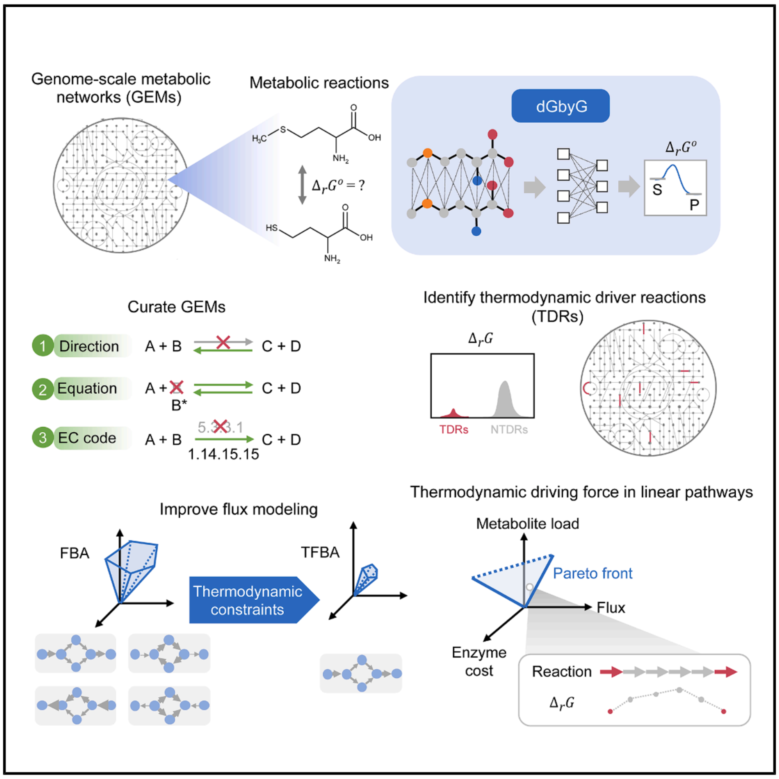
Cell Systems

Article

Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks

Graphical abstract



Highlights

- dGbyG accurately predicts the thermodynamics of metabolic reactions at the genome scale
- Predictions of dGbyG improve the accuracy of genome-scale metabolic modeling
- Thermodynamic driving forces are linked to reaction topology and enzyme expression
- Pareto optimality of enzymes, metabolites, and fluxes shapes reaction thermodynamics

Authors

Wenchao Fan, Yonghong Hao, Xiangyu Hou, Chuyun Ding, Dan Huang, Weiyan Zheng, Ziwei Dai

Correspondence

daizw@sustech.edu.cn

In brief

Fan et al. develop a graph neural network-based algorithm, dGbyG, for predicting standard Gibbs free energy change of metabolic reactions at the genome scale. The accurate reaction thermodynamics data enhance the quality of genome-scale metabolic models and reveal optimality principles governing the distribution of thermodynamic driving forces within metabolic networks.

Fan et al., 2025, Cell Systems *16*, 101393
October 15, 2025 © 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.
https://doi.org/10.1016/j.cels.2025.101393

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

Cell Systems

CellPress

Article

Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks

Wenchao Fan,[1,2] **Yonghong Hao,**[1] **Xiangyu Hou,**[1] **Chuyun Ding,**[1] **Dan Huang,**[1] **Weiyan Zheng,**[1,3] **and Ziwei Dai**[1,4,*]
[1]Department of Systems Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518055, China
[2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[3]Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[4]Lead contact
*Correspondence: daizw@sustech.edu.cn
https://doi.org/10.1016/j.cels.2025.101393

SUMMARY

Our understanding of metabolic thermodynamics is limited by the lack of genome-scale data on the standard Gibbs free energy change ($\Delta_r G^\circ$) of metabolic reactions. Here, we present dGbyG, a graph neural network (GNN)-based model for predicting $\Delta_r G^\circ$ with superior accuracy, versatility, robustness, and generalization ability. Integration of dGbyG predictions into metabolic networks facilitated model curation, improved flux prediction accuracy, and identified thermodynamic driver reactions (TDRs) with substantial negative values of the reaction Gibbs free energy change ($\Delta_r G$). TDRs showed distinctive network topological features and heterogeneous enzyme expression, implying coupling between reaction thermodynamics and network topology for efficient metabolic regulation. We also discovered a universal pattern of thermodynamics in linear metabolic pathways, explained by a multi-objective optimization model balancing the needs to maximize pathway flux and minimize enzyme and metabolite loads. Our work expands accessible thermodynamic data and elucidates optimality principles in metabolism at the genome scale. A record of this paper's transparent peer review process is included in the supplemental information.

INTRODUCTION

Thermodynamics is the primal determinant of the feasibility of all processes in nature, including a variety of biological processes on various temporal and spatial scales, from evolution spanning geological epochs[1] to biochemical reactions[2] and protein folding and allosteric regulation processes[3,4] at the time scale of milliseconds. Recent studies have demonstrated that all aspects of cellular metabolism, including but not limited to the control of metabolic fluxes,[5] the efficiency of metabolic enzymes[6] and metabolic pathways,[7] and the allocation of cellular resources into different metabolic activities,[8] are profoundly affected by thermodynamic features of metabolites, reactions, and metabolic pathways. Moreover, quantitative knowledge about thermodynamic parameters of metabolic reactions is indispensable for accurate reconstruction[9] and curation[10,11] of genome-scale metabolic models (GEMs), which are instrumental for predicting metabolic phenotypes,[12–14] designing synthetic pathways,[15] and developing theory about the operation of metabolism.[16] Therefore, quantitative thermodynamic data for metabolic reactions at the genome scale are indispensable in pursuing a systematic and quantitative understanding of metabolism.[17]

To achieve a quantitative understanding of the thermodynamics of cellular metabolism, knowledge about both standard Gibbs free energy change ($\Delta_r G^\circ$) and reaction Gibbs free energy change ($\Delta_r G$) under physiological conditions is necessary. These two quantities are linked by the relationship $\Delta_r G = \Delta_r G^\circ + RTlnQ$, in which R is the universal gas constant, T is the temperature, and Q is the reaction quotient determined by concentrations of the substrates and products. While $\Delta_r G^\circ$ reflects the difference in thermodynamic stability of the reactants and products, the reaction Gibbs free energy change $\Delta_r G$ indicates the thermodynamic feasibility of the reaction when the actual concentrations of metabolites are given.

In cellular metabolic networks, $\Delta_r G$ have multi-faceted roles in defining how they operate beyond determining the feasibility of reaction steps. For instance, the mode of flux control (i.e., how the pathway flux responds to perturbation of single enzymes) for linear reaction chains[7] and branched metabolic pathways[5] is largely determined by the distribution of $\Delta_r G$ over reactions in the pathway. In canonical metabolic pathways such as glycolysis, reactions with substantial negative $\Delta_r G$ values, such as those catalyzed by hexokinase (HK), phosphofructokinase (PFK), and pyruvate kinase (PK), are considered rate-limiting steps serving as critical control points of the pathway flux, while a few other reactions can also exert significant flux control when they have substantial negative $\Delta_r G$ under certain conditions.[18–20] On the other hand, reactions with $\Delta_r G$ values close

© 2025 Elsevier Inc.
All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

to zero are sensitive to changes in metabolite concentrations, thereby enabling efficient adaptation of the pathway flux.[21] It has also been hypothesized that the total Gibbs free energy dissipated through all metabolic reactions is constrained by an upper limit at the whole-cell level, suggesting that thermodynamics is a global regulator that governs the allocation of metabolic activity into different pathways.[8]

However, while it is intriguing to explore the principles regarding thermodynamics and metabolic regulation, most work focuses on small pathways such as glycolysis instead of the genome-scale metabolic network. This is largely due to the limitation in our quantitative knowledge about both $\Delta_r G^\circ$ and $\Delta_r G$. While the most up-to-date human GEMs, Recon3D and Human1, both contain about 4,000 metabolites and over 10,000 metabolic reactions,[9,22] experimentally measured values for $\Delta_r G^\circ$ in public databases such as the Thermodynamics of Enzyme-Catalyzed Reactions Database (TECRDB)[23] and other databases for standard Gibbs energy of formation[24–26] are only available for a few hundred reactions and metabolites, thereby greatly limiting the scope of study on thermodynamics of metabolism. Estimation of the $\Delta_r G$ requires measurements of absolute concentrations of metabolites[18] or 13C-metabolic flux analysis (MFA) models to quantify the ratio between the forward and backward fluxes,[21,27] which is challenging at the genome scale. This can be partially addressed by assuming more realistic intracellular concentrations of metabolites instead of those considered in the standard condition (i.e., each metabolite has the concentration of 1 mol/L)[28] or by integrating $\Delta_r G^\circ$ of reactions with genome-scale metabolic networks to estimate feasible regions for the metabolite concentrations and corresponding $\Delta_r G$ values. Interdependence between $\Delta_r G$, $\Delta_r G^\circ$, concentrations of metabolites, and direction of a reaction[29] provides a natural point to integrate thermodynamics into GEMs. Several algorithms[30–32] have been developed to achieve this goal.[33] Previous studies[14,34,35] have shown that thermodynamics-based metabolic modeling has unparalleled advantages in metabolic network reconstruction,[10,11] prediction of phenotypes,[12] discovery of biological principles,[8] and so on. Nevertheless, the application of these approaches is greatly limited by the lack of accurate thermodynamics data at the genome scale[36] as well as the limited focus on topics beyond the curation of GEMs by eliminating thermodynamically infeasible reactions. For instance, although $\Delta_r G$ is tightly related to the regulation of metabolic fluxes, our knowledge about its distribution in a genome-scale metabolic network, which defines the mode of metabolic control at the genome scale, is still limited. To answer these questions, both models for accurate prediction of $\Delta_r G^\circ$ and algorithms for reliable estimation of $\Delta_r G$ in realistic intracellular metabolic networks are indispensable.

Several computational methods for prediction of $\Delta_r G^\circ$ of metabolic reactions have been developed based on the additivity rules for the estimation of molecular properties, i.e., approximation of molecular features by summing up different parts of a molecule.[37] Among those methods, the most noteworthy one is the group contribution (GC) method, which relies on an expert-defined set of chemical groups to decompose the structure of molecules and estimates the standard Gibbs free energy using a linear regression model. The GC method has been applied in predicting standard Gibbs free energy of molecules and reactions in the gas phase,[37–39] liquid and solid phase,[40] and biochemical reactions in aqueous solutions.[41] Several algorithms, such as the component contribution (CC) method,[42] fingerprint contribution (FC) method,[43] and automated fragmentation method,[44] have been developed to improve the accuracy of prediction[42,45–47] and bypass the requirement of expert-defined chemical groups.[43,44] Nevertheless, the scope of prediction is still limited to metabolites that only contain chemical groups included in molecules in the training set. For example, eQuilibrator 3.0, the most up-to-date database for Gibbs free energy of metabolic reactions predicted using the CC method, provides standard Gibbs free energy of about 5,000 human metabolic reactions, covering only one-third of the 13,543 reactions in the human GEM, Recon3D.[36] *Ab initio* quantum chemical computation can calculate standard Gibbs free energy change of biochemical reactions[48–50] without the need of training datasets, yet the applicability of these methods to metabolites with complex molecular structure is still limited by the extremely high computational cost.

Because of the limitation of existing methods, developing an algorithm that can efficiently and accurately predict the standard Gibbs free energy of metabolic reactions is an urgent need of the field. The most important reason behind the limitation of the existing methods, almost all of which are derivatives of the GC method, is the dependence of these methods on the ability to decompose the molecular structures in the training set and the molecules to be predicted using the same set of chemical groups. Deep learning models based on graph neural networks (GNNs) offer a promising possibility to address this problem. Instead of considering a molecule as a linear combination of predefined groups, GNN-based models directly treat the molecular structure as a graph and preserve important chemical information at the level of atoms. Hence, GNN-based methods have been continually creating new records on various tasks of predicting properties of molecules[51–55] with broad applications in chemical property prediction,[56] molecular modeling,[57] and drug discovery.[58,59] However, to the best of our knowledge, its application for Gibbs free energy prediction has never been reported.

To address these limitations of existing methods and understand the design principles governing thermodynamics of metabolism on a genome scale, we develop a GNN-based machine learning model, dGbyG, for prediction of $\Delta_r G^\circ$ of metabolic reactions from the molecular structure of metabolites. We apply two strategies, error randomization and weighing of training data, to improve the accuracy and robustness of our model and offer a quantitative estimation of the uncertainty in the predicted values. An unbiased comparison benchmarking the performance of dGbyG and previous models shows that dGbyG outperforms all existing methods in both accuracy and versatility with less training data. $\Delta_r G^\circ$values predicted by dGbyG can be used to curate two state-of-the-art human GEMs, Recon3D[9] and Human1.[22] Despite being the most up-to-date genome-scale models for human metabolism, these models still have inaccuracies in reaction directionality and chemical equations. Moreover, integration of dGbyG predictions with metabolic network models can also improve the accuracy of metabolic flux predictions, highlighting the potential for thermodynamic data to refine and enhance metabolic modeling.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

To investigate the thermodynamic properties of metabolism at the genome scale, we further develop an algorithm that takes $\Delta_r G^\circ$ predicted by dGbyG as the input to estimate $\Delta_r G$ of reactions in genome-scale metabolic networks and identify a set of reactions with substantial negative values of $\Delta_r G$, which we term thermodynamic driver reactions (TDRs). We find a significant association between network topological features and $\Delta_r G$ of metabolic reactions, suggesting that the thermodynamic parameters of reactions might be optimized to achieve certain metabolic objectives. TDRs have distinctive network topological properties and higher variability in abundance of the associated metabolic enzymes, consistent with their potential roles in controlling metabolic fluxes. The association between $\Delta_r G$ and network topology can be explained by a multi-objective optimization model on the trade-off between the maximization of pathway flux, minimization of enzyme cost, and minimization of metabolite load in metabolic pathways. These findings elucidate the design principles of metabolic networks: that reaction thermodynamics, network topology, and enzyme abundance are jointly optimized to improve the efficiency of metabolism.

RESULTS

Development of dGbyG

Experimental data of standard Gibbs free energy for metabolites and metabolic reactions used for training the model were obtained from two sources: the TECRDB[23] published by the National Institute of Standards and Technology (NIST) and a table of standard Gibbs free energy of formation of a set of molecules.[42] To generate a high-quality training dataset robust to variability across different measurements, we conducted a series of pre-processing steps (Figure 1A), in which we removed reactions with undefined or incomplete molecular structures or missing equilibrium constants, curated unbalanced reactions, and standardized the experimental conditions for the measurements (STAR Methods; Table S1). To assess the quality of the resulting training set, we computed the standard deviation (SD) of different measurements for the same reaction or the same metabolite and found that the variation across measurements was substantially reduced by the pre-processing procedures (median SD = 1.2 kJ/mol in the pre-processed data compared with 1.9 kJ/mol in the raw data, Wilcoxon's rank-sum $p = 0.0001$; Figure S1A). Finally, we computed the average values over different measurements of the same reaction or metabolite, thereby generating 671 data points consisting of 449 $\Delta_r G^\circ$ and 222 $\Delta_f G^\circ$ related to 643 unique metabolites, which were used as the training set (Table S2).

We developed a GNN-based model, dGbyG, to predict the standard Gibbs free energy of formation ($\Delta_f G^\circ$) for a metabolite from its molecular structure and estimate the uncertainty in the predictions. We converted the topological structure of molecules to a molecular graph, in which nodes represent atoms and edges represent bonds in the molecule, and parameterized the nodes and edges using chemical features of the atoms and bonds (Table S3). An embedding layer, two message passing (MPNN) layers, three feedforward neural network (FNN) layers, and a pooling layer (Figure 1B) were then included in the GNN model, accepting the molecular graph as the input and computing the $\Delta_f G^\circ$ of the molecule by summing up features of all nodes in the pooling layer. The standard Gibbs free energy change of metabolic reactions, $\Delta_r G^\circ$, was then computed from the weighted sum of the $\Delta_f G^\circ$ of the substrates and products with the stoichiometric coefficients as the weights (STAR Methods).

To estimate the uncertainty in the model prediction caused by noises in the training set and parameter uncertainty of the model, we developed a bootstrapping-based error randomization approach. Briefly, we assumed a Gaussian distribution for the experimental measurements of the standard Gibbs free energy of each reaction and metabolite, estimated the mean and SD of those distributions from the training data, and then randomly sampled 100 data points from those distributions and used them to train the model separately (Figure 1C). We also computed a variability index α for each data point in the training set and assigned weights to different terms in the loss function accordingly (Figures S1B–S1E). This procedure resulted in 100 sets of model parameters and relevant predictions, which were finally used to estimate the mean and variance of the predictions (Figure 1D; STAR Methods).

Taken together, our model integrates several approaches to provide a high-quality dataset for model training, allows simultaneous prediction of $\Delta_f G^\circ$ and $\Delta_r G^\circ$, and characterizes the uncertainty in the model predictions. To facilitate the application of dGbyG to GEMs in various scenarios, we implement dGbyG as a Python package that offers user-friendly modules for prediction with GEMs and accepts a variety of molecular identification types—including Simplified Molecular Input Line Entry System (SMILES), International Chemical Identifier (InChI), Kyoto Encyclopedia of Genes and Genomes (KEGG) entries, and MetaNetX IDs—as well as input formats such as MOL and RXN files.

dGbyG outperforms previous models in accuracy and versatility

To evaluate the performance of dGbyG, we performed 10-fold cross-validation to assess its accuracy and compared its performance with those of other methods.[42–44] We first tested several metrics for the accuracy of prediction, including Pearson's correlation coefficient (Pearson's R), mean absolute error (MAE), and root-mean-squared error (RMSE). The comparison shows that dGbyG improved the accuracy compared with the previous models, including eQuilibrator,[42] dGPredictor,[44] and FC[43] (MAE = 9.18 kJ/mol compared with 16.02 kJ/mol for the best of previous models, RMSE = 25.88 kJ/mol compared with 47.85 kJ/mol for the best of previous models, and Pearson's R = 0.997 compared with 0.994 for the best of previous models; Table S4). We then confirmed that the accuracy of dGbyG has been improved by the introduction of the variability index α in the loss function, while the error randomization approach, which gives an estimate of the uncertainty in the prediction, did not impair the accuracy (Figure 2A). We then compared the distribution of prediction error between dGbyG and other models and again confirmed that dGbyG outperformed all previous models in accuracy (median of absolute prediction error = 4.11 kJ/mol for dGbyG compared with 5.33 kJ/mol for the best of previous models, Mood's median p value = 1.16×10^{-33}; Figures 2B and 2C). We also confirmed that the MPNN architecture implemented in dGbyG performed the best among three variants of MPNN (Table S5), including the original MPNN (MPNN), directed MPNN (D-MPNN),[60] and atom-bond transformer-based MPNN (ABT-MPNN).[61]

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

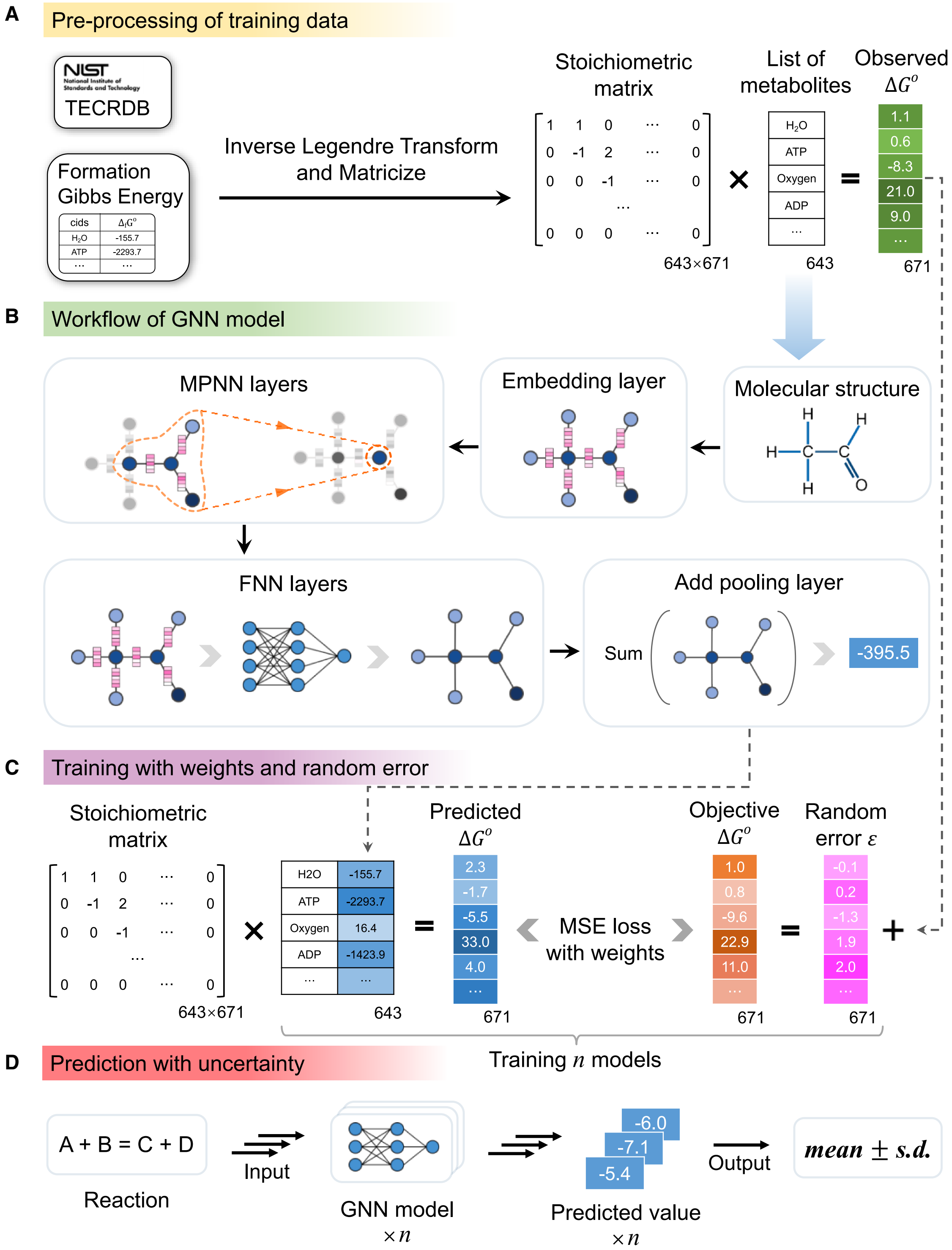


Figure 1. Development of dGbyG

(A) Workflow for pre-processing the training data, which includes 671 experimentally measured standard Gibbs free energy changes associated with 643 unique metabolites.

(B) The architecture of the GNN model used in dGbyG.

(C) The method for error randomization and data weighing used in training the model.

(D) The method for estimating the uncertainty in predictions.

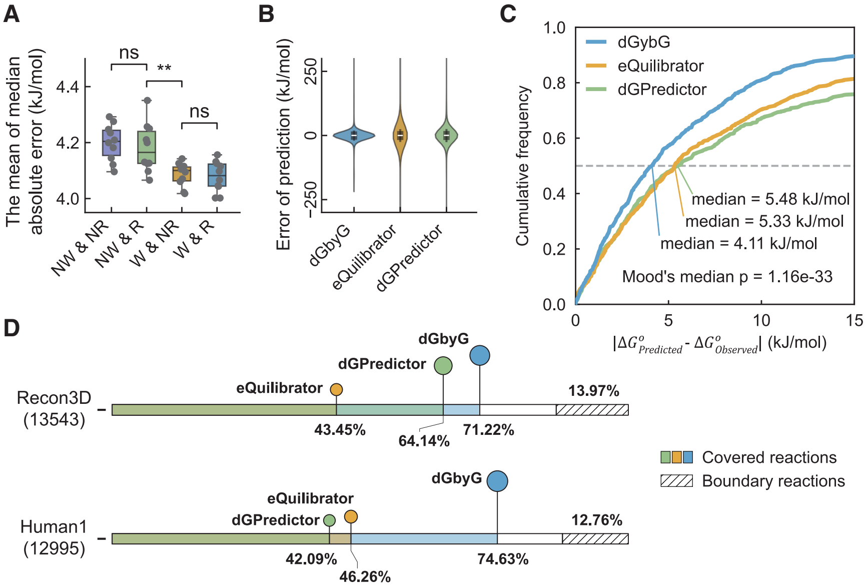


Figure 2. dGbyG outperforms previous methods in accuracy and versatility

(A) Comparison of the mean of median absolute error across different combinations of data weighting and error randomization ($n = 10$ independent replicates for each condition). W, with weighing; R, with error randomization; NW, no weighing; NR, no error randomization. ns: no significance ($p >= 0.05$), *$0.05 > p >= 0.01$, **$p < 0.01$, two-tailed Student's t test.

(B) Distributions of prediction error for dGbyG, eQuilibrator, and dGPredictor. The prediction error of dGbyG was calculated using 10-fold cross-validation, while eQuilibrator and dGPredictor were evaluated using leave-one-out cross-validation.

(C) Distributions of absolute prediction error for dGbyG, eQuilibrator, and dGPredictor. The cross-validation methods were the same as those in (B). Mood's median test was used to compute the p value.

(D) Percentage of reactions in the human GEMs Recon3D and Human1 for which the $\Delta_r G^{\circ}$ can be predicted by dGbyG, eQuilibrator, and dGPredictor.

A major limitation to the previous GC-based methods is that they cannot predict the Gibbs free energy for metabolites with chemical groups unseen in the training set. Therefore, we assessed the versatility of dGbyG and other models by testing their predictive coverage (i.e., ability to make predictions for all reactions) in two up-to-date human GEMs, Recon3D[9] and Human1.[22] We found that dGbyG can predict the standard Gibbs free energy for nearly all reactions in both GEMs, with only 14.80% of reactions in Recon3D and 12.61% in Human1 remaining uncovered (Figure 2D). This is a substantial improvement in predictive coverage compared with eQuilibrator (42.58% uncovered reactions in Recon3D and 40.98% in Human1; Figure 2D) and dGPredictor (21.89% uncovered reactions in Recon3D and 45.15% in Human1; Figure 2D). The reason that prediction by dGbyG is unavailable for a minimal set of reactions is that the molecular structure of at least one metabolite involved in those reactions is either missing or incorrect; hence, structure-based predictions of Gibbs free energy for these reactions are impossible (Figure S2). Therefore, dGbyG has reached the upper bound of reaction coverage for these GEMs. This is because the GNN-based architecture implemented in dGbyG directly models the fundamental atoms and bonds of metabolites instead of relying on predefined chemical groups or moieties, allowing it to generalize better to structurally diverse metabolites in GEMs.

These results together highlight the versatility of dGbyG, extending beyond molecular structures and chemical groups in the training data. To the best of our knowledge, dGbyG has achieved the highest accuracy among all existing machine learning models and the highest possible coverage of reactions in GEMs.

dGbyG outperforms previous models in robustness and generalization ability

We next benchmarked the robustness of dGbyG and other methods by testing the dependence of their accuracy on the size of the training set. We conducted 10-fold cross-validation, 5-fold cross-validation, and 2-fold cross-validation for dGbyG and two previous models, eQuilibrator and dGPredictor. In a k-fold cross-validation, the entire dataset is randomly split into k subsets with equal sizes, and the model is then trained k times, each using $k-1$ subsets as the training set and the remaining subset as the test set. In other words, it means that the proportion of total available data used for training the model was narrowed down to 90%, 80%, and 50%, respectively. We found that when the size of the training set shrunk, the accuracy of dGbyG was not impaired, while both eQuilibrator and dGPredictor continually lost their predictive power (Pearson's R in the 2-fold cross-validation = 0.997 for dGbyG compared with -0.075 for eQuilibrator and -0.082 for dGPredictor; Figure 3A). In fact, using only half of the available data for training, dGbyG still outperformed the accuracy that eQuilibrator and dGPredictor achieved using all available data for training the model, highlighting its robustness in the situation of few training data.

To test whether dGbyG can generate reliable predictions for reactions with mechanisms unseen in the training set, we further performed leave-one-group-out cross-validation for dGbyG and previous methods, in which the reactions were categorized into six groups according to their mechanisms defined by the enzyme commission (EC) codes of the associated enzymes (1 for oxidoreductases, 2 for transferases, 3 for hydrolases, 4 for lyases, 5 for isomerases, and 6 for ligases). In each round of cross-validation, one group of reactions was used as the validation set, while reactions in the other groups were used to train the model. This approach ensures that the accuracy of the model is evaluated on a set of reactions with mechanisms absent in the training set to test its generalization ability. The accuracy of dGbyG in such cross-validation not only substantially exceeded those of eQuilibrator and dGPredictor but also became comparable to that of *ab initio* quantum chemistry computations based on the density functional theory[48] (DFT), which directly models the electronic structure of the atoms and bonds using principles

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

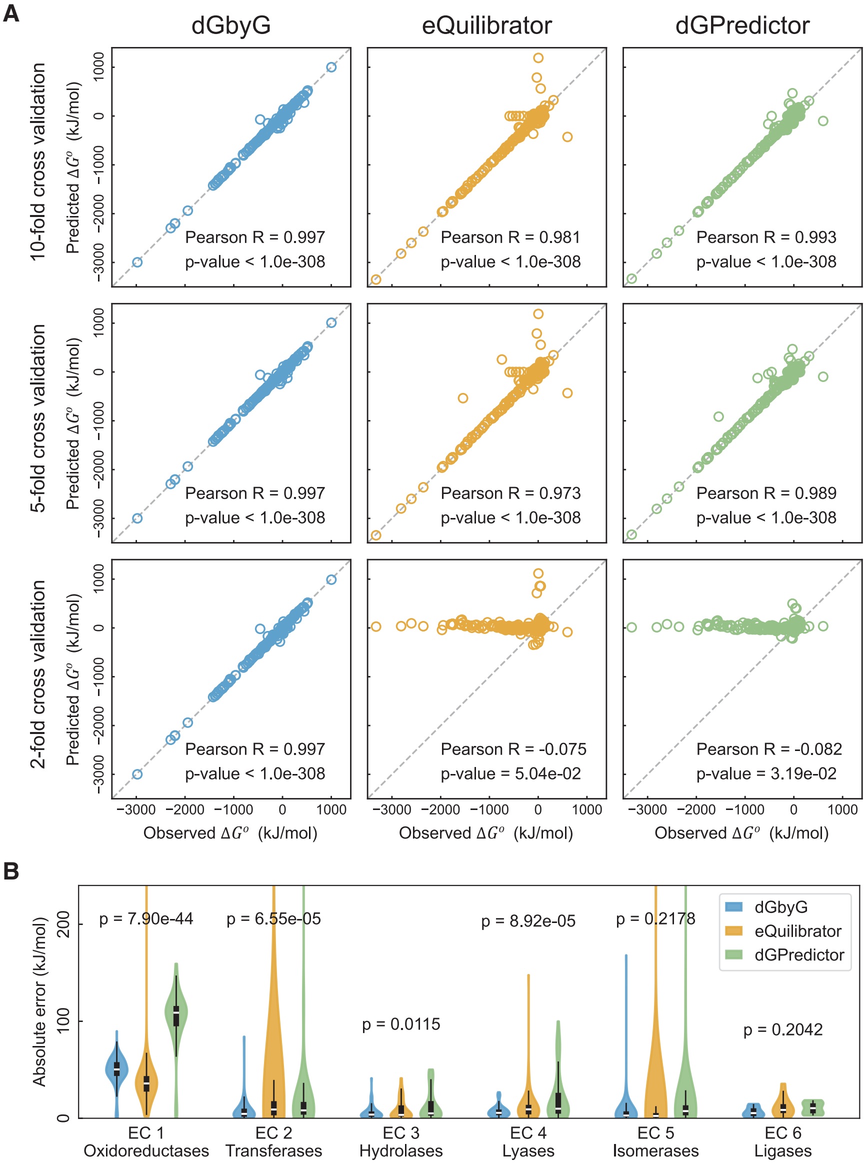


Figure 3. dGbyG outperforms previous methods in robustness and generalization ability

(A) Scatterplots comparing actual and model-predicted $\Delta_r G^\circ$ for dGbyG (left), eQuilibrator (middle), and dGPredictor (right) under 10-fold (top), 5-fold (middle), and 2-fold (bottom) cross-validation. p values were computed using two-sided t test.

(B) Distributions of prediction error for dGbyG, eQuilibrator, and dGPredictor under leave-one-group-out cross-validation as classified by enzyme commission (EC) codes. p values were computed using the Alexander-Govern test.

of quantum mechanics without the need of a training set and is believed to be the most reliable approach for molecular modeling (dGbyG resulted in lower MAE than that of DFT in three out of six groups; Figure 3A; Table S6). However, DFT-based predictions are only available for a minimal set consisting of 150 reactions, while for eQuilibrator and dGPredictor, the removal of reactions with the same mechanism from the training set greatly impaired the accuracy. These results demonstrate that dGbyG outperforms previous GC-based machine learning methods in generalization ability, which allows dGbyG to achieve accuracy comparable to that of first-principles quantum chemistry without requiring training data with similar reaction mechanisms.

dGbyG facilitates genome-scale metabolic modeling

The Gibbs free energy change determines whether a process, such as a metabolic reaction, can occur through a specific direction. For a metabolic reaction, its Gibbs free energy change $\Delta_r G$ is determined by the standard Gibbs free energy change, $\Delta_r G^\circ$, and concentrations of the substrates and products. Therefore, by integrating $\Delta_r G^\circ$ of metabolic reactions with a genome-scale metabolic network, additional thermodynamic constraints on the flux configurations can be introduced[32,62] to address the limitations of the flux balance analysis (FBA) method, which often results in non-unique optimal solutions. A common application of this approach is to assign reversibility and directionality of reactions during reconstruction of GEMs.[63] Its accuracy is limited by the scope and quality of available thermodynamic data for metabolic reactions, but such limitations can be addressed by the thermodynamic data, that is, $\Delta_r G^\circ$ of reactions, predicted by dGbyG at the genome scale.

To demonstrate the potential of dGbyG predictions in curating GEMs, we employed a thermodynamic-based FBA (TFBA) approach that incorporates a set of thermodynamic constraints that link the metabolite concentrations, $\Delta_r G^\circ$ and $\Delta_r G$ of reactions, and directionality of fluxes (STAR Methods). We applied TFBA to predict the reversibility and directionality of reactions in the human GEM Recon3D[9] using dGbyG-predicted values of $\Delta_r G^\circ$ and compared the predictions with built-in labels of reaction directionality in Recon3D. We found that for 13 reactions, their directionalities in the original Recon3D model were inconsistent with predictions by TFBA, suggesting that the directions of these reactions originally labeled in Recon3D are thermodynamically infeasible (Figure 4A; Table S7). We then performed the same analysis on Human1, identifying 12 reactions that are thermodynamically infeasible based on their original chemical equations and directionalities in the model (Figure S4A; Table S8). Only a subset of these potentially inaccurate reactions can be identified with $\Delta_r G^\circ$ values predicted by eQuilibrator (7 out of 13 for Recon3D and 9 out of 12 for Human1; Figure 4B), as $\Delta_r G^\circ$ predictions by eQuilibrator for the remaining reactions were either unavailable or impaired by extremely high uncertainty (Tables S7 and S8). This underscores the importance of

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

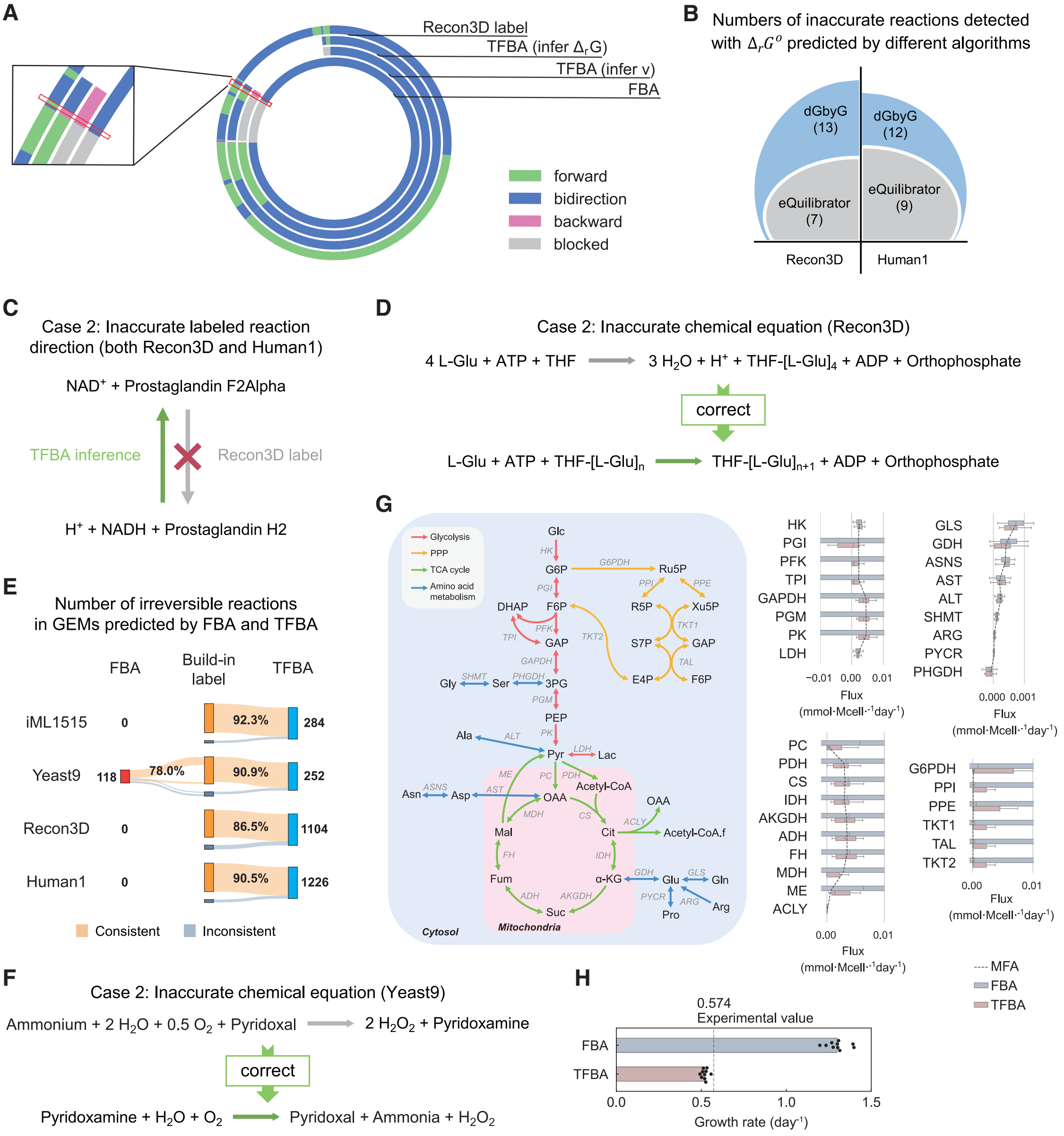


Figure 4. dGbyG facilitates genome-scale metabolic modeling

(A) Comparison between directionalities of reactions in Recon3D predicted by thermodynamics-based flux balanced analysis (TFBA), flux balance analysis (FBA), and their original labels of directionality in the model.

(B) Numbers of thermodynamically infeasible reactions identified by TFBA in Recon3D and Human1 according to $\Delta_r G^\circ$ values predicted by dGbyG and eQuilibrator, respectively.

(C) An example of a reaction with a potentially incorrectly labeled directionality.

(D) An example of a reaction with a potentially inaccurate chemical equation.

(E) Number of irreversible reactions predicted using FBA or TFBA for the GEMs iML1515 (*E. coli*), Yeast9 (*S. cerevisiae*), Recon3D, and Human1.

(F) A reaction with a potentially incorrect chemical equation identified in Yeast9 by TFBA.

(G) Left: central carbon metabolism of CHO cells; right: metabolic fluxes experimentally measured using MFA, computationally predicted using FBA or TFBA. The whiskers of the box plot represent the 5th and 95th percentiles.

(H) Comparison of experimentally measured growth rate of CHO cells to those predicted by FBA and TFBA.

obtaining accurate predictions of $\Delta_r G^\circ$ at the genome scale for the curation of GEMs.

We examined these reactions individually and found that for some of these reactions, their original directionalities in the GEMs contradicted literature-based evidence; hence, their original directionalities might be incorrectly labeled. For example, these reactions include the production of prostaglandin F2alpha ($PGF_{2\alpha}$) from prostaglandin H2 (PGH_2) in arachidonic acid metabolism, a process known to occur in many different types of human cells, such as endothelial cells and adipocytes.[64,65] While the original directionality in both Recon3D and Human1 contradicts with the evidence in literature, TFBA correctly predicted that the reaction should happen in the direction of producing $PGF_{2\alpha}$ from PGH_2 (Figure 4C). For some other reactions, their directionalities in the original GEMs result in *de novo* production of oxygen (Tables S7 and S8), which is only known to occur through photosynthesis in non-human species such as plants. These reactions can be corrected by reversing their original labels of directionality.

Some other reactions predicted to be infeasible by TFBA might have improper stoichiometric coefficients, substrates, and products in their chemical equations, which can be addressed by correcting the chemical equation. For example, predictions by TFBA suggested that a reaction in Recon3D that adds four glutamate molecules onto 5,6,7,8-tetrahydrofolate (THF) to form THF-glutamate$_4$ (THF-[Glu]$_4$) while consuming only one ATP is thermodynamically infeasible. Experimental evidence in literature suggests that the formation of poly-glutamate derivatives of THF consists of multiple reaction steps, in which a single ATP molecule is consumed in each step.[66,67] Thereby, correction of the chemical equation of that reaction can address the thermodynamic infeasibility (Figure 4D; Tables S7 and S8).

To demonstrate the broad applicability of dGbyG and TFBA to metabolic networks of other organisms, we applied our approach to the most up-to-date GEMs, iML1515[68] for *E. coli* and Yeast9[69] for *S. cerevisiae*. The irreversible reactions predicted by TFBA for these GEMs were highly consistent with the built-in directionality annotations (Figure 4E). In addition, we identified one reaction in Yeast9 predicted to be thermodynamically infeasible by TFBA. This reaction had multiple issues: an unbalanced charge, a non-integer stoichiometric coefficient for oxygen, and an implausible reaction mechanism. However, it can be corrected based on the KEGG database entry for the oxygen-dependent conversion of pyridoxamine to pyridoxal (Figure 4F; Table S9). These results suggest that genome-scale thermodynamic data predicted by dGbyG combined with TFBA provides a highly effective protocol for identifying potential thermodynamically infeasible reactions in a variety of GEMs, therefore facilitating the manual curation of these models.

Finally, to evaluate how the thermodynamic data provided by dGbyG affects the accuracy of metabolic flux prediction, we compared metabolic fluxes and growth phenotypes predicted by dGbyG and TFBA to experimentally measured values. While 13C-MFA[27] remains the gold standard for experimental flux determination, its reliance on isotope tracing and computational modeling[70] limits scalability to GEMs due to high costs and sparse genome-scale isotopic data. To address this difficulty, we focused our analysis on central carbon metabolism in Chinese hamster ovary (CHO) cells[71] using a reduced metabolic network, enabling direct comparison of FBA-, TFBA-, and 13C-MFA-derived intracellular fluxes. The results showed that while FBA-predicted fluxes exhibited substantial uncertainty, TFBA can effectively constrain the range of fluxes and yield predictions closely aligned with experimental data (Figure 4G). We further evaluated growth rate predictions using the GEM iCHO1766[72] for CHO cells, finding that the TFBA prediction aligned substantially better with experimental measurement than that from FBA (Figure 4H). These results together underscore how integrating accurate thermodynamic constraints improves the accuracy of metabolic flux and growth phenotype predictions.

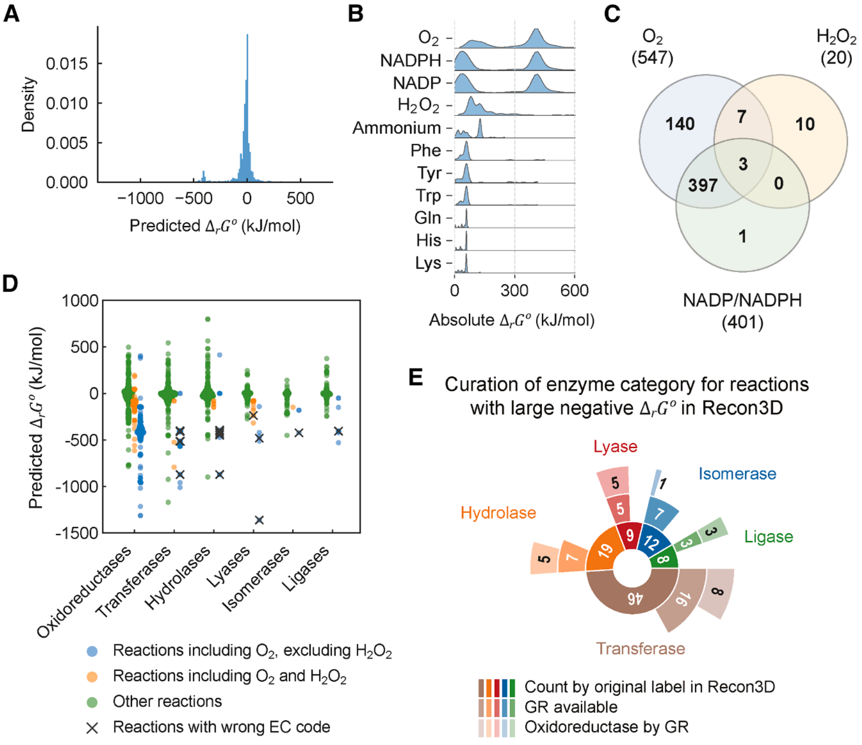
Oxygen is a key determinant of substantial negative $\Delta_r G^\circ$ of metabolic reactions

The accuracy, robustness, and versatility of dGbyG make it a promising tool to characterize the thermodynamic properties of metabolic networks at the genome scale. Hence, we analyzed the $\Delta_r G^\circ$ values of metabolic reactions in the human GEM Recon3D predicted by dGbyG. We first confirmed that the predictive error of dGbyG (defined as the difference between $\Delta_r G^\circ$ predicted by dGbyG and experimentally measured values) was independent of the $\Delta_r G^\circ$ of reactions, and that large predictive errors were not associated with any specific metabolites (Figure S3). We then studied the distribution of $\Delta_r G^\circ$ of reactions in the model, which clearly showed a bimodal distribution with one large peak near 0 kJ/mol and a minor one located near -400 kJ/mol (Figure 5A), suggesting that these reactions are far from thermodynamic equilibrium under standard conditions. To characterize the features of the subset of reactions with large negative $\Delta_r G^\circ$ values, we first investigated the substrates and products participating in these reactions. It has been reported in previous studies with a small set of NAD(P)-related oxidation-reduction reactions that the oxygen-dependent ones tend to have larger negative $\Delta_r G^\circ$ compared with the oxygen-independent ones, but whether this rule applies to reactions with all possible mechanisms remains unknown.[73]

To identify key metabolites participating in these reactions, we selected metabolites that appear in more than 100 reactions and ranked them by the median absolute $\Delta_r G^\circ$ of the reactions associated with that metabolite (Figure 5B), showing that oxygen, NADPH, NADP, and hydrogen peroxide are key participants of reactions with large negative $\Delta_r G^\circ$ values. Next, for the subset of reactions with $\Delta_r G^\circ < -300$ kJ/mol, we counted the numbers of reactions in that subset involving oxygen, NADP, NADPH, and hydrogen peroxide, respectively. We found that although NADP/NADPH are frequently included in these reactions (401 out of 619 reactions), almost all NADP/NADPH-coupled reactions also require oxygen (400 out of 401; Figure 5C). Interestingly, although hydrogen peroxide is also involved in a small set of reactions with substantial negative $\Delta_r G^\circ$ values (20 reactions; Figure 5C), reactions involving both oxygen and hydrogen peroxide have $\Delta_r G^\circ$ values closer to zero compared with those reactions involving oxygen but not hydrogen peroxide (Figure 5D). Similar trends were also observed for $\Delta_r G^\circ$ of reactions in Human1 (Figures S4B and S4C).

Furthermore, comparison of $\Delta_r G^\circ$ across reactions catalyzed by different classes of enzymes showed that the minor peak in the distribution near -400 kJ/mol was predominantly associated with oxidoreductases (Figure 5D). While a small number of reactions annotated to other types of enzymes also exhibited large negative $\Delta_r G^\circ$, many of these reactions were associated with genes encoding oxidoreductases according to their gene-reaction (GR) rules in the model (Figures 5E and 5F; Table S10). This inconsistency suggests potential misannotation of EC numbers for these reactions (Figure 5D). Nevertheless, these results again highlight that the accurate thermodynamic data

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393



F

Reaction ID	Chemical Equation	EC numbers in Recon3D	Gene rule	EC numbers associated with genes in KEGG
COQ7m	O_2 + H^+ + NADPH + 2-Decaprenyl-6-Methoxy-3-Methyl-1,4-Benzoquinone = H_2O + $NADP^+$ + 2-Decaprenyl-5-Hydroxy-6-Methoxy-3-Methyl-1,4-Benzoquinone	3.1.1.1 (Hydrolase)	10229.1	1.14.13.253 (Oxidoreductase)
RE1826M	O_2 + H^+ + NADPH + 5Beta-Cholestane-3Alpha,7Alpha-Diol = H_2O + $NADP^+$ + 5Beta-Cholestane-3Alpha,7Alpha,27-Triol	5.3.3.1 (Isomerase)	1593.1	1.14.15.15 (Oxidoreductase)
FAH3	O_2 + NADPH + Hexadecanoate (N-C16:0) = H_2O + $NADP^+$ + Omega Hydroxy Hexadecanoate (N-C16:0)	6.2.1.2 (Ligase)	1579.1	1.14.14.1 (Oxidoreductase)
C4STMO1r	3.0 O_2 + 3.0 H^+ + 3.0 NADPH + 14-Demethyllanosterol = 4Alpha-Methylzymosterol-4-Carboxylate + 3.0 $NADP^+$ + 4.0 H_2O	4.2.1.17 (Lyase)	6307.1	1.14.18.9 (Oxidoreductase)
RE1096R	O_2 + NADPH + 17Alpha-Hydroxyprogesterone = H_2O + $NADP^+$ + Androst-4-Ene-3,17-Dione + Acetate	2.4.1.17 (Transferase)	1586.1	1.14.14.19 1.14.14.32 (Oxidoreductase)

Figure 5. Oxygen is a key determinant of substantial negative $\Delta_r G^\circ$ of metabolic reactions

(A) Distribution of $\Delta_r G^\circ$ for reactions in the human GEM, Recon3D.

(B) Distribution of absolute value of $\Delta_r G^\circ$ in reactions involving different metabolites.

(C) Number of reactions involving oxygen, hydrogen peroxide, and NADP/NADPH in the set of reactions with $\Delta_r G^\circ < -300$ kJ/mol.

(D) Distributions of $\Delta_r G^\circ$ in reactions catalyzed by different categories of enzymes and involving different combinations of oxygen and hydrogen peroxide.

(E) Numbers of reactions with large negative $\Delta_r G^\circ$ values that are likely oxidoreductase-catalyzed reactions but potentially misannotated with other types of enzymes in the Recon3D model.

(F) Examples of reactions with potentially misannotated EC numbers in the Recon3D model.

generated by dGbyG is crucial for identifying potential inaccuracies in GEMs.

Taken together, these results indicate that oxygen is the key determinant of thermodynamic driving force of metabolic reactions, rendering oxygen-dependent reactions thermodynamically irreversible and potentially rate-limiting steps controlling metabolic fluxes through a pathway.[5] However, since the thermodynamic parameter directly affecting directionality of reactions and mode of flux control is the reaction Gibbs free energy $\Delta_r G$ that relies on both $\Delta_r G^\circ$ and the physiological concentrations of metabolites, whether these reactions truly serve as key control points of metabolic networks needs further investigation.

Network topological and proteomic features of TDRs

Our previous work[5] and other related studies[7] have theoretically demonstrated that reactions with substantial negative $\Delta_r G$ tend to become the rate-limiting step of a linear metabolic pathway, thereby playing an important role in regulating local metabolic fluxes. Such a relationship was confirmed by numerical simulation of thermodynamics and flux control in a kinetic model of glycolysis[5] (Figure S5; STAR Methods). Studies based on control theory have also reported that, to effectively control a dynamic network to evolve to a desired state, the selection of driver nodes, i.e., a minimal set of nodes in the network that is perturbed to control the dynamics of the entire system, largely depends on the topology of the network.[74,75] It is thus tempting to speculate that, for efficient control of metabolic fluxes, reactions with large negative values$\Delta_r G$ should have specific topological features distinct from that of non-thermodynamic-driver reactions in the genome-scale metabolic network. To test this hypothesis, we developed an algorithm to sample thermodynamically and biologically feasible metabolic fluxes and the corresponding $\Delta_r G$ values of metabolic reactions based on GEMs and $\Delta_r G^\circ$ values predicted by dGbyG (Figure 6A; STAR Methods). Briefly, the flux balance, thermodynamic, and osmotic constraints used in the TFBA framework were combined with random objective functions to sample generation flux configurations, optimizing these objective functions. We then define the set of reactions with $\Delta_r G < -200$ kJ/mol and at least one non-zero flux in all flux configurations as TDRs. Comparison between the sets of TDRs and non-TDRs (NTDRs) identified based on $\Delta_r G^\circ$ predicted by dGbyG and those predicted by eQuilibrator shows that dGbyG identified more TDRs than eQuilibrator (Figure 6B).

We next sought to investigate whether TDRs have specific network topological features compared with other reactions. Because a metabolic network is a bipartite graph including two types of nodes corresponding to reactions and metabolites,

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

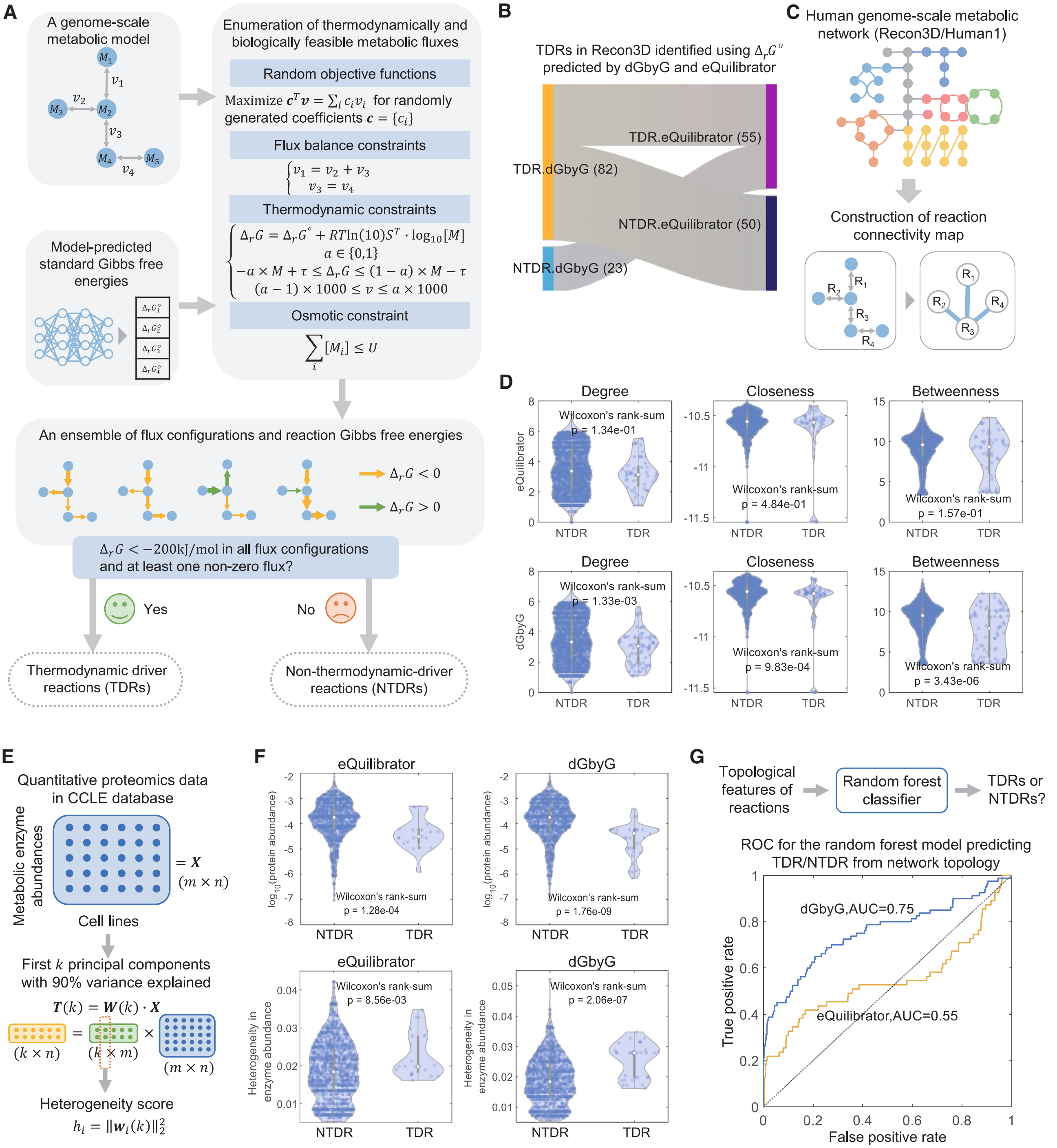


Figure 6. TDRs have distinguished network topological and proteomic features

(A) Workflow for identification of thermodynamic driver reactions (TDRs).

(B) Sankey plots comparing TDRs in Recon3D identified by $\Delta_r G^\circ$ predicted by dGbyG and eQuilibrator. Reactions identified as NTDRs by both dGbyG and eQuilibrator are not shown.

(C) Construction of the reaction connectivity map from a metabolic network.

(D) Violin plots comparing the distributions of log-transformed node degree, closeness, and betweenness between TDRs and NTDRs in Recon3D. *p* values were computed using the two-sided Wilcoxon's rank-sum test.

(E) Structure of the quantitative proteomics data for human cancer cell lines in the CCLE database and the method for calculating the PCA-based heterogeneity score.

(legend continued on next page)

we converted it into a unipartite graph that only includes reactions as the nodes, which we named the reaction connectivity map (Figure 6C; STAR Methods). An edge connecting two reactions in the reaction connectivity map means that velocities of the two reactions are coupled by at least one metabolite that is involved in both reactions. We first compared several metrics for node importance between the TDRs and other reactions, including node degree, closeness, and betweenness. Consistent with a previous study,[76] different categories of enzymes have distinct distributions of these topological features (Figure S6). We found that TDRs tend to have significantly lower degree (Wilcoxon's rank-sum $p = 1.33 \times 10^{-3}$; Figure 6D) as well as the other two node importance metrics (Wilcoxon's rank-sum $p = 9.83 \times 10^{-4}$ for closeness and 3.43×10^{-6} for betweenness; Figure 6D). It is worth mentioning that driver nodes identified solely based on network topology using control theory also tend to avoid the high-degree hubs in the network,[74] suggesting that TDRs identified based on reaction thermodynamics share similar topological features of the driver nodes predicted using control theory. Therefore, we hypothesize that TDRs might be key control points that determine the state of metabolic networks. Consequently, metabolic enzymes catalyzing TDRs should have higher expressional heterogeneity compared with those catalyzing NTDRs. To confirm that abundances of the enzymes catalyzing TDRs can determine the metabolic phenotype of a cell, we obtained quantitative proteomics datasets for 378 human cancer cell lines from the Cancer Cell Line Encyclopedia (CCLE) database[77] and computed a principal-component analysis (PCA)-based heterogeneity score to quantify the contribution of each enzyme to the metabolic heterogeneity of cancer cell lines[78] (Figure 6E). We found that although enzymes catalyzing TDRs have lower abundance compared with other reactions (Wilcoxon's rank-sum $p = 1.76 \times 10^{-9}$; Figure 6F), these enzymes have a higher heterogeneity score (Wilcoxon's rank-sum $p = 2.06 \times 10^{-7}$; Figure 6F), suggesting that the variation in expression of enzymes catalyzing TDRs indeed dominates the metabolic heterogeneity. These results emphasize that the positioning of TDRs within the genome-scale metabolic network is not random, potentially enhancing their ability to regulate the metabolic activity of a cell. To further validate this, we built a random forest regression model that can successfully predict whether a metabolic reaction is a TDR from four variables describing its topological features, including node degree, betweenness, closeness, and distance to nutrient uptake reactions (5-fold cross-validation area under curve [AUC] = 0.75; Figure 6G). These findings were robust to the selection of $\Delta_r G$ threshold for TDRs, as using a different threshold, −250 kJ/mol, did not change the conclusions (Figure S7).

Taken together, by integrating reaction thermodynamics, network topology, and proteomics data, we have shown that TDRs may have important roles in determining the metabolic phenotypes of cells. We also repeated the analysis of TDRs with another human GEM, Human1, and confirmed that TDRs identified for Human1 have the same topological and proteomic features as those in Recon3D, suggesting that the association between thermodynamics, network topology, and enzyme abundance is robust to the selection of GEM (Figure S8). TDRs identified based on $\Delta_r G^\circ$ values predicted by eQuilibrator do not show such distinctive network topological (Figures 6D and 6G) and proteomic (Figure 6F) features compared with NTDRs, highlighting the importance of accurate predictions of $\Delta_r G^\circ$ in studying thermodynamic properties of genome-scale metabolic networks.

Multi-objective optimization of flux, enzyme, and metabolites shapes the distribution of $\Delta_r G^\circ$ in linear metabolic pathways

In theoretical analysis of metabolism, the simplest and most thoroughly explored model is a linear chain of reactions.[79] With the genome-scale thermodynamic data generated by dGbyG, it is therefore intriguing to study the pattern of thermodynamic properties of reactions in linear metabolic pathways. We first used glycolysis as an example of linear metabolic pathway to illustrate the distribution of $\Delta_r G^\circ$ over different reaction steps and confirm that $\Delta_r G^\circ$ values predicted by dGbyG are very close to their actual values (Figure 7A). Next, for each reaction in the human GEMs Recon3D and Human1, we computed its distance to nutrient uptake reactions in the model as a surrogate for the location of that reaction in a linear metabolic pathway. A reaction with a shorter distance to nutrient uptake reactions is more likely located upstream of a metabolic pathway. We then compared the distribution of $\Delta_r G^\circ$ and $\Delta_r G$ across reactions with varying distances to the nutrient uptake reactions and found that both $\Delta_r G^\circ$ and $\Delta_r G$ exhibited an inverted U-shaped trend against the distance to nutrient uptake reactions in both Recon3D (Figures 7B–7D) and Human1 (Figure S9). This relationship suggests that in the human GEMs, reactions at the beginning and end of metabolic pathways are more likely to be driven by stronger thermodynamic forces and may exert stronger control over the pathway flux. On the other hand, reactions located in the middle of the pathway are closer to thermodynamic equilibrium. This is consistent with the thermodynamic property and regulation mode of glycolysis (Figure 7A), in which the upstream reactions HK and PFK and the downstream reaction PK are known to have large negative $\Delta_r G^\circ$ and $\Delta_r G$ values, often considered as rate-limiting steps, and under tight control by allosteric regulations.[80] Such an inverted U-shaped trend can be seen by visualization of $\Delta_r G^\circ$ of reactions on several representative metabolic networks that involve local linear pathways, such as nucleotide and histidine biosynthesis in *E. coli*, branched-chain amino acid (BCAA) metabolism, and central carbon metabolism in humans (Figure 7E).

However, it is still unclear why the distribution of $\Delta_r G^\circ$ of reactions in the metabolic pathway follows such an inverted U-shaped pattern, in which reactions with stronger thermodynamic driving forces tend to locate in the beginning and the

(F) Violin plots comparing log-transformed enzyme abundances and PCA-based heterogeneity scores between TDRs and NTDRs in Recon3D. p values were computed using the two-sided Wilcoxon's rank-sum test.

(G) Receiver operating characteristics (ROC) curves showing the performance of the random forest model predicting TDRs in Recon3D from the network topological features of reactions. TDRs are either identified based on dGbyG (blue curve, AUC = 0.75) or eQuilibrator (yellow curve, AUC = 0.55).

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

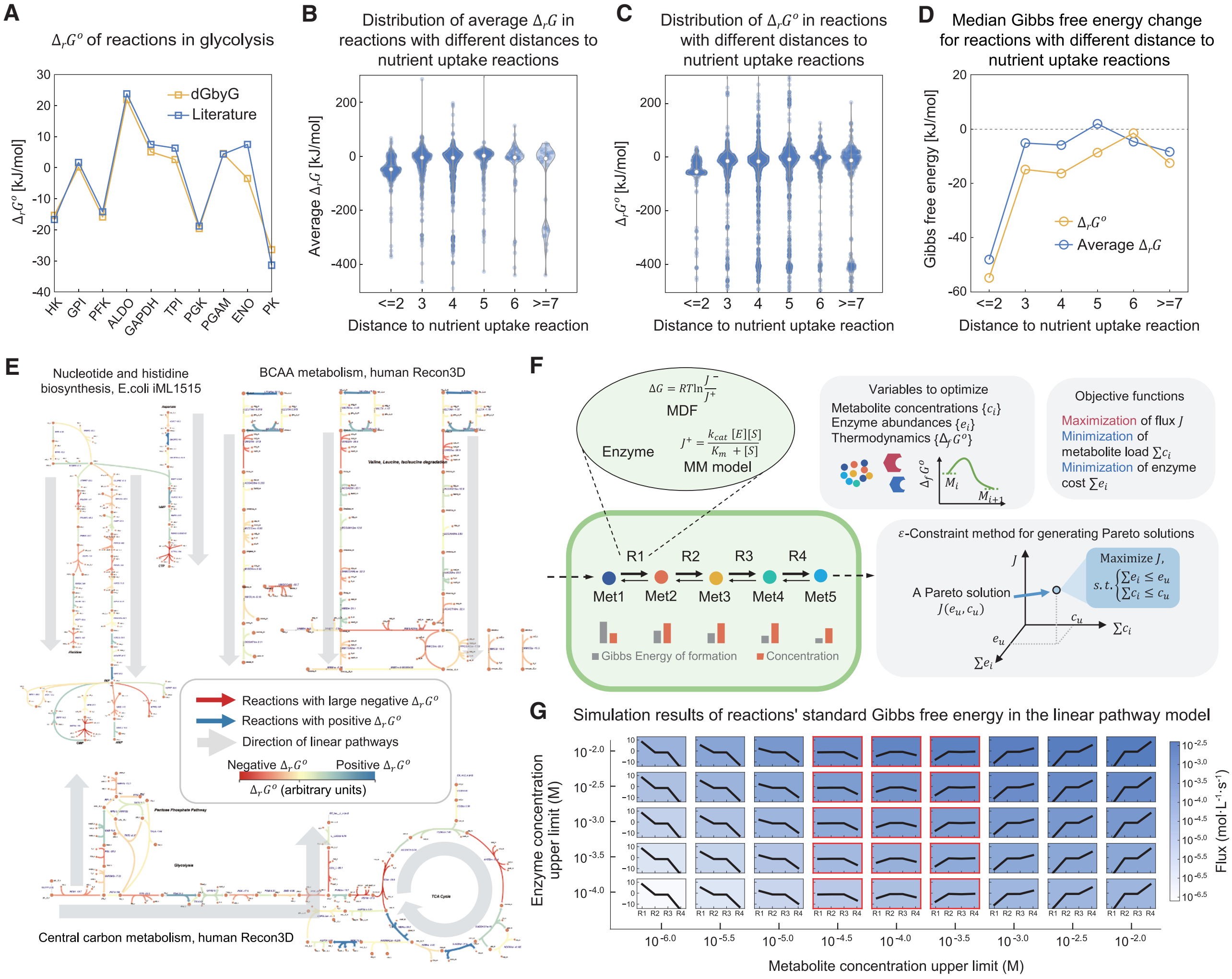


Figure 7. Multi-objective optimization of flux, enzyme, and metabolites shapes the distribution of $\Delta_r G^\circ$ in linear metabolic pathways

(A) Comparison of $\Delta_r G^\circ$ values for reactions in the glycolysis pathway from the literature and those predicted by dGbyG.

(B) Distributions of average $\Delta_r G$ in reactions at varying distances from nutrient uptake reactions in Recon3D.

(C) Distributions of $\Delta_r G^\circ$ in reactions at varying distances from nutrient uptake reactions in Recon3D.

(D) Median values of $\Delta_r G^\circ$ and median values of average $\Delta_r G$ for reactions at varying distances from nutrient uptake reactions in Recon3D.

(E) Visualization of $\Delta_r G^\circ$ on representative metabolic pathways.

(F) Schematic of the linear metabolic pathway model consisting of 5 metabolites and 4 reactions (left), and the algorithm for generating Pareto solutions using the ε-Constraint method (right).

(G) Pareto optimal solutions for $\Delta_r G^\circ$ values of reactions in the linear metabolic pathway model in (F) under different combinations of upper limits on total enzyme abundance and total metabolite concentration. Red frames indicate cases with the inverted U-shaped trend of $\Delta_r G^\circ$ distribution along the pathway.

end of metabolic pathways. Previous studies have shown that the configuration of cellular metabolic networks is shaped by multiple metabolic objectives, including maximization of yield (e.g., production of biomass and ATP) and minimization of cost (e.g., load of metabolites and metabolic enzymes,[81–83] dissipation of energy,[8] and perturbation of fluxes[84]). Hence, we hypothesize that optimization of multiple metabolic objectives in a linear metabolic pathway might result in such a pattern of $\Delta_r G^\circ$ distribution. To test this hypothesis, we developed a mathematical model of a linear metabolic pathway consisting of five metabolites and four reactions (Figure 7F) using reversible Michaelis-Menten kinetics (STAR Methods). Within this model, the concentrations of inner metabolites (Met2, Met3, and Met4), $\Delta_r G^\circ$ of all reactions, and abundance of enzymes can be optimized, whereas concentrations and $\Delta_f G^\circ$ of the initial metabolite (Met1) and the final metabolite (Met5) were fixed as boundary conditions for the model. We considered three metabolic objectives that have been extensively explored in previous research: maximizing pathway flux, minimizing total enzyme abundance, and minimizing total metabolite concentrations. The first objective focuses on maximizing yield represented by the pathway flux, while the other two aim to minimize the costs associated with the load of enzyme and metabolites. These three objectives inherently involve trade-offs, as reducing the abundance of

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

metabolic enzymes or metabolites generally leads to a decrease in the pathway flux.

The trade-off between these objectives can be characterized by constructing the Pareto front, which represents metabolic states where improvement in one objective (e.g., increasing metabolic flux) requires a compromise in at least one other objective (e.g., increasing enzyme or metabolite load). To construct the Pareto front, we applied an ε-constraint approach (Figure 7F) in which maximizing pathway flux was treated as the primary objective function, while the other two objectives, minimization of metabolite concentration and enzyme abundance, were converted to inequality constraints with varying upper limits c_u (upper limit on total metabolite concentration) and e_u (upper limit on total enzyme abundance). We then solved the resulting constrained optimization problems to compute the distribution of $\Delta_r G^\circ$ over the pathway that maximizes the pathway flux J under different combinations of upper limits c_u and e_u (Figure 7F; STAR Methods).

Our results showed that the Pareto optimal solution of $\Delta_r G^\circ$ varies along with changes in the upper limit of total metabolic concentration, c_u, regardless of the amount of total available enzymes, e_u (Figures 7G and S10). An inverted U-shaped trend, which echoes the pattern of $\Delta_r G^\circ$ values in the GEMs (Figures 7B–7E and S9), emerged when the upper bound of total metabolite concentration c_u falls in the range of $10^{-4.5}$ to $10^{-3.5}$ mol/L. Such an inverted U-shaped trend in the Pareto optimal solution of $\Delta_r G^\circ$ for c_u around 10^{-4} mol/L was unaffected by altering the parameters K_m (Figures S10A and S10B), $\Delta_f G^\circ_{Met1}$, and $\Delta_f G^\circ_{Met5}$ (Figures S10C and S10D). Changes in concentrations of Met1 and Met5 had a greater effect on the Pareto optimal solution of $\Delta_r G^\circ$, but the range of c_u that led to the inverted U-shaped trend still fell in the range of roughly 10^{-5} to 10^{-3} mol/L (Figure S11). Interestingly, this is very close to the actual absolute concentration of metabolites in human cells (median concentration of metabolites around 10^{-4} mol/L[18,85]), further validating our model. These results suggest that the trade-off between maximizing pathway flux and minimizing metabolite and enzyme load robustly drives the emergence of the inverted U-shaped $\Delta_r G^\circ$ distribution in real metabolic pathways.

DISCUSSION

The goal of this study is to understand the quantitative principles governing the distribution of thermodynamic driving force in genome-scale metabolic networks. For this aim, we develop a powerful machine learning model, dGbyG, for predicting $\Delta_r G^\circ$ of metabolic reactions at the genome scale. The model has several remarkable advantages compared with previous models: first, the application of GNN in modeling the molecular structures of metabolites enables the model to efficiently extract information at the levels of atoms and bonds and make accurate predictions for metabolites with chemical groups unseen in the training data and metabolic reactions of mechanisms distinct from those in the training data. Second, with the strategies of error randomization and data weighing that address the uncertainty in the training data and predicted values, we were able to further improve the robustness of our model and give an estimation of the uncertainty in the prediction without impairing the accuracy. The comparison with previous methods demonstrates that dGbyG has achieved state-of-the-art performance in accuracy, robustness, versatility, and generalization ability. Furthermore, the combination of dGbyG and thermodynamics-based analysis of GEMs can improve the accuracy of predicted metabolic fluxes and identify potential inaccuracies in these models, hence facilitating the improvement of the quality of them. Since the reconstruction of high-quality GEMs is crucial for a quantitative understanding of metabolism, we believe that the application of our approach could greatly benefit the community.

Leveraging the versatility of dGbyG, we achieved the upper bound of reaction coverage for predicting $\Delta_r G^\circ$ of metabolic reactions in GEMs. These predictions enable us to address key questions regarding the design principles of metabolic networks. Although some results of our analysis align with previous studies, such as the association between substantial negative value of $\Delta_r G^\circ$ and oxygen-dependent reactions,[73] our analysis of TDRs revealed that these reactions have topological features that are distinct from other reactions and dominate the metabolic heterogeneity of human cancer cells. These findings suggest that the distribution of thermodynamic driving force in the human metabolic network is skewed toward these TDRs. This is potentially for efficient control of metabolic fluxes through regulating enzymes catalyzing TDRs, as a strong thermodynamic driving force—the feature of TDRs—is often linked to larger flux control coefficients (FCCs) in the context of metabolic control analysis (MCA), meaning that metabolic fluxes are sensitive to changes in abundance or activity of enzymes catalyzing these reactions.

It is important to note that the regulation of metabolic networks has previously been investigated from two distinct theoretical perspectives. The first is MCA, which focuses on calculating the FCCs that quantify how changes in individual enzymes influence the overall pathway flux.[86] The second perspective, network controllability, assesses the structural controllability of complex networks, identifying whether the network can be directed toward a desired final state by modulating a subset of reactions serving as the driver nodes.[75] While MCA and network controllability are distinct theories that have historically been developed by researchers in different fields, they are inherently connected. First, both aim to identify critical reactions for controlling a metabolic network. MCA highlights reactions with larger FCCs, while network controllability analysis focuses on driver nodes. Second, these two frameworks address complementary aspects of metabolic regulation. To efficiently drive the metabolic network to a desired state by targeting the driver nodes, it is also necessary for the flux through these driver node reactions to be sufficiently responsive to perturbations of enzymes catalyzing them. In other words, driver nodes identified by network controllability analysis should also have large FCCs. Two findings in this study support this hypothesis: TDRs—reactions that likely have large FCCs—and driver nodes share similar network topological features, and metabolic enzymes associated with TDRs dominate the proteomic heterogeneity of human cell lines. Further theoretical work is needed to integrate the two aspects of metabolic regulation and determine the optimal thermodynamic pattern for efficient control of metabolic networks.

Optimality is an important topic in discussions about design principles of biological processes, including metabolism. As reaction thermodynamic parameters such as $\Delta_r G^\circ$ and $\Delta_r G$ change with alterations in network topology and concentrations

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

of metabolites, their choice should be shaped during optimization of the topology and parameters of metabolic networks. In this study, we found an inverted U-shaped pattern of $\Delta_r G^\circ$ and $\Delta_r G$ in the human metabolic network, in which reactions with stronger thermodynamic driving forces tend to be at the beginning and end of metabolic pathways. This pattern can be explained by a multi-objective optimization model on the trade-off between maximizing pathway flux and minimizing enzyme cost and metabolite load in linear pathways. However, due to the topological complexity of the GEM, our theoretical analysis is limited to the simplest linear metabolic pathways. Further analysis with more complicated network structures is needed to extend the relationship between multi-objective optimality and thermodynamics to non-linear metabolic pathways.

Since optimality of biological systems is often investigated from an evolutionary perspective, an intriguing question is whether the thermodynamic patterns we have identified in human metabolic networks in this study are conserved across different phylogenetic lineages, especially for species that are evolutionarily distant from humans. Our analysis focused on human metabolism because high-quality human GEMs and omics datasets are readily available from public databases, while data for less-studied species are much scarcer. However, despite the evolutionary conservation of certain specific metabolic pathways, metabolic networks exhibit significant diversity across species. For instance, anaerobic organisms use alternative electron acceptors such as sulfate, nitrate, and sulfur instead of oxygen.[87] Since these molecules have lower reduction potential than oxygen, reactions using them as electron acceptors may have lower thermodynamic driving forces compared with oxygen-dependent reactions in human metabolism, resulting in a different distribution of $\Delta_r G^\circ$ and $\Delta_r G$ within their metabolic networks. Application of dGbyG to GEMs of other species, such as those available at public databases,[88] will help characterize the thermodynamic properties of non-human metabolic networks and determine whether evolutionary constraints influence the thermodynamics of metabolic networks.

In summary, our work offers a powerful computational tool to obtain high-quality thermodynamic data for metabolic reactions at the genome scale and illustrates the possibility to discover design principles of metabolic networks about the relationship between thermodynamics and network topology.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ziwei Dai (daizw@sustech.edu.cn).

Materials availability

This study did not generate new materials.

Data and code availability

- Datasets used for training dGbyG and analyzed data have been deposited at Zenodo and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table. Original code and detailed instructions for dGbyG and TFBA are also available as Gitee repositories (https://gitee.com/f-wc/dGbyG.git and https://gitee.com/f-wc/ThermoInfer.git). MATLAB scripts analyzing the topological and thermodynamic properties of reactions in Recon3D and Human1 are also available as a GitHub repository (https://github.com/ziweidai/GEM_thermodynamics).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors thank the National Key Research and Development Program of China (2021YFA0911300 and 2021YFF1201000 to Z.D.), the National Natural Science Foundation of China (12371489 to Z.D.), the Guangdong Program (2021QN02Y856 to Z.D.), and the Shenzhen Science and Technology Program (KQTD20180411143432337 to Z.D.) for their generous support. Support for computational resources from the Center for Computational Science and Engineering of Southern University of Science and Technology is gratefully acknowledged.

AUTHOR CONTRIBUTIONS

W.F. and Z.D. designed the study and wrote the manuscript. W.F. developed the dGbyG algorithm, conducted the comparison of performance with previous methods, and performed the analysis with the linear pathway model. Z. D. performed the analysis of thermodynamic driver reactions in Recon3D and Human1, with inputs from D.H. Y.H. performed the 13C-MFA of CHO cells. X.H. curated the misannotated EC codes in Recon3D. W.Z. analyzed the CCLE proteomics data. C.D. and W.F. curated the mistaken reactions in the GEMs found by TFBA. All authors have read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT and DeepSeek in order to enhance the clarity of the text. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - ○ Pre-processing of training data
  - ○ Legendre transformation
  - ○ Structure of GNN model
  - ○ Estimation of uncertainty in the prediction
  - ○ Computation of $\Delta_r G^\circ$ for multi-compartment reactions in GEMs
  - ○ Thermodynamics-based Flux Balance Analysis (TFBA)
  - ○ Predicting directionality of reactions in GEM
  - ○ 13C-Metabolic flux analysis (13C-MFA)
  - ○ Identification of TDRs
  - ○ Topological analysis of metabolic networks
  - ○ Analysis of enzyme abundance in human cancer cell lines
  - ○ Kinetic model of glycolysis
  - ○ Linear metabolic pathway model
  - ○ Constructing Pareto front for the linear metabolic pathway

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2025.101393.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

Received: September 27, 2024
Revised: April 16, 2025
Accepted: August 15, 2025

REFERENCES

1. Demetrius, L. (2000). Thermodynamics and Evolution. J. Theor. Biol. *206*, 1–16. https://doi.org/10.1006/jtbi.2000.2106.
2. Nelson, D.L., Cox, M.M., and Hoskins, A.A. (2021). Lehninger Principles of Biochemistry, Eighth Edition (Macmillan Learning).
3. Shakhnovich, E. (2006). Protein Folding Thermodynamics and Dynamics: Where Physics, Chemistry, and Biology Meet. Chem. Rev. *106*, 1559–1588. https://doi.org/10.1021/cr040425u.
4. Weng, C., Faure, A.J., Escobedo, A., and Lehner, B. (2024). The energetic and allosteric landscape for KRAS inhibition. Nature *626*, 643–652. https://doi.org/10.1038/s41586-023-06954-0.
5. Dai, Z., and Locasale, J.W. (2018). Thermodynamic constraints on the regulation of metabolic fluxes. J. Biol. Chem. *293*, 19725–19739. https://doi.org/10.1074/jbc.RA118.004372.
6. Ooka, H., Chiba, Y., and Nakamura, R. (2023). Thermodynamic principle to enhance enzymatic activity using the substrate affinity. Nat. Commun. *14*, 4860. https://doi.org/10.1038/s41467-023-40471-y.
7. Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., and Milo, R. (2014). Pathway Thermodynamics Highlights Kinetic Obstacles in Central Metabolism. PLoS Comput. Biol. *10*, e1003483. https://doi.org/10.1371/journal.pcbi.1003483.
8. Niebel, B., Leupold, S., and Heinemann, M. (2019). An upper limit on Gibbs energy dissipation governs cellular metabolism. Nat. Metab. *1*, 125–132. https://doi.org/10.1038/s42255-018-0006-7.
9. Brunk, E., Sahoo, S., Zielinski, D.C., Altunkaya, A., Dräger, A., Mih, N., Gatto, F., Nilsson, A., Preciat Gonzalez, G.A., Aurich, M.K., et al. (2018). Recon3D enables a three-dimensional view of gene variation in human metabolism. Nat. Biotechnol. *36*, 272–281. https://doi.org/10.1038/nbt.4072.
10. Martínez, V.S., Quek, L.-E., and Nielsen, L.K. (2014). Network Thermodynamic Curation of Human and Yeast Genome-Scale Metabolic Models. Biophys. J. *107*, 493–503. https://doi.org/10.1016/j.bpj.2014.05.029.
11. Haraldsdóttir, H.S., Thiele, I., and Fleming, R.M.T. (2012). Quantitative Assignment of Reaction Directionality in a Multicompartmental Human Metabolic Reconstruction. Biophys. J. *102*, 1703–1711. https://doi.org/10.1016/j.bpj.2012.02.032.
12. Hamilton, J.J., Dwivedi, V., and Reed, J.L. (2013). Quantitative Assessment of Thermodynamic Constraints on the Solution Space of Genome-Scale Metabolic Models. Biophys. J. *105*, 512–522. https://doi.org/10.1016/j.bpj.2013.06.011.
13. Schultz, A., and Qutub, A.A. (2015). Predicting internal cell fluxes at sub-optimal growth. BMC Syst. Biol. *9*, 18. https://doi.org/10.1186/s12918-015-0153-3.
14. Kiparissides, A., and Hatzimanikatis, V. (2017). Thermodynamics-based Metabolite Sensitivity Analysis in metabolic networks. Metab. Eng. *39*, 117–127. https://doi.org/10.1016/j.ymben.2016.11.006.
15. Hädicke, O., von Kamp, A., Aydogan, T., and Klamt, S. (2018). OptMDFpathway: Identification of metabolic pathways with maximal thermodynamic driving force and its application for analyzing the endogenous CO2 fixation potential of Escherichia coli. PLoS Comput. Biol. *14*, e1006492. https://doi.org/10.1371/journal.pcbi.1006492.
16. Mori, M., Cheng, C., Taylor, B.R., Okano, H., and Hwa, T. (2023). Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions. Nat. Commun. *14*, 4161. https://doi.org/10.1038/s41467-023-39724-7.
17. Ataman, M., and Hatzimanikatis, V. (2015). Heading in the right direction: thermodynamics-based network analysis and pathway engineering. Curr. Opin. Biotechnol. *36*, 176–182. https://doi.org/10.1016/j.copbio.2015.08.021.
18. Park, J.O., Rubin, S.A., Xu, Y.-F., Amador-Noguez, D., Fan, J., Shlomi, T., and Rabinowitz, J.D. (2016). Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. Nat. Chem. Biol. *12*, 482–489. https://doi.org/10.1038/nchembio.2077.
19. Liberti, M.V., Dai, Z., Wardell, S.E., Baccile, J.A., Liu, X., Gao, X., Baldi, R., Mehrmohamadi, M., Johnson, M.O., Madhukar, N.S., et al. (2017). A Predictive Model for Selective Targeting of the Warburg Effect through GAPDH Inhibition with a Natural Product. Cell Metab. *26*, 648–659.e8. https://doi.org/10.1016/j.cmet.2017.08.017.
20. Tanner, L.B., Goglia, A.G., Wei, M.H., Sehgal, T., Parsons, L.R., Park, J.O., White, E., Toettcher, J.E., and Rabinowitz, J.D. (2018). Four Key Steps Control Glycolytic Flux in Mammalian Cells. Cell Syst. *7*, 49–62.e8. https://doi.org/10.1016/j.cels.2018.06.003.
21. Park, J.O., Tanner, L.B., Wei, M.H., Khana, D.B., Jacobson, T.B., Zhang, Z., Rubin, S.A., Li, S.H.-J., Higgins, M.B., Stevenson, D.M., et al. (2019). Near-equilibrium glycolysis supports metabolic homeostasis and energy yield. Nat. Chem. Biol. *15*, 1001–1008. https://doi.org/10.1038/s41589-019-0364-9.
22. Robinson, J.L., Kocabaş, P., Wang, H., Cholley, P.-E., Cook, D., Nilsson, A., Anton, M., Ferreira, R., Domenzain, I., Billa, V., et al. (2020). An atlas of human metabolism. Sci. Signal. *13*, eaaz1482. https://doi.org/10.1126/scisignal.aaz1482.
23. Goldberg, R.N., Tewari, Y.B., and Bhat, T.N. (2004). Thermodynamics of enzyme-catalyzed reactions—a database for quantitative biochemistry. Bioinformatics *20*, 2874–2877. https://doi.org/10.1093/bioinformatics/bth314.
24. Krebs, H.A., and Kornberg, H.L. (1957). Energy Transformations in Living Matter. In Energy Transformations in Living Matter, H.A. Krebs and H.L. Kornberg, eds. (Springer Berlin Heidelberg), pp. 212–298. https://doi.org/10.1007/978-3-642-86577-0_1.
25. Thauer, R.K., Jungermann, K., and Decker, K.H. (1977). Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriol. Rev. *41*, 809. https://doi.org/10.1128/br.41.3.809-809.1977.
26. Alberty, R.A. (2006). Biochemical thermodynamics: applications of Mathematica. Methods Biochem. Anal. *48*, 1–458.
27. Xu, J., Martien, J., Gilbertson, C., Ma, J., Amador-Noguez, D., and Park, J. O. (2020). Metabolic flux analysis and fluxomics-driven determination of reaction free energy using multiple isotopes. Curr. Opin. Biotechnol. *64*, 151–160. https://doi.org/10.1016/j.copbio.2020.02.018.
28. Bar-Even, A., Flamholz, A., Noor, E., and Milo, R. (2012). Thermodynamic constraints shape the structure of carbon fixation pathways. Biochim. Biophys. Acta *1817*, 1646–1659. https://doi.org/10.1016/j.bbabio.2012.05.002.
29. Wiechert, W. (2007). The Thermodynamic Meaning of Metabolic Exchange Fluxes. Biophys. J. *93*, 2255–2264. https://doi.org/10.1529/biophysj.106.099895.
30. Beard, D.A., Liang, S.D., and Qian, H. (2002). Energy Balance for Analysis of Complex Metabolic Networks. Biophys. J. *83*, 79–86. https://doi.org/10.1016/S0006-3495(02)75150-3.
31. Kümmel, A., Panke, S., and Heinemann, M. (2006). Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. Mol. Syst. Biol. *2*, 2006.0034. https://doi.org/10.1038/msb4100074.
32. Henry, C.S., Broadbelt, L.J., and Hatzimanikatis, V. (2007). Thermodynamics-Based Metabolic Flux Analysis. Biophys. J. *92*, 1792–1805. https://doi.org/10.1529/biophysj.106.093138.
33. Soh, K.C., and Hatzimanikatis, V. (2014). Constraining the Flux Space Using Thermodynamics and Integration of Metabolomics Data. MethodsMol. Biol. *1191*, 49–63. https://doi.org/10.1007/978-1-4939-1170-7_3.
34. Dash, S., Olson, D.G., Joshua Chan, S.H., Amador-Noguez, D., Lynd, L.R., and Maranas, C.D. (2019). Thermodynamic analysis of the pathway for

ethanol production from cellobiose in Clostridium thermocellum. Metab. Eng. *55*, 161–169. https://doi.org/10.1016/j.ymben.2019.06.006.

35. Krumholz, E.W., and Libourel, I.G.L. (2017). Thermodynamic Constraints Improve Metabolic Networks. Biophys. J. *113*, 679–689. https://doi.org/10.1016/j.bpj.2017.06.018.

36. Beber, M.E., Gollub, M.G., Mozaffari, D., Shebek, K.M., Flamholz, A.I., Milo, R., and Noor, E. (2022). eQuilibrator 3.0: a database solution for thermodynamic constant estimation. Nucleic Acids Res. *50*, D603–D609. https://doi.org/10.1093/nar/gkab1106.

37. Benson, S.W., and Buss, J.H. (1958). Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties. J. Chem. Phys. *29*, 546–572. https://doi.org/10.1063/1.1744539.

38. Boudart, M. (1977). Thermochemical kinetics, Second Edition., ed. (Wiley Interscience). https://doi.org/10.1002/aic.690230437.

39. Ritter, E.R., and Bozzelli, J.W. (1991). THERMODYNAMIC PROPERTY ESTIMATION FOR GAS PHASE RADICALS and MOLECULES. Int. J. Chem. Kinet. *23*, 767–778.

40. Domalski, E.S., and Hearing, E.D. (1988). Estimation of the Thermodynamic Properties of Hydrocarbons at 298.15 K. J. Phys. Chem. Ref. Data *17*, 1637–1678. https://doi.org/10.1063/1.555814.

41. Mavrovouniotis, M.L., Bayol, P., Lam, T.-K.M., Stephanopoulos, G., and Stephanopoulos, G. (1988). A group contribution method for the estimation of equilibrium constants for biochemical reactions. Biotechnol. Tech. *2*, 23–28. https://doi.org/10.1007/BF01874203.

42. Noor, E., Haraldsdóttir, H.S., Milo, R., and Fleming, R.M.T. (2013). Consistent Estimation of Gibbs Energy Using Component Contributions. PLoS Comput. Biol. *9*, e1003098. https://doi.org/10.1371/journal.pcbi.1003098.

43. Alazmi, M., Kuwahara, H., Soufan, O., Ding, L., and Gao, X. (2019). Systematic selection of chemical fingerprint features improves the Gibbs energy prediction of biochemical reactions. Bioinformatics *35*, 2634–2643. https://doi.org/10.1093/bioinformatics/bty1035.

44. Wang, L., Upadhyay, V., and Maranas, C.D. (2021). dGPredictor: Automated fragmentation method for metabolic reaction free energy prediction and de novo pathway design. PLoS Comput. Biol. *17*, e1009448. https://doi.org/10.1371/journal.pcbi.1009448.

45. Du, B., Zhang, Z., Grubner, S., Yurkovich, J.T., Palsson, B.O., and Zielinski, D.C. (2018). Temperature-Dependent Estimation of Gibbs Energies Using an Updated Group-Contribution Method. Biophys. J. *114*, 2691–2702. https://doi.org/10.1016/j.bpj.2018.04.030.

46. Jankowski, M.D., Henry, C.S., Broadbelt, L.J., and Hatzimanikatis, V. (2008). Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks. Biophys. J. *95*, 1487–1499. https://doi.org/10.1529/biophysj.107.124784.

47. Noor, E., Bar-Even, A., Flamholz, A., Lubling, Y., Davidi, D., and Milo, R. (2012). An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. Bioinformatics *28*, 2037–2044. https://doi.org/10.1093/bioinformatics/bts317.

48. Joshi, R.P., McNaughton, A., Thomas, D.G., Henry, C.S., Canon, S.R., McCue, L.A., and Kumar, N. (2021). Quantum Mechanical Methods Predict Accurate Thermodynamics of Biochemical Reactions. ACS Omega *6*, 9948–9959. https://doi.org/10.1021/acsomega.1c00997.

49. Jinich, A., Rappoport, D., Dunn, I., Sanchez-Lengeling, B., Olivares-Amaya, R., Noor, E., Even, A.B., and Aspuru-Guzik, A. (2014). Quantum Chemical Approach to Estimating the Thermodynamics of Metabolic Reactions. Sci. Rep. *4*, 7022. https://doi.org/10.1038/srep07022.

50. van Speybroeck, V., Gani, R., and Meier, R.J. (2010). The calculation of thermodynamic properties of molecules. Chem. Soc. Rev. *39*, 1764–1779. https://doi.org/10.1039/B809850F.

51. Yang, Q., Ji, H., Lu, H., and Zhang, Z. (2021). Prediction of Liquid Chromatographic Retention Time with Graph Neural Networks to Assist in Small Molecule Identification. Anal. Chem. *93*, 2200–2206. https://doi.org/10.1021/acs.analchem.0c04071.

52. Ma, H., Bian, Y., Rong, Y., Huang, W., Xu, T., Xie, W., Ye, G., and Huang, J. (2022). Cross-dependent graph neural networks for molecular property prediction. Bioinformatics *38*, 2003–2009. https://doi.org/10.1093/bioinformatics/btac039.

53. Fang, X., Liu, L., Lei, J., He, D., Zhang, S., Zhou, J., Wang, F., Wu, H., and Wang, H. (2022). Geometry-enhanced molecular representation learning for property prediction. Nat. Mach. Intell. *4*, 127–134. https://doi.org/10.1038/s42256-021-00438-4.

54. Wang, Y., Wang, J., Cao, Z., and Barati Farimani, A. (2022). Molecular contrastive learning of representations via graph neural networks. Nat. Mach. Intell. *4*, 279–287. https://doi.org/10.1038/s42256-022-00447-x.

55. Fang, Y., Zhang, Q., Zhang, N., Chen, Z., Zhuang, X., Shao, X., Fan, X., and Chen, H. (2023). Knowledge graph-enhanced molecular contrastive learning with functional prompt. Nat. Mach. Intell. *5*, 542–553. https://doi.org/10.1038/s42256-023-00654-0.

56. Roszak, R., Beker, W., Molga, K., and Grzybowski, B.A. (2019). Rapid and Accurate Prediction of pKa Values of C–H Acids Using Graph Convolutional Neural Networks. J. Am. Chem. Soc. *141*, 17142–17149. https://doi.org/10.1021/jacs.9b05895.

57. Park, C.W., Kornbluth, M., Vandermause, J., Wolverton, C., Kozinsky, B., and Mailoa, J.P. (2021). Accurate and scalable graph neural network force field and molecular dynamics with direct force architecture. npj Comput. Mater. *7*, 73. https://doi.org/10.1038/s41524-021-00543-3.

58. Xiong, J., Xiong, Z., Chen, K., Jiang, H., and Zheng, M. (2021). Graph neural networks for automated de novo drug design. Drug Discov. Today *26*, 1382–1393. https://doi.org/10.1016/j.drudis.2021.02.011.

59. Wong, F., Zheng, E.J., Valeri, J.A., Donghia, N.M., Anahtar, M.N., Omori, S., Li, A., Cubillos-Ruiz, A., Krishnan, A., Jin, W., et al. (2024). Discovery of a structural class of antibiotics with explainable deep learning. Nature *626*, 177–185. https://doi.org/10.1038/s41586-023-06887-8.

60. Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., et al. (2019). Analyzing Learned Molecular Representations for Property Prediction. J. Chem. Inf. Model. *59*, 3370–3388. https://doi.org/10.1021/acs.jcim.9b00237.

61. Liu, C., Sun, Y., Davis, R., Cardona, S.T., and Hu, P. (2023). ABT-MPNN: an atom-bond transformer-based message-passing neural network for molecular property prediction. J. Cheminform. *15*, 29. https://doi.org/10.1186/s13321-023-00698-9.

62. Kümmel, A., Panke, S., and Heinemann, M. (2006). Systematic assignment of thermodynamic constraints in metabolic network models. BMC Bioinformatics *7*, 512. https://doi.org/10.1186/1471-2105-7-512.

63. Thiele, I., and Palsson, B.Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. *5*, 93–121. https://doi.org/10.1038/nprot.2009.203.

64. Bresson, E., Lacroix-Pépin, N., Boucher-Kovalik, S., Chapdelaine, P., and Fortier, M.A. (2012). The prostaglandin F synthase activity of the human aldose reductase AKR1B1 brings new lenses to look at pathologic conditions. Front. Pharmacol. *3*, 98. https://doi.org/10.3389/fphar.2012.00098.

65. Michaud, A., Lacroix-Pépin, N., Pelletier, M., Veilleux, A., Noël, S., Bouchard, C., Marceau, P., Fortier, M.A., and Tchernof, A. (2014). Prostaglandin (PG) F2 Alpha Synthesis in Human Subcutaneous and Omental Adipose Tissue: Modulation by Inflammatory Cytokines and Role of the Human Aldose Reductase AKR1B1. PLoS One *9*, e90861. https://doi.org/10.1371/journal.pone.0090861.

66. Tomsho, J.W., Moran, R.G., and Coward, J.K. (2008). Concentration-Dependent Processivity of Multiple Glutamate Ligations Catalyzed by Folylpoly-γ-glutamate Synthetase. Biochemistry *47*, 9040–9050. https://doi.org/10.1021/bi800406w.

67. Sheng, Y., Khanam, N., Tsaksis, Y., Shi, X.M., Lu, Q.S., and Bognar, A.L. (2008). Mutagenesis of Folylpolyglutamate Synthetase Indicates That Dihydropteroate and Tetrahydrofolate Bind to the Same Site. Biochemistry *47*, 2388–2396. https://doi.org/10.1021/bi701670y.

68. Monk, J.M., Lloyd, C.J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., et al. (2017). iML1515, a

knowledgebase that computes Escherichia coli traits. Nat. Biotechnol. *35*, 904–908. https://doi.org/10.1038/nbt.3956.

69. Zhang, C., Sánchez, B.J., Li, F., Eiden, C.W.Q., Scott, W.T., Liebal, U.W., Blank, L.M., Mengers, H.G., Anton, M., Rangel, A.T., et al. (2024). Yeast9: a consensus genome-scale metabolic model for *S. cerevisiae* curated by the community. Mol. Syst. Biol. *20*, 1134–1150. https://doi.org/10.1038/s44320-024-00060-7.
70. Antoniewicz, M.R., Kelleher, J.K., and Stephanopoulos, G. (2007). Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. Metab. Eng. *9*, 68–86. https://doi.org/10.1016/j.ymben.2006.09.001.
71. McAtee Pereira, A.G., Walther, J.L., Hollenbach, M., and Young, J.D. (2018). 13C Flux Analysis Reveals that Rebalancing Medium Amino Acid Composition can Reduce Ammonia Production while Preserving Central Carbon Metabolism of CHO Cell Cultures. Biotechnol. J. *13*, e1700518. https://doi.org/10.1002/biot.201700518.
72. Hefzi, H., Ang, K.S., Hanscho, M., Bordbar, A., Ruckerbauer, D., Lakshmanan, M., Orellana, C.A., Baycin-Hizal, D., Huang, Y., Ley, D., et al. (2016). A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. Cell Syst. *3*, 434–443.e8. https://doi.org/10.1016/j.cels.2016.10.020.
73. Goldford, J.E., George, A.B., Flamholz, A.I., and Segrè, D. (2022). Protein cost minimization promotes the emergence of coenzyme redundancy. Proc. Natl. Acad. Sci. USA *119*, e2110787119. https://doi.org/10.1073/pnas.2110787119.
74. Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2011). Controllability of complex networks. Nature *473*, 167–173. https://doi.org/10.1038/nature10011.
75. Basler, G., Nikoloski, Z., Larhlimi, A., Barabási, A.-L., and Liu, Y.-Y. (2016). Control of fluxes in metabolic networks. Genome Res. *26*, 956–968. https://doi.org/10.1101/gr.202648.115.
76. Gagler, D.C., Karas, B., Kempes, C.P., Malloy, J., Mierzejewski, V., Goldman, A.D., Kim, H., and Walker, S.I. (2022). Scaling laws in enzyme function reveal a new kind of biochemical universality. Proc. Natl. Acad. Sci. USA *119*, e2106655119. https://doi.org/10.1073/pnas.2106655119.
77. Nusinow, D.P., Szpyt, J., Ghandi, M., Rose, C.M., McDonald, E.R., III, Kalocsay, M., Jané-Valbuena, J., Gelfand, E., Schweppe, D.K., Jedrychowski, M., et al. (2020). Quantitative Proteomics of the Cancer Cell Line Encyclopedia. Cell *180*, 387–402.e16. https://doi.org/10.1016/j.cell.2019.12.023.
78. Xiao, Z., Dai, Z., and Locasale, J.W. (2019). Metabolic landscape of the tumor microenvironment at single cell resolution. Nat. Commun. *10*, 3763. https://doi.org/10.1038/s41467-019-11738-0.
79. Heinrich, R., and Rapoport, T.A. (1974). A Linear Steady-State Treatment of Enzymatic Chains. General properties, control and effector strength. Eur. J. Biochem. *42*, 89–95. https://doi.org/10.1111/j.1432-1033.1974.tb03318.x.
80. Locasale, J.W. (2018). New concepts in feedback regulation of glucose metabolism. Curr. Opin. Syst. Biol. *8*, 32–38. https://doi.org/10.1016/j.coisb.2017.11.005.
81. Gustafsson, J., Robinson, J.L., Zetterberg, H., and Nielsen, J. (2024). Brain energy metabolism is optimized to minimize the cost of enzyme synthesis and transport. Proc. Natl. Acad. Sci. USA *121*, e2305035121. https://doi.org/10.1073/pnas.2305035121.
82. Dourado, H., Mori, M., Hwa, T., and Lercher, M.J. (2021). On the optimality of the enzyme–substrate relationship in bacteria. PLoS Biol. *19*, e3001416. https://doi.org/10.1371/journal.pbio.3001416.
83. Zhang, H., Wang, Y., Li, J., Chen, H., He, X., Zhang, H., Liang, H., and Lu, J. (2018). Biosynthetic energy cost for amino acids decreases in cancer evolution. Nat. Commun. *9*, 4124. https://doi.org/10.1038/s41467-018-06461-1.
84. Segrè, D., Vitkup, D., and Church, G.M. (2002). Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. USA *99*, 15112–15117. https://doi.org/10.1073/pnas.232349399.
85. Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., and Milo, R. (2011). The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. Biochemistry *50*, 4402–4410. https://doi.org/10.1021/bi2002289.
86. Rapoport, T.A. (2024). A Berlin-sided retrospective of the origins of metabolic control theory. Interface Focus *14*, 20230024. https://doi.org/10.1098/rsfs.2023.0024.
87. Leng, H., Wang, Y., Zhao, W., Sievert, S.M., and Xiao, X. (2023). Identification of a deep-branching thermophilic clade sheds light on early bacterial evolution. Nat. Commun. *14*, 4354. https://doi.org/10.1038/s41467-023-39960-x.
88. Schellenberger, J., Park, J.O., Conrad, T.M., and Palsson, B.Ø. (2010). BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinformatics *11*, 213. https://doi.org/10.1186/1471-2105-11-213.
89. King, Z.A., Dräger, A., Ebrahim, A., Sonnenschein, N., Lewis, N.E., and Palsson, B.O. (2015). Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. PLoS Comput. Biol. *11*, e1004321. https://doi.org/10.1371/journal.pcbi.1004321.
90. Alberty, R.A. (2002). Inverse Legendre Transform in Biochemical Thermodynamics: Illustrated with the Last Five Reactions of Glycolysis. J. Phys. Chem. B *106*, 6594–6599. https://doi.org/10.1021/jp020764w.
91. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. (2017) Automatic Differentiation in PyTorch.
92. Fey, M., and Lenssen, J.E. (2019) Fast Graph Representation Learning with PyTorch Geometric. Inproceedings.
93. Gurobi Optimization, L.L.C. (2024) Gurobi Optimizer Reference Manual.
94. Koopmans, F., van Nierop, P., Andres-Alonso, M., Byrnes, A., Cijsouw, T., Coba, M.P., Cornelisse, L.N., Farrell, R.J., Goldschmidt, H.L., Howrigan, D.P., et al. (2019). SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. Neuron *103*, 217–234.e4. https://doi.org/10.1016/j.neuron.2019.05.002.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Experimental thermodynamics data	Beber et al.[36]	https://zenodo.org/records/5495826
Training dataset for dGbyG	This paper	https://gitee.com/f-wc/dGbyG; https://doi.org/10.5281/zenodo.15663012
Cross-validation results of dGbyG and other methods	This paper	https://zenodo.org/records/13885478
Recon3D_301	Brunk et al.[9]	https://www.vmh.life/
Human1 (v1.18.0)	Robinson et al.[22]	https://zenodo.org/records/10303455
Yeast9 (version 9.0.2)	Zhang et al.[69]	https://github.com/SysBioChalmers/yeast-GEM
iML1515	Monk et al.[68]	http://bigg.ucsd.edu/models/iML1515
iCHO1766	Hefzi et al.[72]	http://bigg.ucsd.edu/models/iCHOv1
Proteomics data of CCLE cell lines	Nusinow et al.[77]	https://sites.broadinstitute.org/ccle/datasets
Software and algorithms		
Python 3.12.2	Python Software Foundation	https://www.python.org/
NumPy 1.26.4	NumPy Developers	https://numpy.org/
pandas 2.1.4	The pandas development team	https://pandas.pydata.org/
PyTorch 2.2.1	PyTorch Foundation	https://pytorch.org/
PyG 2.5.2	PyG Team	https://www.pyg.org/
RDKit 2023.09.6	Greg Landrum and other RDKit contributors	https://www.rdkit.org/
MATLAB R2022b	MathWorks	https://mathworks.com/
GNU Scientific Library	The GNU Project	https://www.gnu.org/software/gsl/
Calculator plugins, Marvin 24.3.1	Chemaxon Ltd.	https://chemaxon.com/
Gurobi 12.0.1	Gurobi Optimization, LLC.	https://www.gurobi.com/
equilibrator-api 0.6.0	Beber et al.[36]	https://pypi.org/project/equilibrator-api/
dGPredictor	Wang et al.[44]	https://github.com/maranasgroup/dGPredictor
Escher	King et al.[89]	https://escher.github.io/
Code for data pre-processing, cross-validation of dGbyG and other methods, training of dGbyG, and prediction standard Gibbs energy for GEMs	This paper	https://gitee.com/f-wc/dGbyG; https://doi.org/10.5281/zenodo.15663012
Code for MFA and TFBA simulation	This paper	https://gitee.com/f-wc/ThermoInfer; https://doi.org/10.5281/zenodo.15663012
Code for analysis of TDRs	This paper	https://github.com/ziweidai/GEM_thermodynamics; https://doi.org/10.5281/zenodo.15663012
Code for metabolic control analysis of glycolysis	Dai and Locasale[5]	https://github.com/ziweidai/MCA_thermodynamics/tree/master/Glycolysis

METHOD DETAILS

Pre-processing of training data

We first mapped different types of molecular identifiers to SMILES formulas and removed reactions that involved metabolites without clearly defined molecular structures (e.g. tRNA), or lacked chemical equations or equilibrium constants. We then balanced the reactions by adding water molecules to either side of the unbalanced reactions, and discarded reactions that cannot be balanced using this procedure. Based on the pKa values of the molecules predicted by Calculator plugins, Marvin 24.3.1, ChemAxon (https://www.chemaxon.com), we next applied the Legendre transformation[90] to compute the standard Gibbs free energy of the reactions under the same condition of pH = 7.0, T = 298.15 K, ionic strength = 0.25 mol/L, and pMg = 14.0. Finally, we computed the mean and standard deviation (SD) of standard Gibbs free energy change averaged over different measurements of each reaction, $\Delta_r G^{\circ}_{Observed}$, and the mean and SD of standard Gibbs free energy of formation averaged over different measurements of each molecule, $\Delta_f G^{\circ}_{Observed}$. The mean and SD values were used for the training.

Legendre transformation

We applied Legendre transformation to convert all Gibbs free energies experimentally measured under non-standard conditions into the standard condition. Here, we defined pH = 7.0, T = 298.15 K, ionic strength = 0.25 mol/L, and pMg = 14.0 as the standard condition, $\Delta_f G^{\circ}$ as standard Gibbs free energy of formation under the standard condition, and $\Delta_f G'^{\circ}$ as standard formation Gibbs energy under a non-standard condition.

In solutions, metabolites can dissociate into various forms, i.e. pseudoisomers, by losing different numbers of protons. $\Delta_f G^{\circ}$ of a metabolite was computed using a Boltzmann-weighted average of $\Delta_f G^{\circ}$ of its pseudoisomers[47]:

$$\Delta_f G^{\circ} = -RTln \sum_j e^{-\frac{\Delta_f G^{\circ}(j)}{RT}}$$

where $\Delta_f G^{\circ}(j)$ is the standard Gibbs free energy of formation for pseudoisomer j. The Legendre transformation can be therefore applied to each pseudoisomer to compute its $\Delta_f G^{\circ}(j)$ from the value of $\Delta_f G'^{\circ}$ under non-standard conditions:

$$\Delta_f G'^{\circ}(j) = \Delta_f G^{\circ}(j) + \Delta\Theta_H(\mathrm{j}) + \Delta\Theta_{Mg}(\mathrm{j}) + \Delta\Theta_I(\mathrm{j})$$

where

$$\Delta\Theta_H(j) = N_H(j)RT(pH - 7)\ln(10)\Delta\Theta_{Mg}(j) = N_{Mg}(j)RT(pMg - 14)\ln(10)$$

$$\Delta\Theta_I(j) = -2.91482\left(z^2 - N_H\right)\left(\frac{\sqrt{I}}{1+1.6\sqrt{I}} - \frac{\sqrt{0.25}}{1+1.6\sqrt{0.25}}\right)$$

In which $\Delta\Theta_H(j)$ reflects the contribution of protons to the Gibbs energy, $N_H(j)$ is the number of protons associated to this pseudoisomer, $\Delta\Theta_{M_g}(j)$ reflects the contribution of Mg^{2+} ions to the Gibbs energy, $N_{M_g}(j)$ is the number of Magnesium ions associated to this pseudoisomer. $\Delta\Theta_I(j)$reflects the contribution of ionic strength to the Gibbs energy and is defined based on the extended Debye-Hückel equation, which works well in the 0.05 to 0.25 M range of ionic strengths.[47] z is the total charge of this pseudoisomer. I is the ionic strength of the solution. Let $[M_j]$ represent pseudoisomer with j hydrogen atoms and $[M_{j+1}]$ represent pseudoisomer with $j+1$ hydrogen atoms. When the disassociation constant $K_a(j)$ is known, the relationship below holds:

$$\Delta_f G'^{\circ}(j+1) = \Delta_f G'^{\circ}(j) + \Delta_f G'^{\circ}(H^+) - RT\ln(10)pK_a(j)$$

where $\Delta_f G'^{\circ}(H^+)$ is the standard Gibbs free energy of formation for a proton, which is set to be 0 kJ/mol in this framework. Therefore, the difference in Gibbs free energy of formation between the standard and non-standard conditions can be computed as follows:

$$\begin{aligned}\Delta\Delta_f G'^{\circ} = \Delta_f G'^{\circ} - \Delta_f G^{\circ} &= \left(-RT\ln\sum_{j=m} e^{\ln(10)\sum_{i=m}^{j} pK_a(i) - \frac{\Delta\Theta_H(j)+\Delta\Theta_{Mg}(j)+\Delta\Theta_I(j)}{RT}}\right) - \left(-RT\ln\sum_{j=m} e^{\ln(10)\sum_{i=m}^{j} pK_a(i)}\right)\\ &= -RT\left(\ln\sum_{j=m} e^{\ln(10)\sum_{i=m}^{j} pK_a(i) - \frac{\Delta\Theta_H(j)+\Delta\Theta_{Mg}(j)+\Delta\Theta_I(j)}{RT}} - \ln\sum_{j=m} e^{\ln(10)\sum_{i=m}^{j} pK_a(i)}\right)\end{aligned}$$

For experimental values in the training set measured at non-standard conditions, these values were transformed to the standard condition during the pre-processing steps using the following formula:

$$\Delta_f G^{\circ} = \Delta_f G'^{\circ} - \Delta\Delta_f G'^{\circ}$$

As the direct prediction of Gibbs free energy by dGbyG is under the standard condition, predictions for non-standard conditions (e.g. standard Gibbs free energy change for a reaction in a certain cellular compartment) can be calculated as follows:

$$\Delta_f G'^{\circ} = \Delta_f G^{\circ} + \Delta\Delta_f G'^{\circ}$$

Structure of GNN model

dGbyG first obtains molecular structures of metabolites from SMILES or InChI strings, or retrieves the corresponding MOL files based on their database IDs using RDKit (https://www.rdkit.org/) if SMILES or InChI formulas are unavailable for the metabolite. The molecular structures were processed through standardized transformations using the rdkit.Chem.MolStandardize.rdMolStandardize.Normalize function in RDKit to adjust functional groups and optimize charge distributions. This included charge neutralization through protonation and deprotonation steps, followed by the addition of missing hydrogen atoms. dGbyG next converted all structures to molecular graphs with atoms as nodes and bonds as edges, recording atomic number, hybridization, aromaticity, and charge of atoms as node features and type of bonds as edge features. The neural network model begins with one-hot encoding of the molecular graph, which is then transformed into a 300-dimensional latent space for both nodes and edges using embedding layers. Next, two message passing neural network (MPNN) layers propagate and aggregate the message from bonds and neighboring atoms for each atom together with a residual of its own features. This process produces the features of each atom without changing the dimensionality of the latent space. Subsequently, three feedforward neural network (FNN) layers consisting of a ReLU activation function followed by linear layers with input and output dimensions of (300,300), (300,150), and (150,1), respectively, are used to transform the feature vector of each atom to a single value. An additive pooling layer sums up all single values of the atoms to compute the $\Delta_f G^\circ$ as the final output of the neural network. Finally, we compute the standard Gibbs free energy change for a metabolic reaction with the following chemical equation from the model-predicted values of $\Delta_f G^\circ$ of the substrates $\{S_i\}$ and products $\{P_i\}$ and the stoichiometric coefficients $\{a_i\}$ and $\{b_i\}$:

$$\sum_{i=1}^{m} a_i S_i \leftrightarrow \sum_{i=1}^{n} b_i P_i$$

$$\Delta_r G^\circ = \sum_{i=1}^{n} b_i \Delta_f G^\circ_{P_i} - \sum_{i=1}^{m} a_i \Delta_f G^\circ_{S_i}$$

The whole framework is implemented based on PyTorch[91] and PyG.[92]

Estimation of uncertainty in the prediction

To define a metric to quantify the variability in experimental measurements of Gibbs free energy, we quantified the relationship between its standard deviation (SD), standard error of the mean (SEM), absolute value, and sample size (i.e. number of different measurements) for each reaction. We did not observe strong correlation between the standard deviation (SD) of experimentally measured Gibbs energy values and other data features (Figures S2A and S2B), but the significant positive correlation between the standard error of the mean (SEM) and number of measurements suggests that the variability in the mean values decreases as the sample size increases, in other words, measurements with lower SEM are more reliable than the other ones (Figure S2C). Hence, we define a variability index α to assess the variability in the experimental measurements of standard Gibbs energy for each reaction:

$$\alpha_i = \sqrt{SEM_i^2 + \frac{SD^2_{\text{mean}}}{N_i}}$$

where SIM_i is the SEM of the i-th reaction, SD_{mean} is the mean SD of all reactions in the training set, N_i is the number of different measurements for the i-th reaction. The design of α aims to stabilize the variability in the data point by approximating it according to the average SD of all reactions when the sample size N is small. Otherwise, when the sample size N is large, the variability index converges to the actual SEM of the reaction. Under the assumption that the error in the mean value of experimentally measured standard Gibbs energy for each reaction follows a normal distribution with standard deviation α equals to the variability index defined above, we randomly sampled 100 error terms from that distribution for each reaction and added it to the original value in the training set to generate 100 different training sets. We then used those 100 training sets to train 100 different networks, resulting in an ensemble of models with different parameters and predictions. The predictions made by 100 networks in the ensemble of models were used to estimate uncertainty in the prediction.

Computation of $\Delta_r G^\circ$ for multi-compartment reactions in GEMs

Electrostatic potentials difference between cellular compartments could affect the free energy change of metabolite transport and multi-compartment reactions that involve movement of charged metabolites between compartments. This effect was quantified by computing the contribution of charged species movements between compartments to $\Delta_r G^\circ$ using the relationship $\Delta G^\circ_e = N_{\text{charge}} F \Delta \Psi$, where N_{charge} is the net charge of the molecule moving between compartments, F is the Faraday's constant, and $\Delta \Psi$ is the membrane potential difference between compartments. This term ΔG°_e was added to the dGbyG-predicted $\Delta_r G^\circ$ value to compute the total $\Delta_r G^\circ$ of multi-compartment reactions.

Thermodynamics-based Flux Balance Analysis (TFBA)

We developed the framework of TFBA by introducing additional constraints linking fluxes, metabolite concentrations, and reaction thermodynamics to the original mathematical problem of flux balance analysis (FBA). Briefly, we incorporated variables corresponding to $\Delta_f G^\circ$ of each metabolite, $\Delta_r G^\circ$ and $\Delta_r G$ of each reaction, and logarithms of the metabolite concentrations, $z = \log_{10}(c)$, into the original FBA model$\Delta_r G^\circ$. $\Delta_r G^\circ$ was constrained in the range of uncertainty of $\Delta_r G^\circ_p$ for reactions predicted

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

by dGbyG. Thermodynamic constraints on $\Delta_r G$ of each reaction that require the net flux v and the free energy change $\Delta_r G$ to have opposite signs were then added to ensure thermodynamic feasibility of the reactions. Finally, an osmotic constraint on the upper bound of total metabolite concentration in each compartment, and a growth constraint that ensures non-zero biomass production, were added to the model.

The first part of constraints included in the TFBA framework sets the conditions of flux balance and lower and upper bounds of flux for each reaction:

$$Sv = 0,$$

$$v_l \le v \le v_u$$

where S is a $m \times r$ stoichiometric matrix, v is an $r \times 1$ flux vector, m is the total number of metabolites and r is the total number of reactions. v_l and v_u are two $r \times 1$ vectors representing the lower and upper bounds of fluxes.

The next part of TFBA constraints determines directionalities of the reactions. This was done by introducing these constraints below:

$$\begin{gathered} a \in \{0,1\}, \\ (a-1) \times 1000 \le v \le a \times 1000, -a \times M + \tau < \Delta_r G \le (1-a) \times M - \tau, \end{gathered}$$

where a is an $r \times 1$ vector of binary instrumental variables indicating the directionality of each reaction, $\Delta_r G$ is an $r \times 1$ vector of Gibbs free energy change of reactions, M is a sufficiently large positive number, τ is a small positive number that sets a non-zero lower bound of absolute values of $\Delta_r G$. As a is binary, these constraints become $v \ge 0$ and $\Delta_r G \le -\tau$ when $a = 1$, indicating that the reaction occurs through the forward direction. On the other hand, when $a = 0$ these constraints become $v \le 0$ and $\Delta_r G \ge \tau$, meaning that the reaction takes the backward reaction. Here we used the constraints $\Delta_r G \le -\tau$ or $\Delta_r G \ge \tau$ instead of $\Delta_r G < 0$ or $\Delta_r G > 0$ because TFBA is intrinsically a mixed integer linear programming (MILP) problem, and an inherent property of this kind of mathematical problem is that its solution lies in the boundary of its feasible region. Therefore, using constraints of $\Delta_r G \le -\tau$ or $\Delta_r G \ge \tau$ ensures mathematical consistency (i.e. the feasible region contains its boundary) while preserving thermodynamic constraints. We have also confirmed that the selection of this parameter τ did not affect the predictions of TFBA (data not shown).

The next part of TFBA constraints links reaction Gibbs free energies, $\Delta_r G$, to standard Gibbs free energies, $\Delta_r G^\circ$, and 10-base logarithms of metabolite concentrations, z:

$$\begin{gathered} \Delta_r G = \Delta_r G^\circ + RT \ln(10) S^T z, \\ z = \log_{10}(c), \\ z_l \le z \le z_u, \end{gathered}$$

where z_l and z_u are the lower and upper bounds of log10-transformed metabolite concentration z.

The standard Gibbs free energy change $\Delta_r G^\circ$ of most reactions were predicted by dGbyG or retrieved from experimental data. For those reactions, we computed the confidence intervals of the $\Delta_r G^\circ$ values based on the estimation of uncertainty made by dGbyG or experimental data available from multiple sources. According to the confidence intervals, the following constraints were added:

$$\mu_{l,i} \le \Delta_r G_i^\circ - \Delta_r G_{p,i}^\circ \le \mu_{u,i}, \left\{ i = 1, 2, \cdots, n \middle| \Delta_r G_{p,i}^\circ \text{ is known} \right\}$$

where $\Delta_r G_{p,i}^\circ$ are the known values of standard Gibbs energy change available from computational prediction and experimental measurement, and $\Delta_r G_{p,i}^\circ$ represents possible real value of standard Gibbs energy change of reaction i. $\mu_{l,i}$ and $\mu_{u,i}$are the lower and upper bounds of the error in prediction made by dGbyG or other tools of reaction i.

It is also worth noting that the $\Delta_r G^\circ$ of reactions can be computed from $\Delta_f G^\circ$ of its reactants and products. For this reason, we have the linear equality constraint linking these two:

$$\Delta_r G^\circ = S^T \Delta_f G^\circ$$

We also introduced an osmotic constraint that sets the upper limit of total metabolite concentration in each cellular compartment. To ensure the linearity of this constraint to the logarithm of metabolic concentration, z, this was done by approximating the concentration of each metabolite, c_{appr}, using a piecewise linear function of the logarithm of metabolite concentration z (Figure S12):

$$\begin{gathered} b_1, b_2, b_3 \in \{0,1\}, \\ b_1 + b_2 + b_3 = 1, -2b_1 - 3b_2 - 11b_3 \le z \le -b_1 - 2b_2 - 3b_3, \\ c_{appr} = \left(0.09b_1 + 0.009b_2 - 1.25 \times 10^{-4} b_3\right) z + \left(0.19b_1 + 0.028b_2 - 1.375 \times 10^{-3} b_3\right), \\ \sum_k c_{appr,k} \le C_{upper}, \end{gathered}$$

where C_{upper} is the upper limit of total metabolite concentration in each compartment.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

Finally, a lower bound of the biomass biosynthesis flux was set as $v_{biomass} \geq 10$ to ensure non-zero cell growth. Taken together, the TFBA model contains the following constraints:

$$
\begin{gathered}
Sv = 0, \\
v_l \leq v \leq v_u, \\
\Delta_r G^\circ = S^T \Delta_f G^\circ, \\
\Delta_r G = \Delta_r G^\circ + RT \ln(10) S^T z, \\
z_l \leq z \leq z_u \\
a \in \{0, 1\}, \\
(a - 1) \times 1000 \leq v \leq a \times 1000, \; - a \times M + \tau \leq \Delta_r G \leq (1 - a) \times M - \tau, \\
\mu_{l,i} \leq \Delta_r G_i^\circ - \Delta_r G_{p,i}^\circ \leq \mu_{u,i}, \left\{ i = 1, 2, \cdots, n \middle| \Delta_r G_{p,i}^\circ \text{ is known} \right\}, \\
b_1, b_2, b_3 \in \{0, 1\}, \\
b_1 + b_2 + b_3 = 1, \; - 2b_1 - 3b_2 - 11b_3 \leq z \leq - b_1 - 2b_2 - 3b_3, \\
c_{appr} = \left(0.09b_1 + 0.009b_2 - 1.25 \times 10^{-4} b_3\right) z + \left(0.19b_1 + 0.028b_2 - 1.375 \times 10^{-3} b_3\right), \\
\sum_k c_{appr,k} \leq C_{upper}, \\
v_{biomass} \geq 10,
\end{gathered}
$$

Here we set $M = 4000$, $\tau = 0.0001$, and $C_{upper} = 1$. The objective function is defined based on the task requirements.

Predicting directionality of reactions in GEM

All labels of directions of non-boundary reactions in the GEM were removed before predicting directionalities of reactions in GEM. Next, we computed the ranges of net flux carried by reaction i by maximizing or minimizing its flux v_i upon the FBA:

$$
\textit{min or max } v_i
$$

$$
\text{s.t.} \left\{
\begin{array}{c}
Sv = 0, \\
v_l \leq v \leq v_u, \\
v_{\text{biomass}} \geq 10,
\end{array}
\right.
$$

or TFBA framework:

$$
\textit{min or max } v_i \textit{or } \Delta_r G
$$

$$
\text{s.t.} \left\{
\begin{array}{c}
Sv = 0, \\
v_l \leq v \leq v_u, \\
\Delta_r G^\circ = S^T \Delta_f G^\circ, \\
\Delta_r G = \Delta_r G^\circ + RT \ln(10) S^T z, \\
z_l \leq z \leq z_u, \\
a \in \{0, 1\}, \\
(a - 1) \times 1000 \leq v \leq a \times 1000, \\
-a \times M + \tau \leq \Delta_r G \leq (1 - a) \times M - \tau, \\
\mu_{l,i} \leq \Delta_r G_i^\circ - \Delta_r G_{p,i}^\circ \leq \mu_{u,i}, \left\{ i = 1, 2, \cdots, n \middle| \Delta_r G_{p,i}^\circ \text{ is known} \right\}, \\
b_1, b_2, b_3 \in \{0, 1\}, \\
b_1 + b_2 + b_3 = 1, \\
-2b_1 - 3b_2 - 11b_3 \leq z \leq -b_1 - 2b_2 - 3b_3, \\
c_{appr} = \left(0.09b_1 + 0.009b_2 - 1.25 \times 10^{-4} b_3\right) z + \left(0.19b_1 + 0.028b_2 - 1.375 \times 10^{-3} b_3\right), \\
\sum_k c_{appr,k} \leq C_{upper}, \\
v_{biomass} \geq 10,
\end{array}
\right.
$$

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

A positive maximum value of flux v_i means that the i-th reaction can happen in the forward direction, and a negative minimum value of flux v_i means it can happen in the backward direction.

Note that the flux direction of a reaction in TFBA framework was determined by two factors, thermodynamics and topology of network. The latter affects whether the metabolites that participate in the reaction could be produced or consumed by other reactions while satisfying all flux balance constraints in the model. Considering the imperfections of the metabolic network, we also computed the range of Gibbs energy change for reaction i by solving another two optimization problems with objectives of maximizing or minimizing its Gibbs free energy change $\Delta_r G_i$ under the TFBA framework. The potential directions of reactions could be inferred by the range of $\Delta_r G_i$. A positive minimum value of $\Delta_r G_i$ indicates that the reaction can only happen in the backward direction, while a negative maximum value of $\Delta_r G_i$ means that the reaction can only happen in the forward direction. The Gurobi[93] solver was used for solving those MILP problems. According to the directionalities inferred based on the range of v_i and $\Delta_r G_i$, we screened all non-boundary reactions in the GEM, compared the built-in labels with the reaction directionalities inferred by FBA and TFBA, and assigned the direction manually if they were incompatible.

13C-Metabolic flux analysis (13C-MFA)

To compute intracellular metabolic fluxes, we constructed an Elementary Metabolite Unit[70] (EMU) model to predict ^{13}C mass isotopomer distributions (MIDs) of pyruvate and lactate from steady state metabolic fluxes in a network representing central carbon metabolism of CHO cells. The model was fitted to isotope tracing data obtained from a previously published study.[71] To exclude the possibility of solutions being trapped in local minimums, we applied Latin hypercube sampling to generate 100 distinct flux sets. These flux sets were used as initial solutions for minimization of a loss function defined by the discrepancy between simulated and experimental data with flux balance constraints:

$$\min_{v} \Phi = \left(x(v) - x^{\exp}\right)^T \cdot \sum_{x}^{-1} \cdot \left(x(v) - x^{\exp}\right)$$
$$\text{s.t. } Sv = 0$$

The objective function Φ is the covariance-weighted residual sum of squares (RSS), $x(v)$ is the vector simulated measurements, $x^{\exp}$ is the vector experimental measurements including ^{13}C MIDs and extracellular fluxes, $\sum_x$ is a diagonal matrix of variances in the experimental measurements, $S$is the stoichiometry matrix, and v is the flux vector. We employed the interior-point method to solve the above optimization problem, which is implemented as the default algorithm for the MATLAB built-in function fmincon[]. The solution with the smallest loss function among those 100 candidate flux configurations was used as the final solution of the MFA problem. This approach enabled estimation of intracellular fluxes across key metabolic pathways, including glycolysis, pentose phosphate pathway, tricarboxylic acid (TCA) cycle, and amino acid metabolism. Quality of fitting was evaluated using a chi-square goodness-of-fit test, and 95% confidence intervals for the fluxes were estimated by analyzing the sensitivity of the covariance-weighted RSS to the change in each single flux, thereby providing a statistical measure of uncertainty in the fluxes.

Identification of TDRs

Sampling of thermodynamically feasible metabolic fluxes and Gibbs free energy changes of reactions in GEMs includes two steps. First, we computed possible $\Delta_r G^o$ values of reactions that allow the GEM to be thermodynamically feasible with as little deviation from the $\Delta_r G^o$ values predicted by dGbyG as possible. This was done by solving the optimization problem below under the TFBA framework:

$$\min \sum \frac{\left|\Delta_r G_i^{\circ} - \Delta_r G_{p,i}^{\circ}\right|}{1+\varepsilon_i}, \left\{ i = 1, 2, \cdots, n \middle| \Delta_r G_{p,i}^{\circ} \text{ is known} \right\},$$

where $\Delta_r G_{p,i}^{\circ}$ and ε_i are the value of standard Gibbs energy change and its SD predicted by dGbyG for reaction i, respectively. Solving this optimization problem yields values of $\Delta_r G_i^{\circ}$ that were used in the following steps and we use $\Delta_r G_i^{\circ'}$ to refer to these values below.

Second, we generated random sets of metabolic fluxes and Gibbs free energy changes of reactions that satisfy all flux balance, thermodynamic, and osmotic constraints in the framework of TFBA. Here, we add the constraints below to the original TFBA constraints:

$$\Delta_r G_i^{\circ} = \Delta_r G_i^{\circ'}, \left\{ i = 1, 2, \cdots, n \middle| \Delta_r G_{p,i}^{\circ} \text{ is known} \right\},$$

Flux configurations and corresponding $\Delta_r G$ values of reactions in the GEMs were then generated by optimizing random linear objective functions under the TFBA framework. Briefly, the objective function takes the form below:

$$\max \omega^T v$$

where v is the vector of metabolic fluxes, ω is a vector of linear coefficients randomly sampled from the standard normal distribution $N(0,1)$. This procedure was repeated 200 times using 200 randomly sampled vectors of coefficients, ω, resulting in 200 values of v and $\Delta_r G$ for each reaction. TDRs were defined as reactions with all 200 $\Delta_r G$ values below -200 kJ/mol and at least one non-zero value for the flux.

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

Topological analysis of metabolic networks

MATLAB.mat files implementing the human GEM Recon3D and Human1 were retrieved from the online resources http://vmh.life/ and https://github.com/SysBioChalmers/Human-GEM, respectively. For each model, an undirected graph named reaction connectivity map was then constructed based on the stoichiometric matrix S of the model. In constructing the reaction connectivity map, reactions were treated as nodes, and two reactions were considered to be connected with each other if at least one non-hub metabolite is involved in both reactions. Non-hub metabolites were defined by metabolites that are not ions and molecules that participate in many reactions, such as proton, water, phosphate, NAD/NADH, NADP/NADPH, carbon dioxide, and so on. Degree, closeness and betweenness of each reaction node in the reaction connectivity map were computed using the MATLAB built-in functions degree[] and centrality[]. Distances between reactions in the reaction connectivity map were computed using the MATLAB built-in function distances[]. Exchange reactions for glucose, glutamine, homoserine, isoleucine, leucine, lysine, methionine, phenylalanine, serine, threonine, tryptophan, tyrosine, and valine were defined as nutrient uptake reactions. The random forest model predicting thermodynamic driver reactions from topological features of reactions was built with the degree, betweenness, closeness and distance to nutrient uptake reactions as four independent variables, and the binary label of a reaction being a thermodynamic driver reaction or not as the dependent variable. The model was implemented using the MATLAB built-in function fitensemble[], and its performance was evaluated by 5-fold cross-validation using the MATLAB function predict[]. ROC curves were generated using the MATLAB function perfcurve[]. Visualization of $\Delta_r G^o$ values on metabolic maps was done using the Escher tool.[89]

Analysis of enzyme abundance in human cancer cell lines

Quantitative proteomics data of human cancer cell lines measured by mass spectrometry were retrieved from the non-normalized protein data available at https://gygi.hms.harvard.edu/publications/ccle.html, and proportionally scaled by the sum of intensities of all proteins detected in each cell line. We mapped the gene symbols in the proteomic dataset to Entrez IDs using the online tool SynGO,[94] and computed abundances of enzymes associated with each reaction based on the gene-protein-reaction mapping matrix in Recon3D and Human1. We then classified the enzymes into two categories based on their association with the TDRs: one TDR-group consists of enzymes associated with at least one TDR, while the other NTDR-group contains enzymes that are not associated with any TDR. We performed Z-score normalization of the enzyme abundance matrix so that the distribution of enzyme abundance across cell lines for each reaction has the mean value of zero and the standard deviation of one. Principal component analysis (PCA) was then performed with the Z-score normalized enzyme abundance matrix, and a heterogeneity score for each enzyme was computed by squared sum of the loadings of that enzyme in the top 200 principal components.

Kinetic model of glycolysis

Ordinary differential equations (ODEs)-based model for glycolysis was retrieved from our previous work.[5] For computational simulation of this model, 10,000 random sets of kinetic parameters were generated by Latin hypercube sampling, which were used together with dGbyG-predicted $\Delta_r G^o$ values of reactions in glycolysis to parameterize the model. The flux control coefficient of the glycolytic flux with respect to each reaction was estimated using the finite difference approximation:

$$C^J_{v_i} \approx \frac{\log\left[J\left(1.01 V_{max,i}\right)\right] - \log\left[J\left(V_{max,i}/1.01\right)\right]}{2\log 1.01}$$

In which J is the glycolytic flux and $V_{max,i}$ is the maximal velocity of the enzyme catalyzing the i-th reaction. For each reaction, its Gibbs free energy change was computed from its $\Delta_r G^o$ and concentrations of the substrates and products:

$$\Delta_r G = \Delta_r G^\circ + RT \ln \frac{\prod_i [P_i]^{b_i}}{\prod_i [S_i]^{a_i}}$$

Simulation of the model was done with the ODE solver gsl_odeiv2_step_msbdf in the GNU Scientific Library.

Linear metabolic pathway model

The linear metabolic pathway model consists of five metabolites, Met1 to Met5, and four intracellular reactions, R1 to R4, as illustrated in Figure 7E. The model includes a sequence of flux-balanced unimolecular reactions:

$$v_i = v_j = \mathrm{J}, \forall \mathrm{i}, \mathrm{j} \in \{1, 2, \ldots, \mathrm{r}\},$$

In which r is the number of reactions, v_i is the rate of the i-th reaction, J is the steady-state pathway flux. Within this model, the two reactions at the beginning and end of the pathway are defined as boundary reactions that represent the uptake of the first metabolite and secretion of the last metabolite, while the remaining reactions are defined as intracellular reactions that convert the i-th metabolite to the i+ 1-th metabolite. The model employed in this study includes five different metabolites and six reactions (4 intracellular reactions and 2 boundary reactions).

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

Next, we introduce thermodynamic constraints to the model:

$$\Delta_r G_i = \Delta_r G_i^\circ + RT\ \ln\ (10)(z_{i+1} - z_i),$$
$$\Delta_r G_i^\circ = \Delta_f G_{i+1}^\circ - \Delta_f G_i^\circ,$$
$$z_i = \log_{10}\ (c_i),$$
$$c_l \le c_i \le c_u,$$

where $\Delta_r G$ and $\Delta_r G^o$ are $r_{in} \times$ 1vector of Gibbs energy change and standard Gibbs energy change of reactions, $r_{in} = r - 2$ is the number of intracellular reactions in the model, R is the gas constant, T is the temperature,$m = r - 1$ is the number of metabolites in the model, $\Delta_f G^o$ is a $m \times 1$ vector representing for the standard formation Gibbs energy of metabolites, c is a $m \times 1$ vector of the concentrations of metabolites, and z is the base-10 logarithm of c. The lower and upper bounds for metabolite concentration are set as $c_i = 10^{-9}$ mol/L and metabolite concentration upper limit $c_u = 10^{-1}$ mol/L.

For each reaction, the ratio between the forward and backward fluxes is determined by the Gibbs free energy change of this reaction, $\Delta_r G$, through the flux-force relationship:

$$\Delta_r G_i = -RT\ \ln\left(\frac{v_i^+}{v_i^-}\right),$$
$$v_i = v_i^+ - v_i^-,$$

where v_i^+ is the forward flux and v_i^- is the backward flux for the i-th reaction. Furthermore, the forward flux is determined by the kinetic parameters of the enzyme and the concentrations of the substrate and enzyme participating in this reaction through Michaelis-Menten kinetics:

$$v_i^+ = \frac{K_{cat,i} \cdot e_i \cdot c_i}{K_{m,i} + c_i},$$
$$e_i \ge 0$$

where e_i is the concentration of the i-th enzyme, $K_{cat,i}$ and $K_{m,i}$ are kinetic parameters of that enzyme. Overall, the linear metabolic pathway model contains the following constraints:

$$v_i = v_j = J, \forall i, j \in \{1, 2, \cdots, r\},$$
$$\Delta_r G_i = \Delta_r G_i^\circ + RT\ \ln\ (10)(z_{i+1} - z_i),$$
$$\Delta_r G_i^\circ = \Delta_f G_{i+1}^\circ - \Delta_f G_i^\circ,$$
$$z_i = \log_{10}\ (c_i),$$
$$c_l \le c_i \le c_u,$$
$$\Delta_r G_i = -RT\ \ln\left(\frac{v_i^+}{v_i^-}\right),$$
$$v_i = v_i^+ - v_i^-,$$
$$v_i^+ = \frac{K_{cat,i} \cdot e_i \cdot c_i}{K_{m,i} + c_i},$$
$$e_i \ge 0,$$

For simplicity, we use the literature-based median values[85] for these parameters, $K_{cat} = 10\ \text{s}^{-1}$ and $K_{cat} = 1.3 \times 10^{-4}$ mol/L for all enzymes included in the model.

Constructing Pareto front for the linear metabolic pathway

The multi-objective optimization for maximizing flux while minimizing enzyme and metabolite loads in the linear metabolic pathway is converted to a constrained nonlinear optimization problem to construct the Pareto front. The sum of the concentrations of the intermediate metabolites (Met2, Met3, and Met4) is constrained by the upper limit for total metabolite concentrations, denoted as c_u. Similarly, the total abundance of all enzymes (E1 to E4) must not exceed the upper limit for enzyme concentration, e_u. The Pareto front for this multi-objective optimization problem was constructed by systematically varying the combinations of $\{e_u, c_u\}$, spanning the ranges $e_u \in [10^{-4}, 10^{-2}]$ and $c_u \in [10^{-6}, 10^{-2}]$. Plus, for the two metabolites at the beginning and end of the pathway, Met1 and Met5, their concentrations and $\Delta_f G^o$ values are set as constant values. The reason for doing so is that the concentrations of metabolites at the beginning and end of a metabolic pathway usually depend on the concentration of nutrients and wastes in the environment and the $\Delta_f G^o$ values are determined by the molecular structures of the

Please cite this article in press as: Fan et al., Unraveling principles of thermodynamics for genome-scale metabolic networks using graph neural networks, Cell Systems (2025), https://doi.org/10.1016/j.cels.2025.101393

nutrients and wastes, which cannot be optimized by the cell. Therefore, given the values of these two upper limits, c_u and e_u, a Pareto optimal solution can be calculated by maximizing the pathway flux J as follows:

$$\max J$$

$$\text{s.t.}\left\{\begin{aligned}
v_i &= v_j = J, \forall i,j \in \{1,2,\cdots,r\},\\
\Delta_r G_i &= \Delta_r G_i^\circ + RT\ln(10)(z_{i+1} - z_i),\\
\Delta_r G_i^\circ &= \Delta_f G_{i+1}^\circ - \Delta_f G_i^\circ,\\
z_i &= \log_{10}(c_i),\\
c_l &\le c_i \le c_u,\\
\Delta_r G_i &= -RT\ln\left(\frac{v_i^+}{v_i^-}\right),\\
v_i &= v_i^+ - v_i^-,\\
v_i^+ &= \frac{K_{cat,i}\cdot e_i\cdot c_i}{K_{m,i}+c_i},\\
e_i &\ge 0,\\
\sum_{i=2}^{m-1} c_i &\le c_u,\\
\sum_{i=1}^{r_{in}} e_i &\le e_u,\\
c_1 &= c_{Met1},\\
c_5 &= c_{Met5},\\
\Delta_f G_1^\circ &= \Delta_f G_{Met1}^\circ,\\
\Delta_f G_5^\circ &= \Delta_f G_{Met5}^\circ,
\end{aligned}\right.$$

Here in analysis of Figure 7G, we set $c_{Met1} = 10^{-5}$ mol/L and $c_{Met5} = 10^{-4}$ mol/L as the concentrations of Met1 and Met5, and $\Delta_f G_{Met1}^\circ = 0$ kJ/mol and $\Delta_f G_{Met5}^\circ = -10$ kJ/mol as their standard Gibbs free energy values of formation.