Systems biology

An integrated open framework for thermodynamics of reactions that combines accuracy and coverage

Elad Noor1,†, Arren Bar-Even1,†, Avi Flamholz1, Yaniv Lubling2, Dan Davidi1 and Ron Milo1,*

1Department of Plant Sciences and 2Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

Associate Editor: Olya Troyanskaya

ABSTRACT

Motivation: The laws of thermodynamics describe a direct, quantitative relationship between metabolite concentrations and reaction directionality. Despite great efforts, thermodynamic data suffer from limited coverage, scattered accessibility and non-standard annotations. We present a framework for unifying thermodynamic data from multiple sources and demonstrate two new techniques for extrapolating the Gibbs energies of unmeasured reactions and conditions.

Results: Both methods account for changes in cellular conditions (pH, ionic strength, etc.) by using linear regression over the ΔG

The Pseudoisomeric Reactant Contribution method systematically infers compound formation energies using measured K

Availability: Freely available on the web at: http://reactions.reizewijk/ We implement in Python, MySQL, Apache and Django, with all major browsers supported. The framework is open-source code.google.com/p/milo-lab implemented in pure Python and tested mainly on Linux.

Contact: ron.milo@weizmann.ac.il

Supplementary Information: Supplementary data are available at Bioinformatics online.

Received on March 11, 2012; revised on May 21, 2012; accepted on May 23, 2012

1 INTRODUCTION

The study of metabolism has recently regained its central role in diverse areas of cell biology, physiology, medicine and systems biology. The study of metabolic pathways and networks (Haverkorn et al., 2007; Burgard et al., 2003; Oberhardt et al., 2003; Heinrich et al., 2009, 2010) aims to better understand the complex behaviour of living organisms as well as their manipulation for human needs. (Asumi et al., 2008; Bar-Joseph et al., 2011; Sinha et al., 2010; Yim et al., 2011; Chang et al., 2011). The thermodynamics of biochemical reactions is of special interest in the analysis and design of metabolic pathways.

The change in the Gibbs free energy (ΔG) characterizes the balance of biochemical reactions and dictates the direction of net flux (the difference between forward and backward fluxes) in a reaction. It is thus useful for the study of a single enzymatic reaction, for analyzing entire metabolic pathways (Voitnov and von Stockel, 2008), and for the large-scale modeling of whole-cell metabolic networks (Henry et al., 2004).

A reaction at equilibrium carries no net flux. At equilibrium, the apparent reaction quotient Q—the ratio of product to substrate concentrations—is termed the apparent equilibrium constant denoted by K. In ideal dilute solutions, the ‘transformed Gibbs energy of reaction’ is a function of the apparent equilibrium constant: ΔG = −RT ln K (Eq. 1). The ‘standard’ transformed Gibbs energy of reaction (ΔG') is the value of ΔG at standard conditions, i.e. when all compound concentrations are 1 M. Therefore, ΔG = −RT ln K (Eq. 1)

The direction of net flux in a reaction. A negative ΔG' would correspond to a positive (forward) net flux and vice versa. Cell physiology imposes constraints on metabolite concentrations and ΔG' < 0 for any flux and are thus called irreversible reactions. This classification of reactions is especially important in constraint-based modeling that covers whole-cell metabolic networks and depends on the knowledge of reaction directionality for predicting flux distributions, growth rates and other large-scale metabolic phenotypes (Boo et al., 2007; Burgard et al., 2005; Oberhardt et al., 2005). Directionality, annotations typically rely on phenomenological data and arbitrary definitions of reversibility. Recent advances in the field allow incorporating thermodynamic data directly into the model by adding explicit constraints that connect ΔG' concentrations and reaction directionality (Kappe et al., 2005; Henry et al., 2009).

The NIST database for Thermodynamic Data of Enzyme-Catalyzed Reactions (NIST TECR) is the most comprehensive collection of thermodynamic data (Goldberg et al., 2008, 2009). About 400 reactions have measured equilibrium constants and were

© The Author(s) 2012. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
E. Noor et al.

successfully mapped to KEGG identifiers [Kanehisa et al., 2008]. Still, this represents only about 6% of the \(\sim 5500\) biochemically relevant reactions in the KEGG database. Moreover, since equilibrium constants in the NIST-TECR database are measured in a variety of pH and ionic strength levels, it is difficult to extrapolate the equilibrium constant for the conditions prevailing in an organism or system of interest. When \(\Delta_r G^\circ\) are not covered by the NIST-TECR database, several collections of standard Gibbs free energy of formation (\(\Delta_f G^\circ\)) can be consulted and used as in Figure 1—equation II. The largest collection, given by Alberty [1998], contains \(\Delta_f G^\circ\) values for \(\sim 200\) compounds. Unfortunately, many reactions were not measured in the desired conditions (pH, ionic strength, etc.) or have never been experimentally measured at all. If the formation energy of even one of the reactants is unknown, \(\Delta_r G^\circ\) of a reaction cannot be derived. In order to bridge the gap between the known formation energies and the unknown ones, a method based on the group contribution assumption for biochemical compounds in aqueous solutions was described by Mavrovouniotis [1990, 1991] and later greatly improved in terms of coverage and accuracy by Jankowski et al. [2008]. The method is based on the simplifying assumption that each functional group has a characteristic contribution to the overall formation energy of a compound and that these group contributions are independent of each other. Therefore, \(\Delta_f G^\circ\) is estimated by summing all contributions from the different groups composing a compound. The contribution of each such group (\(\Delta_f G^\circ\)) is estimated through linear regression, which uses the known \(\Delta_f G^\circ\) and \(\Delta_r G^\circ\) and the compounds’ partition into groups [Mavrovouniotis, 1991].

The group contribution method is limited in its accuracy. The group independence assumption can be overly simplistic, especially for large compounds or in conjugated systems. In addition, the definitions of the groups rely heavily on heuristics and chemical intuition. Recent improvements of the method included a better choice of groups and the introduction of group interaction corrections [Jankowski et al., 2008]. Recently, a promising new approach based on whole-reaction similarities [Rother et al., 2010] has been shown to be more accurate, but currently does not provide as wide a coverage as group contribution methods.

Previous group contribution studies did not consider the effect of pH on the compounds’ protonation levels [Alberty, 1998; Jankowski et al., 2008; Mavrovouniotis, 1990, 1991] and assumed all measurements were taken at standard aqueous conditions (e.g. pH 7 and ionic strength of 0.25 M). Each compound is actually an

---

**Fig. 1.** An overview of the relationships between the layers of thermodynamic data. Reaction equilibrium constants can be converted to \(\Delta_r G^\circ\) which is calculated as the stoichiometric sum of reactant formation energies (blue shading, equations I and II). Each reactant exists as several pseudoisomers distributed according to the Boltzmann distribution, and its \(\Delta_f G^\circ\) can be found using the Legendre transform (yellow shading, equations III and IV). Using the group contribution assumption, each pseudoisomer can be constructed from its group components (green shading, equations V and VI).
ensemble of ‘pseudoisomers’ differing in their protonation levels (e.g. ATP$^{4-}$, HATP$^{3-}$ and H$_2$ATP$^{2-}$) and its $\Delta G^\circ$ can be found using the Legendre transform (Fig. 1—yellow shading, equations III and IV). As pointed out by Alberty (2003), ignoring the change in the abundance of the different pseudoisomers at changing pH levels can result in errors. This is especially true for the biochemically ubiquitous phosphate groups, which have $pK_a$ values in the range of 6–8 (Robitaille et al. 1991), and hence their pseudoisomer distribution changes considerably even close to pH 7.

The concentrations of the individual pseudoisomers are determined by the Boltzmann distribution (Alberty 2003). $\Delta G^\circ(i)$ of a compound $i$ is, therefore, a function of $\Delta G^\circ(j)$ of the compound’s pseudoisomers (Fig. 1—equation III). Notably, $\Delta G^\circ(j)$ of the pseudoisomers and their distribution is modulated by the pH (Fig. 1—equation IV). For example, at a high pH, the unprotonated form of an acidic compound will have lower $\Delta G^\circ$ and thus be more abundant (Alberty, 2003). In order to use $\Delta G^\circ$ to calculate $\Delta G^\circ$, the different pseudoisomer forms assumed by each compound should be considered.

2 APPROACH

In this study, we present methods to accurately estimate $\Delta G^\circ$ using two major approaches (Fig. 1). The first approach, named Pseudosomeric Reactant Contribution (PRC), recovers pseudoisomer formation energies by applying linear regression to the entire set of reactions available in the literature (as recorded in NIST-TECR). Most measurements in the NIST-TECR database are of ‘apparent’ equilibrium constants, which is the equilibrium apparent reaction quotient of the total compound concentrations—i.e. the sum of all protonation states. Since the pH affects the distribution of species non-linearly, the value of $\Delta G^\circ$ cannot be expressed as a linear sum of the reactants’ formation energies. PRC applies the inverse Legendre transform to linearize the system. The inverse Legendre transform replaces each ensemble of pseudoisomers with a single representative and changes the observed value of $K$ accordingly (Alberty 2003) using the known dissociation constants of the reactants. This is thus possible to find the least-squares solution for the $\Delta G^\circ$ of the pseudoisomers using linear regression (see Supplementary Material for details). Using PRC, we can retrieve $\Delta G^\circ$ values that were previously unknown (Alberty 2003) and enable the calculation of $\Delta G^\circ$ for more reactions.

Compounds that were not measured previously and thus do not appear in the NIST-TECR database require a different approach. We have developed an augmented group contribution method that we call Pseudosomeric Group Contribution (PGC). Unlike previous approaches, we decompose pseudoisomers, not compounds, into functional groups (Fig. 1—green shading). We then estimate $\Delta G^\circ$ of the pseudoisomers by summing over the contributions of their groups (equation V) and calculate the $\Delta G^\circ$ of compounds (equations III and IV). This method provides higher accuracy and can correctly adjust the $\Delta G^\circ$ to the different aqueous conditions described by pH and ionic strength (with pH usually being the most significant). The combination of the two approaches enables a better estimation of $\Delta G^\circ$ for a large variety of compounds in a wide range of aqueous conditions.

3 METHODS

3.1 A PRC method systematically derives $\Delta G^\circ$

The task of estimating the formation energies of compounds given measured reaction equilibrium constants is not straightforward. The difficulty stems from the non-linearity of $\Delta G^\circ$ as a function of $\Delta G^\circ$ and pH (Fig. 1—equations II–IV). However, if $\Delta G^\circ$ has been measured in a specific pH and all but one of the reactants’ $\Delta G^\circ$ values is known, it is straightforward to infer the missing $\Delta G^\circ$.

The extensive list of compound formation energies provided in Alberty (2003) is the product of a meticulous reconstruction such as described above and based on the data provided by many measurements of $K’$ as each new $\Delta G^\circ$ added to this database relies on the previously gathered values; measurement errors for a particular compound are carried on to future calculations.

Alternatively, compound acid dissociation constants ($pK_a$s) can be used to convert the set of equations into a linear system that can be solved computationally. The idea, known as the ‘Inverse Legendre Transform’ (see Supplementary Material), is based on the fact that the difference between the $\Delta G^\circ$ and $\Delta G^\circ$ of any reaction is a function of the $pK_a$s alone and does not depend on the absolute formation energies of the reactants. The resulting linear system can then be solved using multiple linear regression. This application of the inverse Legendre transform to linearize a reaction system was introduced by (Alberty 2003) and to our knowledge was implemented only for a small set of reaction measurements. Here, we perform a global analysis using all the available data in NIST-TECR to achieve the best possible least-squares estimation. We refer to this method as the Pseudosomeric Reactant Contribution (PRC) method.

Using PRC, we were able to obtain values for 407 formation energies using only the 367 reactions in the NIST-TECR dataset and the $pK_a$ values of the participating compounds (described later in ‘Acid dissociation constants’). A detailed analysis of these values in terms of prediction power and accuracy is given in Section 4. Due to linear dependencies between some of the reactions, there are 112 dimensions in the null space of the stoichiometric matrix. For instance, if two compounds always appear together (such as NADox and NAD red), the difference between the $\Delta G^\circ$ of the pair can be inferred, but the absolute formation energy remains unknown. Similarly, element conservation rules contribute one dimension to the null space for every element which appears in the database, namely C, N, O, S and P. Commonly, as is the case in the current study, the ambiguity in the values of the formation energies is solved by defining some compounds as having $\Delta G^\circ$ = 0. Alberty’s table of ~200 formation energies (Alberty 1992) contains 18 such reference points.

3.2 A PGC method covers more reactions and conditions

In order to improve the estimations provided by the group contribution method (Alberty 2003) for highly pH-dependent compounds, we introduce a method that incorporates Alberty’s transformed formation energies (Alberty 2003) into the same framework used by Jankowski et al. (2003). Previous group contribution implementations used apparent equilibrium constants ($K’$) and formation energies ($\Delta G^\circ$) at pH~7, where each compound is actually an ensemble of protonation species—or pseudoisomers (Fig. 1—equation III). For instance, the total concentration of ATP is divided between the pseudoisomers H$_2$ATP$^{2-}$, HATP$^{3-}$ and ATP$^{4-}$ according to the Boltzmann distribution. The formation energy of the ensemble is called the standard transformed Gibbs energy as defined by the International Union of Biochemistry and Molecular Biology (IUBMB) (Alberty et al. 2003). A shift in pH will change the relative abundance of the different pseudoisomers and affect $\Delta G^\circ$ and $K’$ non-linearly. Therefore, in order to normalize the effect of pH across different measurements, we use only formation energies of single pseudoisomers as
input for the linear regression step in our group contribution framework. For example, instead of the ‘transformed’ formation energy of ATP ($-2292.5$ kJ/mol), we use the formation energy of HATP$^-\gamma$, which is $-2768.1$ kJ/mol (Alberty, 2003). This change requires knowing the standard formation energy of each pseudosolomer separately and the exact distribution of protons and charges in its molecular structure. In the case of reactions, we use the inverse Legendre transform—similarly to the PRC method.

In addition, groups that previously had only one form corresponding to the protonation level most abundant at pH 7 can now have multiple forms representing the different protonations at a wide range of pHs. The full list of groups is given in Supplementary Table S4. For example, the terminal phosphate group $-\text{PO}_2^-$ had only one instantiation in previous implementations (Jankowski et al., 2008) ($\Delta_G^\circ = -254$ kJ/mol $= -1063$ kJ/mol), has two versions in PGC namely $\text{PO}_2^-$ and $\text{PO}_2^-$, each with its own contribution ($\Delta_G^\circ = -1024.2$ and $-1073.9$ kJ/mol, respectively). The terminal phosphate of ATP’s major pseudosolomer at pH 7 is semi-protonated and corresponds to the $-\text{PO}_2^-$ group. However, when this group appears in the major pseudosolomer of other compounds such as 3-glucose-6-phosphate, it is fully deprotonated ($-\text{PO}_3^-$). Thus, PGC adjusts the contribution of the two pseudosolomeric groups accordingly. Similarly to previous implementations (Jankowski et al., 2008), the algorithm for determining the group contributions is a least-squares linear regression, except that for the same number of compounds, there are more available values of $\Delta_G^\circ$ and more groups due to the use of pseudosolomers.

We use the NIST-TECR database (Supplementary Table S1), together with a list of formation energies (Albert et al., 2009; Selting and Janousek, 2007; Thauer et al., 2008; Supplementary Table S2) and dissociation constants (Lin et al., 2008; Supplementary Table S3). As part of this study, we have manually annotated each of the species in the list and determined the protonation level of each of their groups. The task was not trivial for compounds with more than one protonation site, since the order of the $p$K$\alpha$ values of the groups determines which will deprotonate first as the pH rises, and these data are not well organized in the literature. For example, the different protonation states of nucleic acids are a particular challenge (Brehm et al., 1979). The details and results of this analysis and the contribution of the pseudosolomeric groups are given in Supplementary Table S4.

Using the PRC method to infer compound formation energies from NIST-TECR, it is possible to obtain predictions for a few hundreds of reactions in the KEGG database (roughly 5%)—not much more than the number of reactions in NIST-TECR itself. However, the PGC method achieves a much better coverage (59%)—not much more than the number of reactions in KEGG database (roughly 5%)—not much more than the number of reactions in KEGG database (roughly 5%).—not much more than the number of reactions in KEGG database (roughly 5%). Thus, PGC provides an efficient way to predict the formation energy of reactions in NIST-TECR itself.

The non-pseudoisomeric group contribution method, which resulted in an RMSE of 11.2 kJ/mol (RMSE). All formation energy and reactions examples were measured. The root mean squared error (RMSE) of each method is calculated by giving each distinct reaction an equal weight (regardless of how many times it has been measured). The reported result is 1.90 kcal/mol (8.0 kJ/mol). In the current article, more reactions from NIST-TECR were used for calculating RMSE since we did not filter observations according to their pH. In addition, we did not include any formation energy data in this evaluation since they are not purely empirical—many are derived from the same data which is already in NIST-TECR, and some values were deposited using a group contribution approach and thus cannot be used to evaluate its precision.

### Table 1. Comparing the different methods for estimating Gibbs free energies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage of reactions</td>
<td>6%</td>
<td>10%</td>
<td>77%</td>
<td>77%</td>
</tr>
<tr>
<td>(out of 5464 in KEGG)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage of reactions</td>
<td>18%</td>
<td>30%</td>
<td>95%</td>
<td>93%</td>
</tr>
<tr>
<td>(out of 729 in $E_{\text{cal}}$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error (RM, kJ/mol)</td>
<td>6.3$^b$</td>
<td>2.4</td>
<td>9.9$^c$</td>
<td>8.5</td>
</tr>
<tr>
<td>Pseudosolomers considered</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Open source</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

$^a$ The relevant reactions in (AP1480).

$^b$ Dithiothreitol (DTT) covers 1073 measurements in NIST-TECR, all other methods cover 2073 reactions.

$^c$ In Jankowski et al. (2008), the reported result is 1.90 kcal/mol (8.0 kJ/mol). In the current article, more reactions from NIST-TECR were used for calculating RMSE since we did not filter observations according to their pH. In addition, we did not include any formation energy data in this evaluation since they are not purely empirical—many are derived from the same data which is already in NIST-TECR, and some values were deposited using a group contribution approach and thus cannot be used to evaluate its precision.

### 3.5 The reversibility index

We introduce a quantitative measure for reversibility that takes the mean squared error (RMSE) of each method is calculated by giving each distinct reaction an equal weight (regardless of how many times it has been measured). The result for a leave-one-out cross-validation test for the PGC method was 11.2 kJ/mol (RMSE). All formation energy and reactions examples were measured. The root mean squared error (RMSE) of each method is calculated by giving each distinct reaction an equal weight (regardless of how many times it has been measured). The reported result is 1.90 kcal/mol (8.0 kJ/mol). In the current article, more reactions from NIST-TECR were used for calculating RMSE since we did not filter observations according to their pH. In addition, we did not include any formation energy data in this evaluation since they are not purely empirical—many are derived from the same data which is already in NIST-TECR, and some values were deposited using a group contribution approach and thus cannot be used to evaluate its precision.

#### 3.5.1 The reversibility index

To quantify the concentration of metabolites and the number of substrates and products into consideration. Further details regarding the application of the index to genome-scale models and motivation for its use are given in Section 4.

We assume that all reactant concentrations lie within a range located symmetrically (in log-scale) around some characteristic physiological concentration $C$—here we use $100\mu$M (Bennett et al., 2008). The variable describing this range is denoted $y$; all substrates are assigned a concentration of $y^C/\sqrt{2}$ and all products have a concentration of $y^C/\sqrt{2}$. Therefore, the log-scale width of the range is $\log(y^{N-1}) - \log(y^C) = \log(y^C/\sqrt{2})$. Thus, a value of $y = 100$ corresponds to the range $10\mu$M–1 mM and $y = 1000$ corresponds to about 0.3–3 mM.

For a reaction with $N_S$ substrates and $N_P$ products, the apparent reaction quotient $Q = \frac{C_{\text{prod}}}{C_{\text{sub}}} = \gamma / \gamma M_N = \gamma M_N$, where we define $N = N_S + N_P$ as the total number of reactants and $Q = C^{N_S} / C^{N_P}$ as the default reaction quotient. We define the reversibility index as $\gamma = \frac{Q}{C^{N_S} / C^{N_P}}$, which is the required concentration range for reversing the reaction, i.e. $Q = C^{N_S} / C^{N_P}$.

The least $\gamma$ from 1 the more irreversible the reaction. The reversibility index of the fructose-bisphosphate aldolase reaction, for example, is 1.04, making it clearly reversible since a change of only 4% in concentrations is required to reverse its direction.
The code for the implementation of the PRC and PGC methods is free and can be found at: http://milo-lab.googlecode.com/svn/branches/bioinfo-2012/

3.6 Source code
The code for the implementation of the PRC and PGC methods is free and open source and can be found at:
http://milo-lab.googlecode.com/svn/branches/bioinfo-2012/

Our software is written completely in Python, depending only on free software such as Open Babel (openbabel.org) and SciPy (scipy.org).

3.6 Source code

4 RESULTS
4.1 Adjustments for pH using $pK_a$ values can increase accuracy

Many organic compounds are weak acids and bases and, as such, assume multiple protonation levels in typical physiological conditions. Formation energies and, as a result, reaction energies are a function of the distribution of reactant pseudoisomers (Fig. 1—equations II and IV). Since the distribution of protonation levels is a function of the prevailing pH and ionic strength, $\Delta G$ and $\Delta G^\circ$ vary with pH and ionic strength as well. For example, the key reaction in gluconeogenesis and C4 plant photosynthesis, pyruvate + ATP + $P_i$ $\rightleftharpoons$ PEP + AMP + $P_i$, is catalyzed by PPDK (pyruvate-phosphate dikinase, EC number 2.7.9.1) and was measured at various pH levels ranging from 6.5 to 8.4 (Fig. 3). Many of the reactants have a $pK_a$ in this pH range, and so $\Delta G^\circ$ of the PPDK reaction changes significantly with pH. Intracellular pH values are typically between 6 and 7.5 for most organisms (Vojinović and von Stockar 2004; but the range can be much wider, e.g. in yeast (5.5–7.5) (Vojinović and Obad 1996; Ryan and Ryan 1972) or bacteria (4.8–9.0) (Breeuwer et al. 1992)). As shown in Figure 3, the PGC and PRC methods can accurately predict the pH dependence of the reaction. Methods that do not account for pseudoisomers can result in a difference of $\pm 20$ kJ/mol for the PPDK reaction. If each compound is assigned only a fixed protonation level, based on the most abundant pseudoisomer at pH 7, the reaction looks as follows:

$$\text{PYR}^- + \text{ATP}^3^- + \text{P}_i^- + \text{H}^+ \rightleftharpoons \text{PEP}^{2+} + \text{AMP}^- + \text{PP}_i^2$$
of groups to the overall simplifying assumption of independence between the contributions could be attributed to the fact that group contribution is based on the Jankowski.

9 kJ/mol for 5 kJ/mol for the PGC method and 9 kJ/mol for the PRC method, there is a good fit between the regression data and independent cross-validation for them. We found that when using previously published methods, there is no way of performing an fact that the training procedures were not made public for the....

K to an observation of K, which is converted into ΔG°, the reaction Gibbs energy in standard conditions (1M concentrations) and in the specific pH and ionic strength of each measurement. The Y-value is calculated using PGC or PRC and adjusted to the appropriate pH and ionic strength. The dashed line marks where estimations are equal to observations. The average estimation error per reaction (RMSE) is 8.5 kJ/mol for PGC and 2.4 kJ/mol for PRC. The estimation error can be explained as follows: measurement noise/bias in the value of K, error in the values of pH and ionic strength, pseudoisomers which are unaccounted for, deviations from the theory of thermodynamics in aqueous solutions and violation of the assumption that the contribution of groups to ΔG° are completely independent (only for PGC).

implies that the reaction energy increases with the pH, since the effect of one proton is kBTln(10)·pH (Fig. 3). However, the PGC and PRC methods, which take the dissociation constants of the reactants into account, show that the response has a negative slope, which corresponds well to measured data (Fig. 3).

4.2 Comparing ΔrG° estimations using the NIST-TECR database

In order to test the accuracy of the two methods described here, we use the NIST-TECR database as a benchmark. Given a measurement of K for a given reaction, we compare the predicted reaction energy (using any of the methods described in Section 3), to the observed reaction energy (Fig. 1—equation 1). The analysis of the PRC and PGC methods is given in Figure 4 and results in root mean squared error (RMSE) of 2.4 and 8.5 kJ/mol, respectively—equivalent to errors of about a factor of 3 and 30 in the estimation of K. Note that all of the different methods, those developed in the past as well as those presented here, have used some or all of the data that appear in NIST-TECR for training the group contributions or formation energies. As a result of these dependencies and the fact that the training procedures were not made public for the previously published methods, there is no way of performing an independent cross-validation for them. We found that when using the PRC method, there is a good fit between the regression data and the observed data in NIST-TECR (RMSE of 2.4 kJ/mol) while the methods based on group contributions do not fit the NIST-TECR data as well (RMSE of 8.5 kJ/mol for the PGC method and 9.9 kJ/mol for the PRC method). The reason for this difference in accuracy could be attributed to the fact that group contribution is based on the simplifying assumption of independence between the contributions of groups to the overall ΔrG°. In addition, there are more free variables in PRC than in PGC (407 compounds versus 99 groups).

A summary of the analysis for the four different estimation methods based on the data provided by NIST-TECR is given in Table 1. The goodness of fit is given by the RMSE for each of the methods. These values should be compared with the baseline method which is to use the average ΔG° of each reaction in NIST-TECR across all its measurements. This baseline achieves the maximum accuracy but is limited in coverage to the scope of NIST-TECR. The baseline RMSE is 1.3 kJ/mol and is the average standard deviation of ΔrG° per reaction.

4.3 Determining the reversibility of reactions

As an example of the usefulness of having a framework that makes all thermodynamic data available in one location and in an open format, we discuss the issue of reaction directionality that plays a pivotal and often problematic role in many metabolic models. (Feist et al., 2003; Oberhardt et al., 2008) and has a crucial effect on their results. A reaction is called irreversible if its net flux flows in the same direction under all allowed physiological conditions. Some reactions are indisputably irreversible, for example the reaction of ribulose-bisphosphate oxygenation (promiscuously catalyzed by the enzyme RuBisCO) which has a ΔrG° of ~530 kJ/mol and therefore K = 10^35. It is thus tempting to use a rule-of-thumb for determining whether a reaction is irreversible (Tanaka et al., 2003), by applying a threshold on its apparent equilibrium constant—e.g. K′ > 1000 (or K′ < 0.001 for irreversible reactions that always flow in the backward direction). This points to the fact that reactant concentrations are bound by physiological considerations and therefore a high-enough K′ would require too much of an imbalance in concentrations between substrates and products for reversing a reaction.
Aside from $K'$ itself, the reaction stoichiometry plays a crucial role in determining reversibility as well. For example, the fructose-bisphosphate aldolase reaction—fructose 1,6-P $\rightleftharpoons$ dihydroxyacetone-P + glyceraldehyde 3-P—has a $\Delta G^\circ$ of 23 kJ/mol and a $K'$ of $9.23\times10^{10}$ (Alberts, 2002). This might be considered irreversible (i.e., always flowing in the backward direction), but when the three reactants are at a concentration of 100 $\mu$M, a typical intracellular concentration for metabolites (Bennett et al., 2002), the $\Delta G^\circ$ is about 0.2 kJ/mol, very close to zero. The reason for the huge difference is due to the fact that this reaction has a different number of products and substrates, and in low-enough concentrations, the effect of the reaction quotient ($Q'$) on $\Delta G^\circ$ is significant. Therefore, any reversibility index should properly account for this stoichiometric effect. Furthermore, if a reaction involves many substrates and products, the dynamic range of $\Delta G^\circ$ given the physiological metabolite concentrations can be much wider. That is, a reaction with one substrate and one product with $\Delta G^\circ = 30$ kJ/mol is much more irreversible than a reaction with three substrates, three products, and the same $\Delta G^\circ$.

Previous studies (Feist et al., 2009) obtained binary reversibility annotations by checking if $\Delta G^\circ$ can attain positive and negative values at different physiological concentrations. Herein, we present a quantitative reversibility index, $\gamma$, which accounts for the effects of stoichiometry and physiological concentrations and defines a convenient metric for comparing the reversibility of reactions. The reversibility index is defined using the formula $\gamma = (K'/Q')^{1/N}$, where $N$ is the total number of reactants (substrates plus products) and $Q'$ is the reaction quotient at characteristic physiological concentrations. For the purpose of calculating $Q'$, all metabolites are taken to have a concentration of 100 $\mu$M (see Section 3 for details). The value of $\gamma$ signifies the fold change that all product and substrate concentrations must undergo in order to reverse a reaction. When considering a range of $\gamma < 10^{12}$, which corresponds to allowing concentrations to span three orders of magnitude around 100 $\mu$M ($\approx 3$ mM), about 55% of the reactions are found to be reversible (see the Supplementary Material for a detailed statistical analysis).

5 DISCUSSION
The advances in metabolic modeling (Hempach et al., 1999; Pfeiffer and Schuster, 2004; Overhults et al., 2009; Hatcher et al., 2011) have created a need for accurate genome-wide values for reaction thermodynamic parameters. As metabolic network models for more organisms and cells emerge, it is increasingly important to have correct predictions for acid and basic environments and for different ionic strengths. This requirement is most prominent when modeling organisms with multi-compartmented cells, and having to adjust the $\Delta G$s to the conditions in each compartment. Most data that do exist are hard to access (e.g. in out-of-print books) and cover only a limited part of the scope of required compounds and reactions. Moreover, the attempts to use group contribution to expand this coverage are based on a closed implementation and are not extendable for outside users to add new group definitions or training examples.

In this article, we explored two approaches to estimating free energies. We show that the formation energies obtained using PRC achieve a good fit to the observed data (RMSE 2.4 kJ/mol), but do not provide genome-wide coverage of metabolic reactions (~30% of E. coli model). As an alternative, PGC does provide free energy estimates for the majority (~77%) of known biochemical reactions, but with larger errors (RMSE 8.5 kJ/mol). Therefore, a combination of the two methods might be beneficial, where PRC values are used whenever possible, and PGC is used to fill the gaps. There is, however, a potential problem with this approach as combinations of reactions can become inconsistent (e.g. stoichiometrically balanced cycles can have a non-zero $\Delta G$). The challenge of combining $\Delta G^\circ$ estimations from different sources and estimation approaches in a unified and consistent manner requires an update to the methods described above (manuscript under preparation).

We hope that, with time, new measurements of reaction equilibrium constants will be published and used to improve the accuracy and coverage of both these methods (PGC and PRC). We thus join the plea of Feist et al. (2009) who published a table of compounds that contain groups with yet unknown contributions.

6 CONCLUSION
We believe that the tools and data that enable thermodynamic analysis of biochemical systems should be easily and freely accessible. In addition to supplying $\Delta G^\circ$ predictions as a table for use in metabolic models, we created a website (http://equilibrator.weizmann.ac.il) with a simple user interface that enables anyone to find reactions by chemical formula or enzyme name (Flamholz et al., 2011). The user can adapt the concentration of reactants and the conditions of the reaction.

The thermodynamics of biochemical reactions has a key role to play in our understanding and manipulation of metabolic pathways. An integrated and open framework that combines accuracy and coverage will facilitate the wide use of this fundamental constraint by physics on the biochemistry of life.

ACKNOWLEDGEMENTS
The authors thank the following people for critically reading this article, their helpful comments and encouraging remarks: Avi Mayo, Tomer Shlomi, Naama Tepper, Robert N. Goldberg, Wolfram Liebermeister, Igor Libourel and Hulda S. Haraldsdóttir.

Funding: The European Research Council [260392 - SYMPAC], The Larson Charitable Foundation, Estate of David Arthur Barton, Anthony Stalbow Charitable Trust & Stella Gelerman, Canada. E.N. is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship. A.B.-E. is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities. Dr. Milo is the incumbent of the Anna and Maurice Boukstein Career Development Chair in Perpetuity.

Conflict of Interest: none declared.

REFERENCES
Christensen, J. J. et al. (1991a) Thermodynamic pH, delta H0, delta S0, and delta Cp0 values for proton dissociation from several purines and their nucleosides in aqueous solution. Biochemistry, 9, 4807–4913.
Christensen, J. J. et al. (1991b) Thermodynamics of proton dissociation in dilute aqueous solution. Part XIV. pH, delta H0, and delta S0 values for proton dissociation from several pyrimidines and their nucleosides at 10 and 40 degrees C. J. Chem. Soc., 1643–1646.