A database of thermodynamic properties is developed, which extends a previous database of glycolysis and tricarboxylic acid cycle by adding the reactions of the pentose phosphate pathway. The raw data and documented estimations of solution properties are made electronically available. The database is determined by estimation of a set of parameters representing species-level free energies of formation. The resulting calculations provide thermodynamic and network-based estimates of thermodynamic properties for six reactions of the pentose phosphate pathway for which estimates are not available in the preexisting literature. Optimized results are made available in ThermoML format. Because calculations depend on estimated hydrogen and metal cation dissociation constants, an uncertainty and sensitivity analysis is performed, revealing 23 critical dissociation constants to which the computed thermodynamic properties are particularly sensitive.

Database URL: http://www.biocoda.org/thermo

Introduction

Reliable and self-consistent databases of thermodynamic properties for biochemical reactions are necessary for accurate analysis of biochemical systems (1–6). A recently developed database of thermodynamic properties for the reactions of glycolysis and the tricarboxylic acid cycle that was constituted from measured equilibrium data (7) represents a refinement to the Alberty database (8) in that it accounts for the ionic strength and interactions of biochemical reactants and metal cations (Mg\(^{2+}\), Ca\(^{2+}\), Na\(^{+}\) and K\(^{+}\)) in estimating the derived properties from the raw data. The database of Li et al. (7) is a framework that can be extended and refined by adding the underlying raw experimental data in the database and/or refining the underlying model assumptions. Here an updated database is developed by adding the reactions of the pentose phosphate pathway into the original database.

As in Li et al. (7), thermodynamic properties (reference \(\Delta rG^0\) and \(\Delta rH^0\)) values are estimated by minimizing the difference between model predictions and experimental data. Apparent equilibrium constants for biochemical reactions are estimated by these derived thermodynamic properties and compared to experimental data measured under non-standard conditions. The basic formulae described in Li et al. (7) are used here, which account for temperature, ionic interactions and ion binding (8–13) effects to convert standard-state reference quantities to experimental state quantities.

Input data

Database of measured equilibrium constants

Raw experimental data are obtained from original reports (14–19). As in Li et al. (7), we preferentially select studies...
rated as A- and B-quality by Goldberg et al. (20–22). Free cation concentrations and ionic strength associated with these original studies are estimated based on the conditions reported in the original sources. All data and calculations are explicitly documented in the database. As we have previously established (7), each measurement entry in the raw-data database provides the following information: (i) enzyme name (EC number); (ii) experimental temperature, pH, ionic strength, apparent equilibrium constant, free metal cation concentrations, buffer and experimental method; (iii) quality rating from Goldberg et al. (20–22); (iv) notes on experiments and strategies of estimations and approximations applied in calculations; and (v) reference information. Up-to-date versions of the experimental database are made available at the URL http://www.biocoda.org/thermo, or by contacting the authors.

**Database of reactions and estimated standard reaction enthalpies**

There are eight reactions in the pentose phosphate pathway: glucose 6-phosphate dehydrogenase (EC 1.1.1.49, G6PD), 6-phosphogluconolactonase (EC 3.1.3.31, PGL), 6-phosphogluconate dehydrogenase (EC 1.1.1.44, PGD), ribose-5-phosphate isomerase (EC 5.3.1.6, R5PI), ribulosephosphate 3-epimerase (EC 5.1.3.1, RUPE), transketolase (EC 2.2.1.1, TKL), transaldolase (EC 2.2.1.2, TAL) and transketolase 2 (EC 2.2.1.1, TKL2). The first three reactions are the oxidative pentose phosphate pathway, and the rest are the reductive pentose phosphate pathway (23). After adding these reactions into our original database, there are 33 reactions in total, which are shown in the Table 1 (first 33 reactions). Table 1 lists EC numbers, reaction names and the abbreviations employed, the reference reaction stoichiometries, and estimated standard reaction enthalpies at T = 298.15 K, I = 0 M. In the reference reaction, a superscript is used to indicate the charge of chemical species. For example, the abbreviation for references species for glucose (GLC) is distinguished from the abbreviation for the biochemical reactant GLC.

Reaction enthalpies (Δ_rH°) for two reactions in the pentose phosphate pathway (PGD and R5PI) can be estimated using van’t Hoff equation because data on apparent equilibrium constants at different temperatures are available. Neither equilibrium data at different temperatures nor prior values of Δ_rH° are available for the other six reactions of the pentose phosphate pathway; hence the symbol ‘#’ is used to denote the absence of data. For these cases, the value is set to zero in further calculations.

**Database of reactant and dissociation constants**

There are seven reactants introduced to the reactant database by adding the pentose phosphate pathway into the thermodynamic database: erythrose 4-phosphate (E4P), 6-phosphoglucono-δ-lactone (PGLT), 6-phospho-D-gluconate (PGN), ribose 5-phosphate (R5P), ribulose 5-phosphate (RUP), sedoheptulose 7-phosphate (S7P) and xylulose 5-phosphate (X5P). For the five sugar phosphates (E4P, R5P, RUP, S7P and X5P), cation dissociation constants for only R5P can be found in NIST database. Since E4P, RUP, S7P and X5P are structurally similar substances to R5P (all have similar near neighbors to the phosphate group), a pragmatic approach to estimate the necessary dissociation constants is to use the values for R5P (24). Specifically, E4P, RUP, S7P and X5P are assumed to have dissociation properties equal to those for R5P in our calculations. For PGLT, we use the pK_H value of 5.99 from Aliberty (10). For PGN, Casazza et al. (15) report that the hydrogen ion dissociation constant for the carboxylic acid is > 10^-4, while the constant for the phosphate ester is the same as that of glucose 6-P [pK_H = 5.99 (10)]. Here we arbitrarily assign the value of 4.995 to the pK_H for PGN, which is the average of 4 and 5.99. The uncertainties of these assignments are considered in the uncertainty and sensitivity analysis below.

Basic thermodynamic and ion binding data for biochemical reactants and associated reference species are listed in Table 2. Each entry in the table contains the following information: (i) detailed name of reactant; (ii) reference species abbreviation; (iii) reactant abbreviation; (iv) number of protons in reference species; and (v) the dissociation constants (provided as pK) and the corresponding dissociation enthalpies Δ_rH_w. Dissociation constants and enthalpies are tabulated at 298.15 K and 0.1 M ionic strength. The symbol ‘#’ is used to indicate absence of data. For these cases, the pK’s are assumed to be infinite (no binding) with corresponding dissociation constants equal to zero. In the calculations, all the pK and Δ_rH_w values are adjusted to a common reference state of T = 298.15 K and I = 0.

**Estimation of standard-state thermodynamic quantities**

Given the compiled raw experimental data on the reactions and reactants, the thermodynamic model is used to estimate reference Δ_rG° and Δ_rG° values for the reference reactions and species. Note that in our thermodynamic model the Δ_rG° values are adjustable parameters estimated to obtain the best fit to the biochemical equilibrium data. As in previous studies (8), values of Δ_rG° for oxidized species of certain redox pairs are arbitrarily set to zero. Thus, these values are not true free energies of the reactions of formation for these chemical species; instead, they are parameters for which the thermodynamic model makes optimal predictions for these interdependent biochemical reactions.
### Table 1. Reactions with standard reaction enthalpies $\Delta_r H^o$ estimated for $I = 0$

| EC No. | Reaction name                      | Reaction abbreviation | Reference reaction | $\Delta_r H^o$ (kcal/mol) |
|--------|------------------------------------|-----------------------|--------------------|---------------------------|
| EC 2.7.1.1 | Glucokinase                        | GLK                   | GLC$^+$ + ATP$^-$ = G6P$^+$ + ADP$^-$ + H$^+$ | -23.8$^b$             |
| EC 5.3.1.9 | Phosphoglucone isomerase           | PGI                   | G6P$^-$ = F6P$^-$ | 11.53$^b$               |
| EC 2.7.1.11 | Phosphofructokinase               | PFK                   | F6P$^-$ + ATP$^+$ = Fructose$^+$ + ADP$^-$ + H$^+$ | -9.5$^c$              |
| EC 4.1.2.13 | Fructose-1,6-biphosphate aldolase | FBA                   | F16P$^-$ = DHAP$^-$ + GAP$^-$ | 48.97$^b$             |
| EC 5.3.1.1 | Triosephosphate isomerase         | TPI                   | GAP$^+$ = DHAP$^-$ | 2.73$^d$               |
| EC 4.1.2.13 | Fructose-1,6-biphosphate aldolase 2| FBA2                  | F16P$^-$ = 2DHAP$^-$ | 51.70$^b$             |
| EC 1.2.1.12 | Glyceraldehyde-3-P dehydrogenase  | GAP                   | GAP$^+$ + HPO$_4$$^-$ + NAD$_{ox}^{-}$ = BPG$^-$ + NAD$_{red}^{-}$ + 2H$^+$ | $^a$                  |
| EC 2.7.2.3 | Phosphoglycerate kinase           | PGK                   | GAP$^+$ + HPO$_4$$^-$ + NAD$_{ox}^{-}$ + ADP$^-$ = PG3$^-$ + NAD$_{red}^{-}$ + ATP$^-$ + H$^+$ | $^a$                  |
| EC 5.4.2.1 | Phosphoglycerate mutase           | PGYM                  | PG2$^-$ = PG3$^-$ | 28.05$^b$             |
| EC 4.2.1.11 | Enolase                           | ENO                   | PG2$^-$ = PEPP$^-$ + H$_2$O$^0$ | 15.1$^b$            |
| EC 2.7.1.40 | Pyruvate kinase                   | PYK                   | PYR$^+$ + ATP$^-$ = PEPP$^-$ + ADP$^-$ + H$^+$ | 5.415$^a$           |
| EC 4.1.3.7 | Citrate synthase                  | CITS                  | OAA$^-$ = AcCoA$^0$ + H$_2$O$^0$ = CIT$^-$ + COAS$^-$ + 2H$^+$ | $^a$                  |
| EC 4.2.1.3 | Aconitrase hydratase              | ACON                  | ISCT$^-$ = CIT$^-$ | -20.0$^b$          |
| EC 1.1.1.42 | Isocitrate dehydrogenase          | IDH                   | ISCT$^-$ + NADP$_{ox}^{-}$ + H$_2$O$^0$ = AKG$^-$ + NADP$_{red}^{-}$ + CO$_2$$^+$ + 2H$^+$ | -22.1$^b$         |
| EC 6.2.1.4 | Auccinate-CoA ligase              | SCS                   | GTP$^+$ + SUCC$^+$ + COAS$^-$ + H$^+$ = GDP$^-$ + HPO$_4$$^-$ + SUCC$^+$ | -30.9$^b$         |
| EC 4.2.1.2 | Fumarate hydratase                | FUM                   | FUM$^+$ + H$_2$O$^0$ = MAL$^-$ | -13.18$^b$        |
| EC 1.1.1.37 | Malate dehydrogenase             | MDH                   | MAL$^-$ + NAD$_{ox}^{-}$ = OAA$^-$ + NAD$_{red}^{-}$ + H$^+$ | 51.29$^b$          |
| EC 2.7.4.6 | Nucleoside-diphosphate kinase     | NDK                   | ATP$^+$ + GDP$_{ox}^{-}$ = ADP$^-$ + GTP$^-$ | $^a$                  |
| EC 1.6.1.1 | NADP transhydrogenase            | NPTH                  | NAD$_{ox}^{-}$ + NADP$_{red}^{-}$ + NADP$_{ox}^{-}$ + NADP$_{red}^{-}$ | -4.1$^c$          |
| EC 1.1.1.40 | Malic enzyme                      | MLE                   | MAL$^-$ + NADP$_{ox}^{-}$ + H$_2$O$^0$ = PYR$^+$ + NADP$_{red}^{-}$ + CO$_2$$^+$ + 2H$^+$ | $^a$                  |
| EC 1.1.1.37 | Malate dehydrogenase 2           | MDH2                  | MAL$^-$ + AcCoA$^0$ + NAD$_{ox}^{-}$ + H$_2$O$^0$ = CIT$^-$ + COAS$^-$ + NAD$_{red}^{-}$ + 3H$^+$ | $^a$                  |
| EC 2.7.1.23 | NAD$^+$ kinase                    | NADK                  | ATP$^+$ + NAD$_{ox}^{-}$ = ADP$^-$ + NADP$_{ox}^{-}$ + H$^+$ | $^a$                  |
| EC 3.6.1.32 | ATPase                            | ATPS                  | ATP$^+$ + H$_2$O$^0$ = ADP$^-$ + HPO$_4$$^-$ + H$^+$ | -20.5$^b$          |
| EC 3.1.3.1 | Alkaline phosphatase/G6P hydrolysis| G6PH                 | G6P$^-$ + H$_2$O$^0$ = GLC$^+$ + Pi$^-$ | 0.91$^b$            |
| EC 6.4.1.1 | Pyruvate carboxylase              | PCL                   | PYR$^+$ + ATP$^-$ + CO$_2$$^-$ = OAA$^-$ + ADP$^-$ + Pi$^-$ | $^a$                  |
| EC 1.1.1.49 | Glucose 6-phosphate dehydrogenase | G6PD                  | G6P$^-$ + NADP$_{ox}^{-}$ + H$_2$O$^0$ = PGL$^-$ + NAD$_{red}^{-}$ + H$^+$ | $^a$                  |
| EC 3.1.3.11 | 6-Phosphofluconolactonase         | PGL                   | G6P$^-$ + NADP$_{ox}^{-}$ + H$_2$O$^0$ = PGL$^-$ + NADP$_{red}^{-}$ + 2H$^+$ | $^a$                  |
| EC 1.1.1.44 | 6-Phosphofluconate dehydrogenase  | PGD                   | PGN$^+$ + NADP$_{ox}^{-}$ + H$_2$O$^0$ = RUSP$^-$ + NADP$_{red}^{-}$ + CO$_2$$^+$ + 2H$^+$ | 37.47$^b$          |
| EC 5.3.1.6 | Ribose-5-phosphate isomerase      | RPSI                  | R5P$^-$ = RUSP$^-$ | 12.86$^b$          |
| EC 5.1.3.1 | Ribulose-phosphate 3-epimerase    | RUPE                  | R5P$^-$ + X5P$^-$ | $^a$                  |
| EC 2.2.1.1 | Transketolase                     | TKL                   | S7P$^-$ + GAP$^-$ = R5P$^-$ + X5P$^-$ | $^a$                  |
| EC 2.2.1.2 | Transaldolase                     | TAL                   | S7P$^-$ + GAP$^-$ = E4P$^-$ + F6P$^-$ | $^a$                  |
| EC 2.2.1.1 | Transketolase 2                   | TKL2                  | F6P$^-$ + GAP$^-$ = E4P$^-$ + X5P$^-$ | $^a$                  |
| EC 1.2.4.1+EC 2.3.1.12+EC 1.8.1.4 | Pyruvate dehydrogenase complex    | PDH                   | PYR$^+$ + COAS$^-$ + NAD$_{ox}^{-}$ + H$_2$O$^0$ = CO$_2$$^+$ + AcCoA$^0$ + NAD$_{red}^{-}$ + H$^+$ | $^a$                  |
| EC 1.1.1.41 | Isocitrate dehydrogenase          | IDH2                  | ISCT$^-$ + NAD$_{ox}^{-}$ + H$_2$O$^0$ = AKG$^-$ + NAD$_{red}^{-}$ + CO$_2$$^+$ + 2H$^+$ | -26.27$^d$         |
| EC 1.2.1.52 | $\alpha$-Ketoglutarate dehydrogenase | AKGDH               | AKG$^-$ + NAD$_{ox}^{-}$ + COAS$^-$ + H$_2$O$^0$ = SUCC$^+$ + NAD$_{red}^{-}$ + CO$_2$$^+$ + H$^+$ | $^a$                  |
| EC 1.3.5.1 | Succinate dehydrogenase           | SDH                   | SUC$^-$ + CO$_2$$^+$ = FUM$^+$ + CoQH$_2$ | $^a$                  |

$^a$Goldberg et al. (21, 33).
$^b$Calculated value based on experimental data at different temperatures.
$^c$Values obtained from Goldberg et al. (20, 21) where associated ionic strength is not reported.
$^d$Value calculated from sum of dependent reactions.
$^e$Value not available.
Table 2. Reactant database\(^a\)

| Reactant                        | Reference species abbreviation | Reactant abbreviation | \(N_1\) | \(pK_{H1}\) | \(pK_{H2}\) | \(pK_{H3}\) | \(pK_{M1}\) | \(pK_{M2}\) | \(pK_{M3}\) | \(pK_{H4}\) | \(pK_{H5}\) | \(pK_{H6}\) | \(pK_{H7}\) | \(pK_{H8}\) | \(pK_{H9}\) |
|--------------------------------|--------------------------------|-----------------------|---------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Acetyl-coenzyme A              | ACoA\(^0\)                    | ACoA                  | 3       | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Adenosine diphosphate          | ADP\(^-\)                     | ADP                   | 12      | 6.496      | -2          | 3.87        | 16          | 3.3         | -15         | 1.59        | -7.5        | 1.27        | -11.76      | 1           | #           |
| Adenosine triphosphate         | ATP\(^-\)                     | ATP                   | 12      | 6.71       | -2          | 3.99        | 15          | 4.28        | -18         | 2.32        | -9.6        | 1.7         | -17.52      | 1.17        | -1          |
| 1,3-Bisphosphoglycerate        | BPG\(^-\)                     | BPG                   | 4       | 7.1\(^1\)  | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Citrate                        | GT\(^-\)                       | GT                    | 5       | 5.67       | -1.9        | 4.35        | 3.1         | 3.517       | -8          | 1.8         | #           | #           | #           | 0.6         | -3.54       |
| Glycine A-SH                   | COAS\(^-\)                    | COAS                  | 0       | 8.17\(^c\) | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Carbon dioxide (total)         | CO\(_2\)\(^-\)                | CO\(_2\)_tot         | 0       | 9.9\(^c\)  | 16.1\(^c\)  | 6.15\(^c\)  | 8.27\(^c\)  | #           | #           | #           | #           | #           | #           | #           | #           |
| Dihydroxyacetone phosphate     | DHAP\(^-\)                    | DHAP                  | 5       | 5.9        | #           | #           | #           | 1.57        | #           | #           | #           | #           | #           | #           | #           |
| \(\Delta\)-fructose 6-phosphate| F6P\(^-\)                     | F6P                   | 11      | 5.89\(^a\) | -0.559\(^b\) | 1.1         | #           | 1.74\(^e\)  | -9.72\(^d\) | #           | #           | #           | #           | #           | #           |
| \(\Delta\)-fructose 1,6-phosphate| F16P\(^-\)                  | F16P                  | 10      | 6.64       | #           | 5.92        | #           | 2.7         | #           | 2.12        | #           | #           | #           | #           | #           |
| Fumarate                       | FUM\(^-\)                     | FUM                   | 2       | 4.09\(^f\) | -1.56       | 2.86        | 1.08        | #           | #           | #           | #           | #           | #           | #           | #           |
| \(\Delta\)-glucose 6-phosphate| G6P\(^-\)                     | G6P                   | 11      | 5.89\(^a\) | -0.559\(^b\) | #           | #           | 1.74\(^h\)  | -9.72\(^d\) | #           | #           | #           | #           | #           | #           |
| \(\Delta\)-glyceraldehyde 3-phosphate| GAP\(^-\)     | GAP                   | 5       | 5.27\(^c\) | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Guanine diphosphate            | GDP\(^-\)                     | GDP                   | 12      | 6.505      | -2.14       | 2.8         | #           | 3.4         | -7.1        | #           | #           | #           | #           | #           | #           |
| \(\Delta\)-glucose            | GLC\(^-\)                     | GLC                   | 12      | #          | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Guanine triphosphate           | GTP\(^-\)                     | GTP                   | 12      | 6.63\(^a\) | -3          | 2.93        | 7.1         | 4.31        | -17         | 2.31        | #           | #           | #           | #           | #           |
| Water                          | H\(_2\)O\(^0\)                | H\(_2\)O              | 2       | #          | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Isocitrate                     | ISOT\(^3\)                    | ISOT                  | 5       | 5.765      | #           | 4.29        | #           | 2.625       | #           | 1.43        | #           | #           | #           | #           | #           |
| \(\alpha\)-Ketoglutarate       | AKG\(^2\)                     | AKG                   | 4       | #          | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| Malate                         | MAL\(^-\)                     | MAL                   | 4       | 4.715      | -0.58       | 3.265       | 3.4         | 1.71        | -6.16       | 0.9\(^g\)  | #           | #           | #           | 0.18        | -2.86       |
| NAD\(_{ox}\)                   | NAD\(_{ox}\)\(^-\)           | NAD\(_{ox}\)          | 26      | #          | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| NAD\(_{red}\)                  | NAD\(_{red}\)\(^-\)          | NAD\(_{red}\)         | 27      | #          | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| NADP\(_{ox}\)                  | NADP\(_{ox}\)\(^-\)          | NADP\(_{ox}\)         | 25      | 6.255\(^h\) | #           | 3.874\(^h\) | #           | #           | #           | #           | #           | #           | #           | #           | #           |
| NADP\(_{red}\)                 | NADP\(_{red}\)\(^-\)         | NADP\(_{red}\)        | 26      | 6.255\(^h\) | #           | 3.874\(^h\) | #           | #           | #           | #           | #           | #           | #           | #           | #           |

(Continued)
Table 2. Continued.

| Reactant                        | Reference species abbreviation | Reactant abbreviation | $N_{i}^{a}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ | $\Delta_{H}^{0}$ | $pK_{i}^{c}$ |
|--------------------------------|--------------------------------|-----------------------|-------------|-------------|-----------------|-------------|-----------------|-------------|-----------------|-------------|-----------------|-------------|-----------------|-------------|-----------------|-------------|
| Oxaloacetate OAA 2             | OAA                            |                       | 2           | 3.5         | 5.24            | 2.26        | 16.62           | 1.02        |                 |              |                 |              |                 |              |                 |              |
| Orthophosphate HPO 4 2         | Pi                             |                       | 1           | 6.78        | 4.6             | 1.945       | 8.7             | 1.823        | 9.518           | 0.669        |                 |              |                 |              |                 |              |
| 2-Phospho-o-glycerate PG 2     | PG2                            |                       | 4           | 7           |                 | 3.554       |                 | 2.45         |                 | 0.5           |                 | 0.61          |                 | 0.0856        | 1.08           | 1.745         | 9.518         | 0.921         | 10.759         |
| 3-Phospho-o-glycerate PG 3     | PG3                            |                       | 4           | 6.89        |                 | 3.647       |                 | 2.21          |                 | 0.87          |                 |              |                 |              |                 |              |
| Phosphoenolpyruvate PEP 3-     | PEP                            |                       | 2           | 6.245       |                 | 3.45        |                 | 2.26         |                 | 1.08          |                 |              |                 |              |                 |              |
| Pyruvate PYR                    | PYR                            |                       | 3           | 2.26        | 12.8            |                 | 1.1             |                 |              |              |              |              |              |              |              |              |
| Succinate SUCCO 4              | SUCCO                          |                       | 4           | 5.275       | 0.41            | 4.02        | 3               | 1.355        | 0.62           | 0.43          |                 | 2.76          | 0.421           | 2.759         | 1.405          | 8.939         | 0.65           | 8.8            |
| Succinyl-CoA SUCCOA             | SUCCOA                         |                       | 4           | 3.99        |                 |              |                 |              |                 |              |                 |              |                 |              |                 |              |
| Erythrose 4-phosphate E4P      | E4P                            |                       | 9           | 6.255       | $-9.75^c$       | Z′           | #               | 1.58         | $-9.71^d$      | #             |                 |              |                 |              |                 |              |
| 6-Phosphoglucono-6-lactone PG 7| PG 7                           |                       | 11          | 5.99        |                 |              |                 |              |                 |              |                 |              |                 |              |                 |              |
| 6-Phosphogluconate PGN 5       | PGN                            |                       | 13          | 4.99        |                 |              |                 |              |                 |              |                 |              |                 |              |                 |              |
| Ribose 5-phosphate RSP 2       | RSP                            |                       | 11          | 6.255       | $-9.75^c$       | 2             # | 1.58           | $-9.71^d$    |                 |              |                 |              |                 |              |                 |              |
| Ribulose 5-phosphate RUSP 2    | RUSP                           |                       | 11          | 6.255       | $-9.75^c$       | Z′           | #               | 1.58         | $-9.71^d$      | #             |                 |              |                 |              |                 |              |
| Sedoheptulose 7-phosphate STP 7| STP                            |                       | 15          | 6.255       | $-9.75^c$       | Z′           | #               | 1.58         | $-9.71^d$      | #             |                 |              |                 |              |                 |              |
| Xylulose 5-phosphate XSP 2     | XSP                            |                       | 11          | 6.255       | $-9.75^c$       | Z′           | #               | 1.58         | $-9.71^d$      | #             |                 |              |                 |              |                 |              |

$^a$Dissociation $pK$ and $\Delta_{H}^{0}$ are reported for $T=298.15$ K and $I=0.1$ M. Unless indicated, values are the average number obtained from NIST database (27). Dissociation enthalpies are reported in units of kJ/mol. Subscripts on '$pK$' and '$\Delta_{H}$' entries are defined as follows: 'H': hydrogen, 'Mg': magnesium, 'K': potassium, 'Na': sodium, 'Ca': calcium, '1': first ion dissociation, '2': second ion dissociation, 'HMg': hydrogen ion binds to the ligand before magnesium ion binds to the ligand. '#' denotes value is not available.

$^b$Alberty (8).

$^c$Alberty (10, 34).

$^d$Tewari et al. (35).

$^e$G6P and F6P are assumed to have equivalent H$^+$ and Mg$^{2+}$-dissociation properties.

$^f$From NIST database (36) at $T=293.15$ K.

$^g$Larson-Ražnikiewicz (37).

$^h$Merrill et al. (38).

$^i$E4P, RUSP, STP and XSP are assumed to have equivalent dissociation properties as RSP (24).

$^j$Casazza et al. (15).

$^k$Briggs et al. (39).

$^l$NADP$_{red}$ is assumed to have equivalent dissociation properties as NADP$_{ox}$.
Table 3. Values of $\Delta_r G^0_i$ ($I = 298.15$ K, $I = 0$) used in this study and that were taken from Alberty (8) (Table 3.2)

| Species | $\Delta_r G^0_i$ (kJ/mol) |
|---------|--------------------------|
| $\text{ACoA}^0$ | $-188.52$ |
| $\text{ADP}^3-$ | $-1906.13$ |
| $\text{CO}_3^{2-}$ | $-527.81$ |
| $\text{GTP}^4-$ | $-2768.1$ |
| $\text{H}_2\text{O}$ | $-237.19$ |
| $\text{NAD}^+$ | $0^a$ |
| $\text{HPO}_4^{2-}$ | $-1096.1$ |
| $\text{H}^+$ | $0$ |
| $\text{CoQ}^0$ | $0^a$ |
| $\text{CoQH}_2^0$ | $-89.92$ |

*Property value is based on the arbitrary assignment of zero.

Because the calculated $\Delta_r G^0_i$ are interrelated [e.g. in the reductive pentose phosphate pathway, five carbon sugars (X5P and R5P) are converted into three carbon (GAP) and six carbon (F6P) sugars which can then be utilized by the pathway of glycolysis], the database of $\Delta_r G^0_i$ may be rigorously extended only by recalculating the entire database using all of the raw data. Therefore, model fitting is based on total 686 data entries for the network of all 33 reactions (the first 33 reactions listed in Table 1). Standard free energies of formation for all the reference species are unable to be estimated independently because there are 29 stoichiometrically independent reactions and in total 40 reactants in our system. There are four reactions PDH, IDH2, AKGDH and SDH (last four reactions in Tables 1 and 4) in the TCA cycle for which direct measurements are not available in the literature. For these reactions, $\Delta_r G^0_i$ values are calculated using Alberty’s database (8) and set as constraints to perform the optimization of the thermodynamic model. IDH2 is not an independent reaction in the overall network of 33 reactions. Therefore, values of $\Delta_r G^0_i$ for 32 references species may be estimated from data on the 29 reactions with the three constraints. The values of $\Delta_r G^0_i$ for eight species are set to fixed values which are obtained from Alberty (8), as shown in Table 3 (values of $\Delta_r G^0_i$ for CoQ$^0$ and CoQH$_2^0$ do not come into these calculations). A constrained nonlinear optimization procedure with the fmincon solver (Mathworks, Inc.) is used to analyze the whole data set. By weighting in inverse proportion to the number of data points available for a given reaction and minimizing the difference between model predictions and experimental data, a simultaneous solution of standard reaction Gibbs energies is obtained for the entire data set.

Results

Estimated Gibbs free energies of reaction and formation

Figure 1 illustrates model predictions versus experimental data for all data used in the analysis. The predicted $K'$ values are obtained based on accounting for the biochemical state associated with a given experimental measurement in the raw-data database. Data points in the pentose phosphate pathway are shown as filled squares. Other data points are shown as open squares. The data span almost 14 orders of magnitude, and reveal good agreement between model predictions and measured data.

The optimal estimates of $\Delta_r G^0_i$ listed in Table 4 ($\Delta_r G^0_i$), compared to the values from Li et al. (7) ($\Delta_r G^0_i$). $\Delta_r G^0_i$ of first 25 reactions are almost the same as the results obtained previously (7). For the remaining eight reactions of the pentose phosphate pathway, since the $\Delta_r G^0_i$ of the reference species for E4P, PGLT, PGN and S7P are not available in the Alberty (8) or Goldberg’s database (25), the symbol ‘#’ is used to denote the lack of a value. For reactions R5PI and RUPE, the absolute difference between our current model predictions and the values reported by Goldberg et al. (25) is within a reasonable margin, 0.21 and 0.07 kJ/mol, respectively. Predicted $K'$ versus experimental measurements are plotted in Figure 2 for these two reactions (R5PI and RUPE). The current database and the Goldberg database yield similar results. Recall that the dissociation properties of R5P, RUSP and XSP are assumed to be the same; the binding polynomials for the reactants on the left- and right-hand sides of these reactions are identical.
Table 4. Optimal reaction free energies for reference chemical reactions (ΔrG°) and for biochemical reactions (ΔrG°) under physiological conditions

| EC No. | Reaction | ΔrG° | ΔrG° | ΔrG° | T | I | pH | P |
|--------|----------|------|------|------|---|---|----|---|
| EC 2.7.1.1 | GLK 16.19 16.03 | –19.22 –35.41 | 1.61 1.57 | –41.56 2.97 |
| EC 2.7.1.9 | PGI 3.12 3.13 | 2.78 –0.34 | –0.34 0 0 0 |
| EC 2.7.1.11 | PFK 26.79 26.79 | –15.62 –42.41 | 1.46 –4.70 –41.56 2.40 |
| EC 4.1.2.13 | FBA 18.79 18.80 | 24.64 5.85 | –1.21 6.26 0 0.81 |
| EC 5.3.1.1 | TPI –7.01 –7.01 | –7.57 –0.56 | –0.39 0 0 –0.17 |
| EC 4.1.2.13 | FBA2 11.79 11.79 | 17.07 5.28 | –1.61 6.26 0 0.64 |
| EC 1.2.1.12 | GAP 51.37 51.37 | 2.60 –48.77 | 2.07 –9.39 –41.56 0.12 |
| EC 2.7.2.3 | PGK 34.37 34.37 | –19.00 –53.37 | 1.38 –9.39 –41.56 –7.97 |
| EC 5.4.2.1 | PGYM –5.89 –5.90 | –6.35 –0.46 | –1.37 0 0 0.91 |
| EC 4.2.1.11 | ENO –4.54 –4.53 | –4.47 0.07 | –0.79 0 0 0.86 |
| EC 2.7.1.40 | PYK 66.91 66.90 | 27.18 –39.73 | 2.48 –1.57 –41.56 0.92 |
| EC 4.1.3.7 | OCS 60.16 60.32 | –36.60 –96.76 | 2.42 –6.26 –83.13 –9.80 |
| EC 4.2.1.3 | ACON –5.75 –5.76 | –7.58 –1.83 | 0.57 0 0 –2.41 |
| EC 1.1.1.42 | IDH 97.05 97.06 | –3.29 –100.34 | 4.80 –6.26 –83.13 15.76 |
| EC 6.2.1.4 | SCS –56.55 –56.56 | 0.07 56.62 | –1.03 6.26 41.56 9.82 |
| EC 4.2.1.2 | FUM –3.38 –3.38 | –3.52 –0.14 | 0.39 0 0 –0.53 |
| EC 1.1.1.37 | MDH 71.41 71.09 | 28.04 –43.37 | 0.81 –3.13 –41.56 0.52 |
| EC 2.7.4.6 | NDK 0.02 0.01 | –0.56 –0.58 | 0 0 0 –0.58 |
| EC 1.6.1.1 | NPTH –3.53 –3.58 | –0.41 3.12 | 0.02 3.13 0 –0.03 |
| EC 1.1.1.40 | MLE 104.53 103.45 | 2.00 –102.53 | 4.21 –7.83 –83.13 –15.78 |
| EC 1.1.1.37 | MDH2 131.57 131.41 | –6.50 –138.07 | 5.30 –9.39 –124.69 9.28 |
| EC 2.7.1.23 | NADK 29.11 26.90 | –9.95 –39.06 | 1.17 –1.57 –41.56 2.89 |
| EC 3.6.1.32 | ATP5 4.99 4.67 | –32.42 –37.41 | 1.03 1.57 –41.56 1.56 |
| EC 3.1.3.1 | G6PH –11.20 –11.36 | –13.10 –1.90 | –0.49 0 0 –1.41 |
| EC 6.4.1.1 | PCL –24.60 –24.12 | –4.57 20.03 | –0.99 3.13 0 17.89 |
| EC 1.1.1.49 | G6PD 38.69 #d | –7.51 –46.20 | 1.56 –6.26 –41.56 0.07 |
| EC 3.1.1.31 | PGL 69.16 #d | –21.89 –91.05 | 2.78 –10.96 –83.13 0.25 |
| EC 1.1.1.44 | PGD 104.64 #d | 1.23 –103.41 | 2.70 –6.26 –83.13 –16.73 |
| EC 5.3.1.6 | R5P 0.99 1.2a | 0.52 –0.47 | –0.47 0 0 0 |
| EC 5.1.3.1 | RUPE –1.28 –1.21a | –1.33 –0.05 | –0.05 0 0 0 |
| EC 2.2.1.1 | TKL 1.58 #d | 1.23 –0.35 | 0.06 0 0 –0.41 |
| EC 2.2.1.2 | TAL 2.72 #d | 2.63 –0.09 | 0.11 0 0 –0.20 |
| EC 2.2.1.1 | TAL2 8.98 #d | 8.71 –0.27 | 0.36 0 0 –0.63 |
| EC 1.2.4.1+EC 2.3.1.12+EC 1.8.1.4 | PDH 15.78c | 17.50 –39.26 –55.04 | 0.64 –4.70 –41.56 –9.43 |
| EC 1.1.1.41 | IDH2 93.52 93.48 | –3.70 –97.22 | 4.82 –3.13 –83.13 –15.79 |
| EC 1.2.1.52 | AKGDH 15.85c | 15.28 –37.66 –53.51 | 0.64 –3.13 –41.56 –9.45 |
| EC 1.3.5.1 | SDH –1.35d | –3.10 –0.59 | 0.76 –0.05 0 0.81 |

The physiological conditions (26) are as follows: T = 310.15 K, I = 0.18 M, pH = 7, [Mg2+] = 0.8 mM, [K+] = 140 mM, [Na+] = 10 mM, [Ca2+] = 0.001 mM. The unit of the free energy is kJ/mol.

ΔrG° is the optimal reaction free energies for reference chemical reactions in this work, ΔrG° is the optimal reaction free energies for reference chemical reaction in Li et al. (7).

This column lists the actual values of each non-standard physiological condition contribution (in kJ/mol) to the difference between the physiological free energies (ΔrG°) and the standard free energies (ΔrG°). T denotes the temperature contribution, I denotes the ionic strength contribution, pH denotes the pH contribution and P denotes the binding polynomial contribution.

Value is not available.

Calculated from Goldberg’s database (25): the ΔrG° values for RSP, RUSP and XSP in Goldberg’s database are −1582.57, −1581.37 and −1582.58 kJ/mol, respectively.

Values calculated from Alberts’s database (8) are used as model constraints for which there are no equilibrium data in the raw-data database. Since the ΔrG° values for CoQ and CoQH2 are not predicted here, these values are set to 0 and −89.92 kJ/mol, respectively [from Alberts (8)] to computed ΔrG° for the SDH reaction.
Thus, the computed apparent equilibrium constants for these reactions do not depend on the experimental ionic composition. The validity of this assumption will be considered in the uncertainty and sensitivity analysis below. In addition, the predicted thermodynamic properties are not affected by the ionic strength for these reactions.

Optimal predicted $\Delta_f G^0$ associated with the $\Delta_i G^0$ predictions are listed in Table 5, compared to the optimal values of our previous version of the database (7). After the seven reactants of the pentose phosphate pathway and three constraints ($\Delta_i G^0$ of reaction PDH, AKGDH and SDH) are introduced, most of the estimated values of $\Delta_i G^0$ are shifted substantially. However, from Table 4 we can see that these shifts do not change the predicted $\Delta_i G^0$ compared to previous predictions. This is because the shifts in $\Delta_i G^0$ do not change the optimized results (apparent $\Delta_i G^0$ or $K'$) of our model. Recall that here the estimated $\Delta_i G^0$ values represent parameters in a thermodynamic model for this set of interdependent biochemical reactions. Since the number of independent reactions for which data exist (29 reactions) is smaller than the number of $\Delta_i G^0$ values to represent the system (40 reactions), values for several reference species are set to either existing values or values reported elsewhere (Table 3). As a result, these $\Delta_i G^0$ values are not physical constants. Rather, they are parameters in a thermodynamic network model that together form a self-consistent picture of the thermodynamics of reactions of the set of reactants studied here.

The robustness of these calculations is checked by repeating the optimization with the $\Delta_i G^0$ for a single species constrained to a value $\pm 10\%$ different from the estimated optimal value. The degree to which the experimental data can be matched with one $\Delta_i G^0$ value $10\%$ different from the optimal values reported in Table 5 provides a measure of the sensitivity of the estimate. We define a sensitivity measure $S_i$ for the estimate of $\Delta_i G^0$ as

$$E = \sum_{i=1}^{M} \left[ \frac{1}{N_i} \sum_{j=1}^{N_i} \left( \Delta_i G^0_{\text{model}} - \Delta_i G^0_{\text{exp}} \right)^2 \right],$$

$$S_i = \frac{\max \left( \left| E_i(\Delta_i G^0) \pm 0.1 \Delta_i G^0 \right| - E_i(\Delta_i G^0) \right)}{0.1 E_i(\Delta_i G^0)},$$

where $E(\Delta_i G^0)$ is the optimal value of the error function (for values listed in Table 5) and $E(\Delta_i G^0 \pm 0.1 \Delta_i G^0)$ is the error with $\Delta_i G^0$ set to a 90% or 110% of its optimal value, $M$ is the number of reactions, and $N_i$ is the number of experimental measures for each reaction. Sensitivity values are listed in Table 5 for each species, revealing that estimates of $\Delta_i G^0$ for GLC$^0$, NAD$^+$, PYR$^-$, AKG$^{$2$-}$, SUC$^2$-, FUM$^2$- and COAS$^0$ are not highly sensitive to the data.

**Predicted apparent Gibbs free energies under physiology conditions**

The fifth column in Table 4 reports the predicted apparent $\Delta_r G^0$ ($= -RT \ln K'$) at physiological conditions representative of a muscle cell (26) ($T = 310.15$ K, $I = 0.18$ M, pH = 7, $[\text{Mg}^{2+}] = 0.8$ mM, $[\text{K}^+] = 140$ mM, $[\text{Na}^+] = 10$ mM, $[\text{Ca}^{2+}] = 0.0001$ mM). Differences between the physiological free energies ($\Delta_r G^0$) and the standard free energies ($\Delta_i G^0$) are also shown in Table 4. These differences come from the effects of non-standard physiological conditions—temperature ($T$), ionic strength ($I$), pH and cation bindings. Therefore, the physiological free energy $\Delta_i G^0$ can be expressed by the summation of the standard free

![Figure 2](image-url). Model-predicted $K'$ versus experimental $K'$ for the ribose-5-phosphate isomerase and ribulose-phosphate 3-epimerase reactions. Open circles (GT-based $K'$) are computed based on the Goldberg’s database (25); filled squares (optimized $K'$) are computed based on the optimized values of $\Delta_i G^0$ from Table 5.
Table 5. Optimal predicted $\Delta_r G_0^0$ for reference species ($T = 298.15$ K, $I = 0$ M)

| No. | Species | $\Delta_r G_{f_{00}}^0$ | $\Delta_r G_{f_{0n}}^0$ | $|\Delta_r G_{f_{00}}^0 - \Delta_r G_{f_{0n}}^0|$ | Sensitivity |
|-----|---------|---------------------|---------------------|---------------------------------|-------------|
| 1   | GLC$^0$ | −719.37             | −916.39             | 197.02                          | 0.33        |
| 2   | ATP$^{3-}$ | −2770.03           | −2769.71            | 0.32                            | 56099.52    |
| 3   | F6P$^{2-}$ | −1563.96           | −1760.81            | 196.85                          | 1.52        |
| 4   | F16P$^{4-}$ | −2401.08           | −2597.60            | 196.52                          | 3.55        |
| 5   | DHAP$^{2-}$ | −1194.65           | −1292.91            | 98.26                           | 3.51        |
| 6   | GAP$^{3-}$ | −1187.64           | −1285.90            | 98.26                           | 3.49        |
| 7   | BPG$^{4-}$ | −2276.28           | −2354.55            | 78.27                           | 19.44       |
| 8   | NAD$_{red}^{2-}$ | 43.91              | 23.91               | 20.00                           | 0.14        |
| 9   | PG$^{3-}$ | −1429.38           | −1507.96            | 78.58                           | 7.83        |
| 10  | PG$^{3-}$ | −1423.49           | −1502.06            | 78.57                           | 7.83        |
| 11  | PEP$^{3-}$ | −1190.84           | −1269.40            | 78.56                           | 5.54        |
| 12  | PYR$^{-}$ | −393.85            | −472.72             | 78.87                           | 0.64        |
| 13  | OAA$^{2-}$ | −714.06            | −792.13             | 78.07                           | 2.06        |
| 14  | CIT$^{3-}$ | −1022.44           | −1157.52            | 135.08                          | 1.38        |
| 15  | ISCIT$^{3-}$ | −1016.69           | −1151.76            | 135.07                          | 1.38        |
| 16  | NAD$_{ox}^{3-}$ | −834.79           | −836.68             | 1.89                            | 1837.57     |
| 17  | AKG$^{2-}$ | −676.45            | −791.57             | 115.12                          | 0.83        |
| 18  | SUC$^{2-}$ | −589.56            | −685.56             | 96.00                           | 0.91        |
| 19  | GDP$^{3-}$ | −1904.22           | −1904.53            | 0.31                            | 9560.62     |
| 20  | FUM$^{2-}$ | −500.99            | −598.74             | 97.75                           | 0.67        |
| 21  | MAL$^{2-}$ | −741.56            | −839.30             | 97.74                           | 1.44        |
| 22  | GAP$^{3-}$ | −1567.08           | −1763.94$^b$        | 196.86                          | 1.52        |
| 23  | COA$^0$ | −57.17             | 0$^b$               | 57.17                           | 0.03        |
| 24  | NAD$_{red}^{4-}$ | −787.35           | −809.19$^b$         | 21.84                           | 41.94       |
| 25  | SUCCoA$^{-}$ | −471.06            | −509.59$^b$         | 38.53                           | 3.47        |
| 26  | E4P$^{2-}$ | −1306.62           | #                   | #                               | 2.39        |
| 27  | PGLT$^{2-}$ | −1575.83           | #                   | #                               | 1.85        |
| 28  | PGN$^{3-}$ | −1782.55           | #                   | #                               | 2.34        |
| 29  | R5P$^{2-}$ | −1435.72           | #                   | #                               | 1.86        |
| 30  | RUSP$^{2-}$ | −1434.72           | #                   | #                               | 1.86        |
| 31  | S7P$^{2-}$ | −1685.66           | #                   | #                               | 1.31        |
| 32  | X5P$^{2-}$ | −1436.00           | #                   | #                               | 1.85        |

$\Delta_r G_{f_{00}}^0$ is the optimal free energies of formation in this work, $\Delta_r G_{f_{0n}}^0$ is optimal free energies of formation in Li et al. (7). The unit of the free energy is kJ/mol.

$\Delta_r G_0^0$ of species is set as fixed value which is obtained from Alberty's database as in Li et al. (7).

The energy $\Delta_r G_0^0$ and the non-standard contributions $\Delta_T$, $\Delta_I$, $\Delta_{ph}$ and $\Delta_P$:

$$\Delta_r G_0^0(T, I) = -RT \ln K'(T, I) = \Delta_r G_0^0(T_0, I_0) + \Delta_T + \Delta_I + \Delta_{ph} + \Delta_P,$$

where $T_0 = 298.15$ K, $I_0 = 0$ M, $K'$ is the apparent equilibrium constant for the associated biochemical reaction, $\Delta_T$ denotes the temperature contribution, $\Delta_I$ denotes the ionic strength contribution, $\Delta_{ph}$ denotes the pH contribution and $\Delta_P$ denotes the binding polynomial contribution.

An example calculation [for reaction GLK (EC 2.7.1.1)] is described in the Appendix A.

The contributions of non-standard elements to the differences between the standard free energy for the reference chemical reaction ($\Delta_r G_0^0$) and the physiological energy values ($\Delta_r G_0^0$) are listed in the last four columns of Table 4. The results can be divided to three cases: (i) if the stoichiometric coefficient of $H^+$ is non-zero in the reference chemical reaction ($v_H \neq 0$), the differences are substantial and pH contributes significantly, accounting for at least 73% of the difference, e.g. reactions GLK, PFK and GAP; (ii) if $v_H = 0$
\( \Delta pH = 0 \) and \( \sum v_i z_i^2 = 0 \) \( \Delta u = 0 \), the differences tend to be relatively small compared with case (i) and temperature and/or binding polynomials contribute most to the differences, e.g. reactions G6PD, PGL and PGD show substantial difference between free energy for the reference chemical reaction \((\Delta_r G^0)\) and for biochemical reaction \((\Delta_r G^0)\) under physiological conditions. These three reactions have \( v_i \neq 0 \). They are the oxidative portion of the pentose phosphate pathway (15, 23), which produce NADP_{red} and are essentially 'irreversible' in vivo (23).

**Dissociation constants uncertainty and sensitivity analysis**

The \( pK \) values listed in Table 2 are taken as the average value when there are several values \((\geq 2)\) available in NIST database (27). For these \( pK \) values, the average value may not represent the best choice to be used in the model, i.e. some value among those available values may be more accurate than others. For some \( pK \) values, there exists only one estimate or no direct estimates. In order to predict the impact of uncertainty of these values on the model output, an uncertainty and sensitivity analysis is performed.

The following equation is used as a measure of uncertainty in a \( pK \) value when several independent measures are available:

\[
U = u_i = \frac{pK_{i,\text{max}} - pK_{i,\text{min}}}{pK_i},
\]

(4)

where \( pK_{i,\text{max}} \) and \( pK_{i,\text{min}} \) refer to the maximum and minimum value of \( pK_i \), respectively. Table 6 shows the computed uncertainties for these \( pK \)s.

| Reactant abbreviation | \( U_{pK_{i,\text{H}}} \) | \( U_{pK_{i,\text{L}}} \) | \( U_{pK_{\text{Mg1}}} \) | \( U_{pK_{\text{Ca1}}} \) |
|----------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| ADP                  | 0.0539                  | 0.0545                  | 0.0174                  | 0.0174                  |
| ATP                  | 0.0849                  | 0.1051                  | 0.1215                  | 0.1215                  |
| CIT                  | 0.0423                  | 0.0597                  | 0.0480                  | 0.0480                  |
| GDP                  | 0.0630                  | 0.0799                  | 0.0835                  | 0.0835                  |
| DHAP                 | 0.0052                  | 0.0153                  | 0.0052                  | 0.0052                  |
| GTP                  | 0.0148                  | 0.0153                  | 0.0052                  | 0.0052                  |
| PG2                  | 0.0251                  | 0.0463                  | 0.1975                  | 0.1975                  |
| PG3                  | 0.0133                  | 0.0149                  | 0.1433                  | 0.1433                  |
| PEP                  | 0.0133                  | 0.0149                  | 0.1433                  | 0.1433                  |
| R5P                  | 0.0133                  | 0.0149                  | 0.1433                  | 0.1433                  |

When only one \( pK \) value estimate is available, the uncertainty is defined as the average number \( \sigma \) of all calculated \( u_i \):

\[
U = \sigma = \sum u_i / n
\]

(5)

According to Table 6, \( \sigma \) is equal to 0.0609.

The sensitivities of the computed thermodynamic database due to a 10% change of \( pK \) values are calculated (28):

\[
S_i = \frac{\partial E_i(x_i)}{\partial x_i} = \frac{\max(|E_i(x_i + \pm 0.1x_i) - E_i(x_i)|)}{0.1 E_i(x_i)}
\]

(6)

where \( E \) is shown in equation (1), and \( x_i \) is the value of the \( i^{th} \) \( pK \). Table 7 lists calculated sensitivities >0.1. Sensitivity values for all potassium and sodium ion dissociation constants are <0.1. Calcium ion dissociation constants are not included in these calculations because \([Ca^{2+}] = 0 \) for all reactions in our raw-data database. Thus calculated sensitivities of calcium ion dissociations are equal to 0.

The product \( US \) combining uncertainty \( U \) and sensitivity \( S \) can be used to check the overlapping effect of uncertainty and sensitivity. For example, recall that we arbitrarily assign the value of 4.995 to the \( pK_{i,\text{H}} \) for PGN. Since the value is not available in NIST database, its uncertainty \( U \) is set to the average number 0.0609. If we consider the theoretical range of 4.5-5.99 discussed above, then the

| Reactant abbreviation | \( S_{pK_{i,\text{H}}} \) | \( S_{pK_{i,\text{L}}} \) | \( S_{pK_{\text{Mg1}}} \) | \( S_{pK_{\text{Ca1}}} \) |
|----------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| ADP                  | 0.4936                  | 0.3696                  | 0.1709                  | 0.1709                  |
| ATP                  | 0.2151                  | 0.2665                  | 0.2210                  | 0.2210                  |
| CIT                  | 0.0602                  | 0.0602                  | 0.0602                  | 0.0602                  |
| COAS                 | 0.5344                  | 0.5344                  | 0.5344                  | 0.5344                  |
| DHAP                 | 0.6591                  | 0.3772                  | 0.3772                  | 0.3772                  |
| F16P                 | 1.1839                  | 0.3181                  | 1.3545                  | 1.3545                  |
| G6P                  | 0.6372                  | 0.4524                  | 0.4524                  | 0.4524                  |
| GDP                  | 0.1873                  | 0.1873                  | 0.1873                  | 0.1873                  |
| GTP                  | 0.1310                  | 0.1310                  | 0.1310                  | 0.1310                  |
| ISCIT                | 0.1230                  | 0.1230                  | 0.1230                  | 0.1230                  |
| NADP_{ox}            | 1.7916                  | 1.7916                  | 1.7916                  | 1.7916                  |
| NADP_{red}           | 0.7611                  | 0.7611                  | 0.7611                  | 0.7611                  |
| Pi                   | 0.1248                  | 0.2015                  | 0.2015                  | 0.2015                  |
| PG2                  | 1.2172                  | 1.2172                  | 1.2172                  | 1.2172                  |
| PG3                  | 0.3698                  | 0.3698                  | 0.3698                  | 0.3698                  |
| PEP                  | 0.2293                  | 0.2293                  | 0.2293                  | 0.2293                  |
| R5P                  | 0.1818                  | 0.1818                  | 0.1818                  | 0.1818                  |
| R5P                  | 0.1007                  | 0.1007                  | 0.1007                  | 0.1007                  |
| S7P                  | 0.1433                  | 0.1433                  | 0.1433                  | 0.1433                  |
| X5P                  | 0.1597                  | 0.1597                  | 0.1597                  | 0.1597                  |


calculated uncertainty \( U \) is \(-0.4\). For this case, because the computed \( \text{pK}_{H1} \) sensitivity is 0.0225, the \( US \) product is \(<0.01\), which is small enough that the value of \( \text{pK}_{H1} \) for PGN has no substantial effect on the model output. Similarly, recall that E4P, RUSP, S7P and X5P are assumed to have dissociation properties equivalent to R5P in our calculations. Although the computed \( \text{pK} \) sensitivities of E4P, RUSP, S7P and X5P are 3.81E-6, 0.1007, 0.1433 and 0.1597, respectively, all the \( US \) products of the \( \text{pK}_{H1} \) for these reactants are \(<0.01\). Therefore, the assumption of equivalent dissociation properties for E4P, R5P, RUSP, S7P and X5P does not substantially effect our calculations.

Figure 3A illustrates that \( US \) products span eight orders of magnitude. Figure 3B illustrates the detailed distribution of the \( US \) products \( >0.01\). All \( US \) products are \(<0.11\). There are 23 cases for which \( US >0.01\). These 23 \( US \) values belong to 15 reactants and four \( \text{pKs} \) as listed in Table 8. They are a subset of the \( \text{pKs} \) with sensitivity \( >0.1\). They demonstrate the most important \( \text{pKs} \) which can make obvious impact on the model output. The largest four \( US \) values are \( \text{pK}_{NADPox}^{H1} \), \( \text{pK}_{F16P}^{H1} \), \( \text{pK}_{PG2}^{H1} \) and \( \text{pK}_{PG3}^{H1} \), as indicated in Figure 3B.

**Database dissemination**

ThermoML is an extensible markup language (XML)-based approach, which is an IUPAC standard for storage and exchange of thermodynamic property data (29–32). Our optimized results are stored in the standard ThermoML format with two small extensions to the current ThermoML schema (32): (i) adding ‘pseudo-Gibbs free energy of formation, \( \text{kJ/mol} \)’ in the list of ePropName in BioProperties of PureOrMixtureData; and (ii) adding ‘biochemical network calculation’ in ePredictionType of Prediction. We use the term ‘pseudo-Gibbs free energy of formation’ because in our thermodynamic model the estimated \( \Delta_G^0 \) values represent adjustable parameters for the given set of inter-dependent biochemical reactions. In our ThermoML data files, the abbreviation name of reactant is also added in sCommonName in Compound. In the ThermoML reaction database, if the value of \( \Delta, H^0 \) is not available, it is specified in sPredictionMethodDescription, and both nPropValue and nPropDigits are set to 0 in PropertyValue of NumValues. The ThermoML schema and our ThermoML data files are provided in Supplementary Data.

**Table 8.** The product \( US >0.01 \) in dissociation constants uncertainty and sensitivity analysis

| Reactant abbreviation | pKs       |
|-----------------------|-----------|
| ADP                   | \( \text{pK}_{\text{ADP}}^{H1}, \text{pK}_{\text{ADP}}^{Mg1}, \text{pK}_{\text{ADP}}^{Mg2} \) |
| ATP                   | \( \text{pK}_{\text{ATP}}^{H1}, \text{pK}_{\text{ATP}}^{Mg1}, \text{pK}_{\text{ATP}}^{Mg2} \) |
| CIT                   | \( \text{pK}_{\text{CIT}}^{Mg1} \) |
| COAS                  | \( \text{pK}_{\text{COAS}}^{H1} \) |
| DHAP                  | \( \text{pK}_{\text{DHAP}}^{Mg1} \) |
| F16P                  | \( \text{pK}_{\text{F16P}}^{H1}, \text{pK}_{\text{F16P}}^{Mg1}, \text{pK}_{\text{F16P}}^{Mg2} \) |
| G6P                   | \( \text{pK}_{\text{G6P}}^{Mg1} \) |
| GDP                   | \( \text{pK}_{\text{GDP}}^{Mg1} \) |
| GTP                   | \( \text{pK}_{\text{GTP}}^{Mg1} \) |
| ISCIT                 | \( \text{pK}_{\text{ISCIT}}^{Mg1} \) |
| Pi                    | \( \text{pK}_{\text{Pi}}^{Mg1} \) |
| NADP\_ox              | \( \text{pK}_{\text{NADP}}^{H1} \) |
| NADP\_red             | \( \text{pK}_{\text{NADP}}^{H1} \) |
| PG2                   | \( \text{pK}_{\text{PG2}}^{H1} \) |
| PG3                   | \( \text{pK}_{\text{PG3}}^{H1} \) |
Discussion

We have updated our biochemical thermodynamic database by adding the reactions of the pentose phosphate pathway. To build the new database, each original publication has been studied on a case-by-case basis, case-specific assumptions and approximations determined and documented, and case-specific calculations performed and documented. Raw data and documented estimations of solution properties are made electronically available so that the updated database remains transparent and extensible.

The developed database is optimally self-consistent and consistent with the data available for reactions considered and the constraints. Theoretical predictions of apparent equilibrium constants optimally match experimental data on equilibrium constants. These apparent equilibrium constants are predicted based on the estimated species-level Gibbs free energies of formation and accounting for the effects of temperature, ionic interactions and hydrogen and metal cation binding. The new database provides thermodynamic and network-based estimates of thermodynamic properties for six reactions of the pentose phosphate pathway for which estimates are not available in the pre-existing literature.

These calculations demonstrate how network thermodynamic calculations are effectively extended. Adding raw experimental data on reactions and reactants of the pentose phosphate pathway into corresponding raw-data database, reaction database and reactant database, respectively, a previous thermodynamic network model was extended to include these elements.

Although the new optimal estimates of $\Delta G^0$ for reactants of glycolysis and tricarboxylic acid cycle are equal to previous estimates (7), most of the optimal predicted $\Delta r G^0$ values which are associated with the $\Delta G^0$ predictions are shifted compared to the previous version. This result demonstrates that reliable and self-consistent extensions require the recalculation of entire set of species-level parameters. Furthermore, sensitivity analysis reveals that some formation energies can vary substantially without changing the optimized objective value significantly. While this set of optimal values represents a self-consistent set, combining these estimated formation energies with independently estimated parameters from other studies would require re-optimizing a combined database.

The uncertainty and sensitivity analysis of dissociation constants reveals 23 pKs most important to the model output. Additional experimental measurements of these parameter values are desirable.

Supplementary Data

Supplementary data are available at Database Online.
Appendix A

Calculating the physiological free energy \( \Delta r G^0 \)

The standard free energy \( \Delta r G^0 \) can be calculated based on the equilibrium constant \( K \), for a reference chemical reaction:

\[
\Delta r G^0(T_0, I_0) = -RT_0 \ln(K(T_0, I_0)),
\]

where \( T_0 = 298.15 \text{ K}, I_0 = 0 \text{ M}, R = 8.3145 \text{ J K}^{-1} \text{ mol}^{-1}. \)

The apparent equilibrium constant, \( K' \), for the associated biochemical reaction can also be calculated based on \( K \):

\[
K' = K[H^+]^{-\nu_1} \prod_{j=1}^{N} P_j^{\nu_j},
\]

where \( \nu_j \) is the stoichiometric coefficient associated with \( H^+ \) in the reference reaction, \( P_j \) is the binding polynomial associated with species \( j \) and \( \nu_j \) is the stoichiometric coefficient of species \( j \). If assuming that the activity coefficient for hydrogen ion is equal to 1, i.e. \( \text{pH} = -\log_{10}([H^+]) \), the physiological free energy \( \Delta r G^0 \) can then be obtained based on equations (A.1) and (A.2):

\[
\Delta r G^0(T,I) = -RT \ln(K'(T,I)) = -RT \ln(K(T,I)) - 2.303RT \times \nu_1 \times \text{pH} - RT \ln \left( \prod_{j=1}^{N} P_j^{\nu_j} \right).
\]
The first term in the right-hand side of equation (A.3) can be separated into four terms as shown in equation (A.6) based on equations (A.4) and (A.5) (7):

\[
pK(T, I) = pK(T_0, I_0) + \frac{\Delta_H^p(l_0)}{2.303 T_0} \left( \frac{1}{T} - \frac{1}{T_0} \right) + \\
\alpha(T) \left( \frac{1}{1 + B_0} - \frac{1}{1 + B I / T} \right) \sum v_i z_i^2,
\]

where \( pK \) is the negative base-10 logarithm of equilibrium constant \( K \), \( \Delta_H^p(l_0) \) is the standard reaction enthalpy at ionic strength \( I_0 \), \( z_i \) is the valence of species \( i \), \( B \) is an empirical constant taken to be 1.6 M \( \cdot \) \( \text{mol}^{-1} \), and \( \alpha(T) \) is the coefficient in the Debye equation \( \gamma = -\alpha z^2 / T \).

The non-standard element (temperature, ionic strength, \( pH \) and binding polynomial) contributions \( \Delta_T, \Delta_l, \Delta_{\text{ph}} \) and \( \Delta_p \) are defined as follows:

\[
\Delta_T = RT \ln K(T_0, I_0) \left( \frac{T_0}{T} - 1 \right) + \Delta_H^p(l_0) \left( \frac{1}{T} - \frac{1}{T_0} \right) + \\
RT \alpha(T) \left( \frac{1}{1 + B_0} - \frac{1}{1 + B I / T} \right) \sum v_i z_i^2,
\]

\[
\Delta_l = RT \alpha(T) \left( \frac{1}{1 + B_0} - \frac{1}{1 + B I / T} \right) \sum v_i z_i^2,
\]

\[
\Delta_{\text{ph}} = -2.303 RT \times \nu_i \times \text{pH},
\]

\[
\Delta_p = -RT \ln \left( \prod_{i=1}^{N} P_i \right).
\]

Then the physiological free energy \( \Delta_r G^0 \) can be expressed by the summation of the standard free energy \( \Delta_r G^0 \) and the non-standard element contributions \( \Delta_T, \Delta_l, \Delta_{\text{ph}} \) and \( \Delta_p \):

\[
\Delta_r G^0(T, I) = \Delta_r G^0(T_0, I_0) + \Delta_T + \Delta_l + \Delta_{\text{ph}} + \Delta_p.
\]

Example

The calculations the physiological free energy \( \Delta_r G^0 \) for reaction glucokinase GLK (EC 2.7.1.1) is illustrated as an example. As mentioned in the article, the physiological conditions are \( T = 310.15 \text{ K}, \ I = 0.18 \text{ M}, \ \text{pH} = 7, \ [\text{Mg}^{2+}] = 0.8 \text{ mM}, \ [\text{K}^-] = 140 \text{ mM}, \ [\text{Na}^+] = 10 \text{ mM}, \ [\text{Ca}^{2+}] = 0.0001 \text{ mM}. \) The optimized standard free energy (\( \Delta_r G^0 \)) of GLK is 16.19 kJ/mol, as shown in Table 4. Equilibrium constant for the chemical reference reaction can be computed as \( K = \exp(-\Delta_r G^0 / RT) = 1.46 \times 10^{-3} \). The chemical reference reaction of GLK is

\[
\text{GLC}^0 + \text{ATP}^4^- = \text{G6P}^2^- + \text{ADP}^3^- + \text{H}^+.
\]

\[\nu_i = 1 \text{ in this reference reaction, and } \sum v_i z_i^2 = 1 \times (-2)^2 + 1 \times (-3)^2 + 1 \times 1^2 - 1 \times 0^2 - 1 \times (-4)^2 = -2. \] Thus, the \( pH \) and ionic strength contributions can be calculated as follows:

\[
\Delta_{\text{ph}} = -2.303 RT \times \nu_i \times \text{pH} = -2.303 \times \left( \frac{8.3145 \times 10^{-3}}{310.15 \times 1 \times 7} \right) \times 41.56 \text{ (kJ/mol)}
\]

\[
\alpha(T) = 1.10708 - (1.54508 \times 10^{-3}) T + (5.95584 \times 10^{-6}) T^2.
\]

\[
\Delta_I = RT \alpha(T) \left( \frac{1}{1 + B_0} - \frac{1}{1 + B I / T} \right) \sum v_i z_i^2
\]

\[
= \left( \frac{8.3145 \times 10^{-3}}{310.15 \times 1 \times 7} \right) \times (2.0181 / 1.6 \times 0.1817^2) \times (-2) = 1.57 \text{ (kJ/mol)}.
\]

As shown in Table 1 in the article, \( \Delta_r H^p(l = 0) = -23.8 \text{ kJ/mol}. \) Then temperature contribution can be obtained

\[
\Delta_T = RT \ln K(T_0, I_0) \left( \frac{T_0}{T} - 1 \right) + \Delta_H^p(l_0) \left( \frac{1}{T} - \frac{1}{T_0} \right)
\]

\[
= \left( \frac{8.3145 \times 10^{-3}}{310.15 \times 1 \times 7} \right) \times (23.8) \times \left( \frac{1}{1.298.15} \right)
\]

\[= 1.61 \text{ (kJ/mol)}. \]

In order to calculate binding polynomial contribution \( \Delta_p \), the binding polynomial \( P_j \) for each reactant \( j \) should be first calculated (7):

\[
P_j = 1 + \frac{10^{-pK_{dj}}}{K_{d1}} + \frac{10^{-2pK_{dj}}}{K_{d1}K_{d2}} + \frac{[\text{Mg}^{2+}]}{K_{d1}K_{Mg1}} + \frac{10^{-pK_{Mgj}}}{K_{Mg1}K_{Mg2}} + \ldots
\]

where \( K_{di} \) is the dissociation constant of cation ion \( d \). The negative base-10 logarithm of dissociation constant \( pK_d \) \( (pK_{dij} = -\log K_{dij}) \) for all reactants are listed in Table 2 in the article. Temperature and ionic strength should be done to transform all the dissociation constants to \( T = 310.15 \text{ K} \) and \( I = 0.18 \text{ M} \) before calculating the binding polynomial \( P_j \) For reaction GLK,

\[
\Delta_P = -RT \ln \left( \frac{P_{\text{GLC}^0 \times P_{\text{ATP}^4^-}}}{P_{\text{G6P}^2^- \times P_{\text{ADP}^3^- \times \text{H}^+}}} \right) = \left( \frac{8.3145 \times 10^{-3}}{310.15} \right) \times \ln \left( \frac{1.0978 \times 3.4834}{1 \times 12.1245} \right) = 2.97 \text{ (kJ/mol)}.
\]
Therefore, the physiological free energy for reaction glucokinase GLK is

\[ \Delta_r G^0 = \Delta_r G^0 + \Delta_T + \Delta_I + \Delta_{\text{pH}} + \Delta_P \]
\[ = 16.19 + 1.61 + 1.57 - 41.56 + 2.97 \]
\[ = -19.22 \text{ (kJ/mol).} \]

The total difference is \( \Delta_{\text{tot}} = \Delta_T + \Delta_I + \Delta_{\text{pH}} + \Delta_P = 1.61 + 1.57 - 41.56 + 2.97 = -35.41 \text{ (kJ/mol).} \) The percentages of the contributions of non-standard elements to this difference are

\[ T\% = \frac{\Delta_T}{\Delta_{\text{tot}}} \times 100\% = -5\%, \]
\[ I\% = \frac{\Delta_I}{\Delta_{\text{tot}}} \times 100\% = -4\%, \]
\[ \text{pH}\% = \frac{\Delta_{\text{pH}}}{\Delta_{\text{tot}}} \times 100\% = 117\%, \]
\[ P\% = \frac{\Delta_P}{\Delta_{\text{tot}}} \times 100\% = -8\%. \]

**Nomenclature**

- \( B \): an empirical constant taken to be 1.6 M\(^{-1/2}\)
- \( E \): the minimum squared difference between model simulations and experimental data
- \( \Delta_r G^0 \): standard Gibbs free energy of reaction
- \( \Delta_r G^0_i \): apparent Gibbs free energy of formation of species \( i \)
- \( \Delta_r G^0_0 \): apparent Gibbs free energy of reaction
- \( \Delta_{\text{d}H_{K_d}} \): enthalpy of proton/metal cation dissociation
- \( \Delta_r H^0 \): standard enthalpy of reaction
- \( I \): ionic strength
- \( K \): chemical (reference reaction) equilibrium constant
- \( K' \): apparent equilibrium constant
- \( K_{a_d} \): the dissociation constant of cation ion \( d \)
- \( M \): the number of reactions
- \( N \): the number of \( pK \)s which have several values available in NIST database
- \( N_j \): the number of experimental data in each reaction
- \( P_j \): binding polynomial associated with reactant \( j \)
- \( pK \): negative logarithm of the dissociation constant
- \( R \): gas constant, 8.3145 J K\(^{-1}\) mol\(^{-1}\)
- \( S \): sensitivity
- \( T \): temperature
- \( U \): uncertainty
- \( v_i \): stoichiometric coefficient for species \( i \) in a given reaction
- \( z_i \): valence of species \( i \)
- \( \Delta_I \): the ionic strength contribution
- \( \Delta_{\text{pH}} \): the \( \text{pH} \) contribution
- \( \Delta_P \): the binding polynomial contribution
- \( \Delta_T \): the temperature contribution