX, Weber WJ (1994) Potential biomass limitations on rates of degradation of monoaromatic hydrocarbons by indigenous microbes in subsurface soils. Water Res 28:1415–1423

Demirdogen BC, Adali O (2005) Characterization and modulation by drugs of sheep liver microsomal flavin monooxygenase activity. Cell Biochem Funct 23:245–251

De Vries W, Niekus HGD, Boellaard M, Stouthamer AH (1980) Growth yield and energy generation by *Campylobacter sputorum* subspecies *bubulus* during growth in continuous culture with different hydrogen acceptors. Arch Microbiol 124:221–227

Dixon M, Kleppe K (1965) D-Amino acid oxidase I. Dissociation and recombination of the haloenzyme. Biochim Biophys Acta 96:357–367

Egli T (2001) Biodegradation of metal-complexing aminopolycarboxylic acids. J Biosci Bioeng 92(2):89–97

Enroth C, Neujahr H, Schneider G, Lindqvist Y (1998) The crystal structure of phenol hydroxylase in complex with FAD and phenol provides evidence for a concerted conformational change in the enzyme and its cofactor during catalysis. Structure 6:605–17

Ensign SA, Small FJ, Allen JR (1998) New roles for CO_2 in the microbial metabolism of aliphatic epoxides and ketones. Arch Microbiol 169:179–187

Eubanks EF, Forney FW, Larson AD (1974) Purification and characterization of nocardial acetylesterase involved in 2-butanone degradation. J Bacteriol 120:1133–1143

Fosdike WLJ, Smith TJ, Dalton H (2005) Adventitious reactions of alkene monooxygenase reveal common reaction pathways and component interactions among bacterial hydrocarbon oxygenase. FEBS J 272:2661–2669

Hamamura N, Storfa R, Semprini L, Arp DJ (1999) Diversity in butane monooxygenases among butane-grown bacteria. Appl Environ Microbiol 65:4586–4593

Heijnen JJ (1984) Biological industrial waste water treatment minimizing biomass production and maximizing biomass concentration. Ph.D. dissertation. Delft University of Technology, Delft, The Netherlands

Heijnen JJ (1994) Thermodynamics of microbial growth and its implications for process design. Trend Biotechnol 12:483–492

Heijnen JJ (1999) Bioenergetics of microbial growth. In: Flickinger MC, Drew SW (eds) Encyclopedia of bioprocess technology: fermentation, biocatalysis, and bioseparation. John Wiley and Sons, Inc., New York

Heijnen JJ, Roels JA (1981) A macroscopic model describing yield and maintenance relationships in aerobic fermentation processes. Biotechnol Bioeng 23:739–763

Heijnen JJ, van Dijken JP (1992) In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. Biotechnol Bioeng 39:833–858

Heijnen JJ, van Dijken JP (1993) Response to comments on ''in search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms''. Biotechnol Bioeng 42:1127–1130

Henneken L, Nortemann B, Hempel DC (1998) Biological degradation of reaction EDTA: kinetics and technical approach. J Chem Technol Biotechnol 73:144–152

Henze M, Grady CPL, Gujer W, Marais GVR, Matsuo T (1987) IAWPRC scientific and technical report, No 1. Activated sludge model No 1. London, IAWPRC

Henze M, Gujer W, Mino T, Matsuo T, Wentzel MC, Marais GVR (1995) IAWQ scientific and technical report, No 3. Activated sludge model No 2. London, IAWQ

Hernandez E, Johnson MJ (1967) Energy supply and cell yield in aerobically grown microorganisms. J Bacteriol 94(4):996–1001

Kappler O, Janssen PH, Kreft JU, Schink B (1997) Effects of alternative methyl group acceptors on the growth energetics of the o-demethylating anaerobe *Holophaga foetida*. Microbiology (UK) 143:1105–1114

Kappler U, Friedrich C, Truper H, Dahl C (2001) Evidence for two pathways of thiosulfate oxidation in *Starkeya novella* (formerly *Thiobacillus novellus*). Arch Microbiol 175:102–111

Kari FG, Giger W (1996) Speciation and fate of ethylenediaminetetraacetate (EDTA) in municipal wastewater treatment. Water Res 30:122–134

Kauppi B, Lee K, Carredano E, Parales RE, Gibson DT, Eklund H, Ramaswamy S (1998) Structure of an aromatic-ring-hydroxylating dioxygenase¨C naphthalene 1,2-dioxygenase. Structure 6:571–86

Liebig T, Wagner M, Bjerrum L, Denecke M (2001) Nitrification performance and nitrifier community composition of a chemostat and a membrane-assisted bioreactor for the nitrification of sludge reject water. Bioproce Biosyst Eng 24:203–210

Linton JD, Stephenson RJ (1978) A preliminary study on growth yields in relation to the carbon and energy content of various organic growth substrates. FEMS Microbiol Lett 3:95–98

Mavrovouniotis ML (1990) Group contributions for estimating standard gibbs energies of formation of biochemical compounds in aqueous solution. Biotechnol Bioeng 36:1070–1082

Mavrovouniotis ML (1991) Estimation of standard Gibbs energy changes of biotransformations. J Biol Chem 266:14440–14445

McCarty PL (1965) Thermodynamics of biological synthesis and growth. In: Baers J (ed) Advances in water pollution research: proceedings of the 2nd international conference on water pollution research. Pergamon Press, Inc., Oxford England, pp 169–199

McCarty PL (2007) Thermodynamic electron equivalents model for bacterial yield prediction: modifications and comparative evaluations. Accepted and in Press, Biotechnology and Bioengineering.

Muller A, Wentzel MC, Loewenthal RE, Ekama GA (2003) Heterotroph anoxic yield in anoxic aerobic activated sludge systems treating municipal wastewater. Water Res 37:2435–2441

Noguera DR, Brusseau GA, Rittmann BE, Stahl DA (1998) A unified model describing the role of hydrogen in the growth of *Desulfovibrio vulgaris* under different environmental conditions. Biotechnol Bioengr 59:732–746

Peters J, Witholt B (1994) Solubilization of the overexpressed integral membrane protein alkane monooxygenase of the recombinant *Escherichia coli W3110[pGEc47]*. Biochim Biophys Acta 1196:145–53

Pronk JT, Meesters PJW, van Dijken JP, Bos P, Kuenen JG (1990) Heterotrophic growth of *Thiobacillus acidophilus* in batch and chemostat cultures. Arch Microbiol 153:392–398

Reardon KF, Mosteller DC, Rogers JDB (2000) Biodegradation Kinetics of Benzene, Toluene, and Phenol as Single and Mixed Substrates for *Pseudomonas putida* F1. Biotechnol Bioeng 69:385–400

Rittmann BE, McCarty PL (2001) Environmental biotechnology: principles and applications. McGraw-Hill Book Company, New York

Roels JA (1983) Energetics and kinetics in biotechnology. Elsevier biomedical press, Amsterdam, the Netherlands

Rutgers M (1990) Control and thermodynamics of microbial growth. Ph.D. dissertation. University of Amsterdam. The Netherlands

Sanders HK, Becker GE, Nason A (1972) Glycine-cytochrome c reductase from *Nitrobacter agilis*. J Biol Chem 247:2015–2025

Schink B (1984) Fermentation of 2,3-butanediol by *Pelobacter carbinolicus* sp. nov. and P*elobacter propionicus* sp. nov., and evidence for propionate formation from C_2 compounds. Arch Microbiol 137:33–41

von Stockar U, Liu JS (1999) Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. Biochem Biophys Acta 1412:191–211

Timkovich R, Thrasher JS (1988) Carbon monoxide oxygenase activity of *cytochrome cd1*. Biochemistry 27:5383-8

van Verseveld HW (1979) Influence of environmental factors on the efficiency of energy conservation in *Paracoccus denitrificans*. Ph.D. thesis, Free Universit y of Amsterdam, Amsterdam, The Netherlands

van Verseveld HW, Meijer EM, Stouthamer AH (1977) Energy conservation during nitrate respiration in *Paracoccus denitrificans*. Arch Microbiol 112:17–23

VanBriesen JM (2001) Thermodynamic yield predictions for biodegradation through oxygenase activation reactions. Biodegradation 12:265–281

VanBriesen JM (2002) Evaluation of yield predictions using thermodynamics. Biodegradation 13:171–190

VanBriesen JM, Rittmann BE (2000) Modeling speciation effects on biodegradation in mixed metal/chelate systems. Biodegradation 10:315–330

Verduyn C, Stouthamer AH, Scheffers A, vanDijken JP (1991) A theoretical evaluation of growth yields of yeasts. Anton Leeuw 59:49–63

Verduyn C (1991) Physiology of yeasts in relation to biomass yields. Anton Leeuw 60:325–353

Woo SH, Rittmann BE (2000) Microbial energetics and stoichiometry for biodegradation of aromatic compounds involving oxygenation reactions. Biodegradation 11:213–227

Xiao J, VanBriesen JM (2006) Expanded thermodynamic model for microbial true yield prediction. Biotechnol Bioeng 93(1):110–121

Yuan Z, VanBriesen JM (2002) Yield prediction and stoichiometry of multi-step biodegradation reactions involving oxygenation. Biotechnol Bioeng 80:100–113

Yuan Z (2004) The role of oxygenation and intermediates in biodegradation of chelating agents. Ph.D. dissertation. Carnegie Mellon University, US