Published in final edited form as:

Biotechnol Bioeng. 2009 June 15; 103(3): 532–541. doi:10.1002/bit.22285.

Thermodynamic Analysis of Biodegradation Pathways

Stacey D. Finley¹, Linda J. Broadbelt¹, and Vassily Hatzimanikatis²

Vassily Hatzimanikatis: vassily.hatzimanikatis@epfl.ch

¹Department of Chemical and Biological Engineering, McCormick School of Engineering and Applied Sciences, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208

²Laboratory of Computational Systems Biotechnology, Ecole Polytechnique Fédérale de Lausanne (EPFL) and Swiss Institute of Bioinformatics (SIB), CH H4 625, Station 6, CH-1015 Lausanne, Switzerland; telephone: +41-21-693-98-70; fax: +41-21-693-98-75

Abstract

Microorganisms provide a wealth of biodegradative potential in the reduction and elimination of xenobiotic compounds in the environment. One useful metric to evaluate potential biodegradation pathways is thermodynamic feasibility. However, experimental data for the thermodynamic properties of xenobiotics is scarce. The present work uses a group contribution method to study the thermodynamic properties of the University of Minnesota Biocatalysis/Biodegradation Database. The Gibbs free energies of formation and reaction are estimated for 914 compounds (81%) and 902 reactions (75%), respectively, in the database. The reactions are classified based on the minimum and maximum Gibbs free energy values, which accounts for uncertainty in the free energy estimates and a feasible concentration range relevant to biodegradation. Using the free energy estimates, the cumulative free energy change of 89 biodegradation pathways (51%) in the database could be estimated. A comparison of the likelihood of the biotransformation rules in the Pathway Prediction System and their thermodynamic feasibility was then carried out. This analysis revealed that when evaluating the feasibility of biodegradation pathways, it is important to consider the thermodynamic topology of the reactions in the context of the complete pathway. Group contribution is shown to be a viable tool for estimating, a priori, the thermodynamic feasibility and the relative likelihood of alternative biodegradation reactions. This work offers a useful tool to a broad range of researchers interested in estimating the feasibility of the reactions in existing or novel biodegradation pathways.

Keywords

complex systems; metabolic engineering; network analysis; synthetic biology; systems biology

Introduction

Increasing amounts of anthropogenic chemicals, called xenobiotics, are released into the environment in many different forms, including industrial pollutants and pesticides (Alexander, 1999). Microorganisms found in nature are able to degrade some of these chemicals (Jain et al., 2005; Paul et al., 2005), but a growing number of anthropogenic compounds are not biodegradable (Pieper and Reineke, 2000). Accumulation of these recalcitrant xenobiotics leads to contamination of air, water, and soil (Chen et al., 2005). Due to the adverse environmental effects of waste treatment methods that are commonly used such as incineration, landfilling, and air-stripping (Dua et al., 2002; Paul et al., 2005), there is a strong effort to develop alternative methods of biodegradation that are efficient, minimally hazardous, and economical (Boopathy, 2000; Sutherland et al., 2004). One such method is bioremediation, which exploits the ability of microorganisms to utilize organic substances for maintenance and growth (Dua et al., 2002; Jain et al., 2005; Watanabe, 2001). Many microorganisms and consortia of microorganisms utilize a variety of pathways for the biodegradation of xenobiotics under different physiological conditions, for example, aerobic or anaerobic conditions.

The information gap between the number of compounds thought to be biodegradable and information on the reactions through which microbial biodegradation occurs can be addressed using computational tools to predict biodegradation. Several tools have been developed to predict the biodegradability of organic compounds and the reactions involved in biodegradation, including BNICE, BDPServer, CATABOL, META, MetabolExpert, METEOR, and the Pathway Prediction System (PPS) (Darvas, 1987; Gomez et al., 2007; Greene et al., 1999; Hatzimanikatis et al., 2005; Hou et al., 2003; Jaworska et al., 2002; Klopman et al., 1994). Once the predicted biodegradation pathway is obtained, it must be characterized to determine its feasibility. A thermodynamic analysis provides an estimate of the energetic feasibility of potential biodegradation pathways.

The thermodynamic properties of a large number of biochemical compounds and reactions have been calculated experimentally under different conditions (Alberty, 1998, 2003; Goldberg et al., 2004; Thauer et al., 1977; Wagman, 1982), but for most of the known biochemical compounds such information is lacking. Group contribution is a tool used to estimate physical properties of organic compounds for which there is little or no experimental data available (Benson, 1976; Mavrovouniotis, 1990, 1991), and there have been several successful demonstrations of the application of these methods. Dolfing and Janssen (1994) used group contribution to classify conversion steps in a biodegradation pathway of halogenated compounds as endergonic or exergonic and determine whether the reactions yield adequate energy to sustain microbial growth. Shelley et al. (1996) used group contribution to evaluate thermodynamic properties of the 2,4,6-trinitrotoluene (TNT) biodegradation pathways. Thermodynamic analysis has also been used to compare biodegradation routes involving reductive dechlorination to those involving oxidative and fermentative degradation reactions (Dolfing, 2000, 2003). A new group contribution method has recently been developed (Jankowski et al., 2008), where experimental data (Alberty, 1998, 2003; Dolfing and Harrison, 1992; Dolfing and Janssen, 1994; Goldberg et al., 2004; Thauer, 1998; Thauer et al., 1977; Wagman, 1982) was utilized to expand the basis set of

groups and refit the group contribution values obtained by Mavrouniotis. This method provides a means of estimating the thermodynamic feasibility of biochemical reactions and has been used to study aromatic amino acid pathways (Hatzimanikatis et al., 2005), glycolysis (Maskow and von Stockar, 2005), and a genome-scale model of *E. coli* (Feist et al., 2007; Henry et al., 2007, 2006).

The present work reports a large-scale thermodynamic analysis of the compounds and biodegradation reactions compiled in the University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) (Ellis and Wackett, 2006), which includes 1,124 compounds, 1,205 reactions, and 170 pathways (June 2008). This analysis includes a study of the estimated Gibbs free energies of formation and reaction, the cumulative free energy of biodegradation pathways, and effect of metabolite activity. Additionally, the thermodynamic feasibility of the biotransformation rules used in the PPS was estimated. The results obtained here provide valuable data to researchers who are studying known biodegradation routes and can be used to evaluate biodegradation routes obtained from prediction methods to determine the thermodynamic feasibility of the proposed reactions.

Methods

Estimating Thermodynamic Properties

The standard Gibbs free energy of reaction, $\Delta_r G^\circ$, is a means of assessing the thermodynamic feasibility of reactions. The reaction energy is related to the equilibrium constant K :

$$\Delta_r G^\circ = -RT \ln K = -RT \ln \left(\prod_{i=1}^m a_i^{n_i} \right) \quad (1)$$

where R is the ideal-gas constant, T is the temperature, m is the number of reactants, and a_i and n_i are the activity and stoichiometric coefficient of compound i , respectively. Standard state is an aqueous solution at $T = 298$ K, where the concentration of each species is 1 M. The concentration of H^+ is set to 10^{-7} M (pH 7) for biological conditions, and the corresponding Gibbs free energy value is denoted by $\Delta_r G'^\circ$.

The Gibbs reaction energy can be calculated from the difference in Gibbs formation energies, $\Delta_f G'^\circ$, of the products and reactants. However, experimental values for the formation energy of most biochemical compounds are largely unavailable (Kummel et al., 2006; Mavrouniotis, 1990). The lack of thermodynamic data is even more pronounced when considering xenobiotics and other compounds relevant to biodegradation. For example, for the 1,124 compounds catalogued in the UM-BBD, experimental data for the formation energy is available for 7% of these compounds. Therefore, the Gibbs free energy of formation, $\Delta_f G'^\circ_{\text{est}}$, is found using group contribution. Similarly, the Gibbs free energy of reaction, $\Delta_r G'^\circ_{\text{est}}$, is estimated by decomposing the reactants and products into groups and using the group changes that occur in the reaction, taking all cofactors into account. The uncertainty in the estimated reaction energy, SE_r , is calculated based on the standard error for each group contribution value (Jankowski et al., 2008):

$$SE_r = \sqrt{\sum_{i=1}^m (n_i SE_{gr,i})^2} \quad (2)$$

where $SE_{gr,i}$ is the standard error for group i and n_i is the number of times group i appears.

Special Considerations for Biodegradation Reactions

Many of the reactions in the UM-BBD (27%) are catalyzed by oxygenases, which incorporate one or two atoms of oxygen into the substrate. These reactions are often the first step in the degradation of organic compounds such as aromatic hydrocarbons (Bugg and Winfield, 1998) and thus are important in biodegradation. Much of the energy released in oxygenase reactions is associated with the reduction of oxygen to water, and since this energy is not coupled to the generation of electron carriers, it is not available to the host organism (Yuan and VanBriessen, 2002). To account for this, the estimated Gibbs free energy change for oxygenase reactions is reduced by the energy associated with the reduction of oxygen, -312.56 kcal/mol (VanBriessen, 2001). This allows the reactions to be compared solely based on the energy available to the organism for cell maintenance and growth.

In microbial biodegradation the xenobiotic is generally the electron donor; however, there are instances, such as anaerobic respiration, when the xenobiotic is the electron acceptor. In these cases, there is a choice of which electron donor to use. The UM-BBD uses electrons to charge balance reactions when the specific electron source is unknown. The contribution value for electrons is zero; however, it is more appropriate to specify the electron source explicitly to obtain free energy estimates, since the specific electron donor will change the free energy value. Therefore, we chose to use a “standard” electron donor for these reactions. Here, acetate was chosen as the electron donor because its degree of reductance and Gibbs free energy of dissipation are close to the regularity values observed by Minkevich and Eroshin (1973). Supplementary Table S2 reports the Gibbs free energy of the reactions for which an electron source was not specified in the UM-BBD, with and without acetate as the electron donor. This approach can be used for any electron donor of interest.

Evaluation of Group Contribution Method

The available experimental data for the Gibbs free energy of formation for compounds catalogued in the UM-BBD (79 compounds, 7% of the database) were used in the data set for the development of the group contribution method (Jankowski et al., 2008). Certain groups were only found in UM-BBD compounds, and thus all of the experimental data available for UM-BBD compounds was used to regress their group contribution values. The estimated formation energy obtained using the group contribution method was compared to the literature value for all 79 compounds, and the deviation between the estimated and literature value was used to assess how well the group contribution method fit the experimental data. $\Delta_r G'_{est}$ for 77% of the compounds is within the range of one standard deviation (7.9 kJ/mol) of the experimentally observed value; 95% of estimated formation energies are within two standard deviations. This is consistent with the confidence interval calculated for all of the experimental data used in the development of the group contribution

method, where 85% and 96% of the data fall within the range of one standard deviation and two standard deviations, respectively, indicating that the estimated values are in good agreement with the available experimental data. A parity plot of the estimated and experimental Gibbs formation energy values is shown in Supplementary Figure S1; the experimental data are also given in Supplementary Table S1. There are three UM-BBD compounds for which the estimated free energy of formation is not within two standard deviations. These compounds do not share a common group, and there does not appear to be any trends that explain why the deviation is larger than expected. The same holds true for the compounds and reactions with large errors in the complete data set used to regress the contribution values (Jankowski et al., 2008).

Jankowski and coworkers performed an extensive cross validation analysis of the group contribution method to determine the ability to estimate thermodynamic properties of compounds and reactions not used in the training data set. Additionally, we identified another set of experimental data that can be used as a separate measure of the validity of the method. Holmes et al. (1993) estimated the aqueous Gibbs free energies of formation for polychlorinated biphenyls (PCBs) using experimental data for the standard state enthalpy of formation of the gas. They found the formation energy of biphenyl and 4-chlorobiphenyl to be 275.2 and 250.0 kcal/mol, respectively. We estimated these values to be 281.8 and 243.6 kcal/mol; both estimates are within the standard error, further validating the estimation method.

Results and Discussion

Distribution of the Estimated Gibbs Free Energy of Formation, $\Delta_f G'_{\text{est}}$

The Gibbs free energy of formation, $\Delta_f G'_{\text{est}}$, was estimated for 914 (81%) compounds in the database (Supplementary Fig. S2 and Supplementary Table S1). $\Delta_f G'_{\text{est}}$ of 812 (72%) compounds ranges from -1.3×10^3 to 525 kJ/mol, and $\Delta_f G'_{\text{est}}$ of 102 compounds is more negative than -1.9×10^3 kJ/mol. Ninety-four of these low-energy compounds contain a CoA moiety, and their large negative formation energy results from the contribution of CoA (Supplementary Fig. S3). One low-energy compound is hydrogen (1*R*, 2*R*)-1-glutathio-2-hydroxypropylphosphonic acid, while the other five low-energy compounds participate in the methanogenesis pathway (Donnelly and Wolfe, 1986). These large molecules are composed of many groups, including carboxyl and phosphate groups, ketones, and secondary alcohols. The combination of these groups in particular results in the large negative free energy value. Note that compounds with a large negative or positive free energy estimate do not necessarily have large uncertainty. Rather, the uncertainty in the free energy is related to the number of groups that comprise the compound and their individual uncertainties, as shown in Equation (2).

The formation energy was not estimated for 210 (19%) compounds. These compounds were not fully decomposed into the basis set of groups used to develop the method. Among these compounds, 54 contain the $-\text{NO}_2$ group (5%), which has an unknown group contribution value. In addition to these compounds, the group contribution method is unable to estimate the thermodynamic properties of compounds containing structural elements listed in Table I.

Given the limited experimental data available, it is not possible to determine their contribution value (Jankowski et al., 2008). However, this summary of the missing groups offers salient targets for experimentalists who are interested in measuring free energies of formation, since experimental values of a small number of compounds containing these groups would have a profound impact on our ability to estimate thermodynamic properties for a broader range of species.

Distribution of the Estimated Gibbs Free Energy of Reaction, $\Delta_r G'_{\text{est}}^{\circ}$

Using the group contribution method, the Gibbs free energy of reaction was estimated for 902 reactions (75%) (Supplementary Table S2). A sample calculation for the phenylacetate hydroxylase reaction is given in the Supplementary Materials (Supplementary Fig. S4).

Most of the biodegradation reactions are thermodynamically favorable; 78% of $\Delta_r G'_{\text{est}}^{\circ}$ values are ≤ 0 (Fig. 1). Because the database contains only observed biodegradation reactions, the reactions estimated to be unfavorable were examined more closely in subsequent analyses, and we found that incorporating a range of species' concentrations relevant to biodegradation results in a range of feasible Gibbs free energy values for nearly all of these reactions. Examining the distribution of $\Delta_r G'_{\text{est}}^{\circ}$ according to the Enzyme Classification (EC) provides an idea of the types of enzymes employed in biodegradation and the energetics of these enzyme-catalyzed reactions. This analysis shows that oxidoreductases catalyze 55% of the reactions for which the free energy can be estimated (Supplementary Fig. S5). Of these, 52% are oxygenase reactions. Additionally, hydrolases and lyases account for 25% of the reactions with an estimated free energy value.

Based on the current state of the group contribution method, the free energy of 25% of the reactions in the database could not be estimated, which can be separated into three groups (Table II). The first group of reactions involves compounds that are not fully decomposed into the basis set of groups. Here, groups with unknown contribution values are created or destroyed. The second group includes reactions that are missing intermediate steps and are not atom balanced. This includes reactions where the proposed biodegradation pathway shows that the substrate is degraded to a compound such as carbon dioxide, but the intermediate steps are not known. When estimating the free energy of reaction, the reaction must be atom balanced such that all group changes that occur from substrates to products are accounted for. The third group of reactions is missing structural information for cofactors such as rubredoxin or AQDSH₂ or an unspecified R-group or acceptor.

The residual substructures needed to fully decompose all of the compounds and reactions in the database include groups containing elements such as arsenic, bromine, silicon, and tin, and structural features such as multiple fused rings. However, the subset of structural groups needed to estimate the thermodynamic properties of the compounds and reactions in the database that do not presently have an estimate is not expected to significantly impact the other group values used in the method.

Cumulative $\Delta_r G'_{\text{est}}$ of Biodegradation Pathways

Using the reactions for which the free energy could be estimated, we investigated the energetic feasibility of biodegradation pathways of xenobiotic compounds reported in the UM-BBD. In this analysis, the directionality of the reactions was defined based on the flow of degradation and the length of the pathway as the number of reactions in the longest branch. The cumulative Gibbs free energy could be calculated for 89 (51%) pathways of varying length. The free energy landscape of these pathways varied widely depending on the substrate (Supplementary Fig. S6). This analysis can be illustrated using the 4-chlorobiphenyl (cbp) pathway (Fig. 2). 4-Chlorobiphenyl has been used as a model substrate to examine aerobic biodegradation of PCBs (Ellis and Wackett, 2006), a class of synthetic compounds composed of biphenyl molecules with 1–10 chlorine atoms (Erickson and Mondello, 1993). The pathway in the UM-BBD shows the degradation of cbp to pyruvate and acetaldehyde. Additionally, this pathway illustrates the interconnectedness of the pathways within the UM-BBD; 4-chlorobenzoate, released in the fourth step, enters the 2,4-dichlorobenzoate (dcb) pathway where it is metabolized to 4-hydroxybenzoate. This compound is subsequently transformed into 3-carboxy-*cis,cis*-muconate in the vanillin (van) pathway. Alternatively, 4-hydroxybenzoate can be further metabolized to acetaldehyde and pyruvate. Thus the endpoints of cbp biodegradation are acetaldehyde, pyruvate, and 3-carboxy-*cis,cis*-muconate. These compounds have known intermediary metabolism and are further degraded to carbon dioxide, hydrochloric acid, and water, the endpoints of complete mineralization of cbp.

Estimating the thermodynamic landscape has several applications in biodegradation. Since many of the pathways in the UM-BBD include branching or competing pathways, the free energy landscape can be used to compare the possible routes. Additionally, this work is able to identify the specific reactions that account for the wide range of cumulative free energy values. Lastly, thermodynamics can be used to determine the reactions that are regulated by the cell, where it has been proposed that enzymes catalyzing reactions with large Gibbs free energy values are likely to be tightly regulated because of their influence on flux control (Kummel et al., 2006).

Effect of Metabolite Activity

Substrate concentration is known to affect the biodegradation of some chemicals and xenobiotics (Alexander, 1985; Boethling and Alexander, 1979; Kovar et al., 2002; Rapp and Timmis, 1999). Given that there is a threshold concentration below which these compounds are not metabolized, it is useful to estimate the range of possible free energy values over a range of species' concentration. Additionally, it may be useful to assess the thermodynamic feasibility even when the exact concentration of each reactant and product cannot be measured and as the substrate concentration changes during the course of biodegradation.

The minimum and maximum values for the estimated free energy change, $\Delta_r G'_{\text{est},\text{min}}$ and $\Delta_r G'_{\text{est},\text{max}}$, are calculated using the equations described by Henry et al. (2006). These equations incorporate metabolite activity, as well as the uncertainty in the reaction energy within two standard deviations:

$$\Delta_r G'_{\text{est,max}} = \sum_{i=1}^n n_i \Delta_f G'^{\circ}_{\text{est}} + RT \sum_{i=1}^{\text{products}} n_i \ln(x_{\text{max}}) + RT \sum_{i=1}^{\text{reactants}} n_i \ln(x_{\text{min}}) + U_{r,\text{est}} \quad (3)$$

$$\Delta_r G'_{\text{est,min}} = \sum_{i=1}^m n_i \Delta_f G'^{\circ}_{\text{est}} + RT \sum_{i=1}^{\text{products}} n_i \ln(x_{\text{min}}) + RT \sum_{i=1}^{\text{reactants}} n_i \ln(x_{\text{max}}) - U_{r,\text{est}} \quad (4)$$

where $U_{r,\text{est}}$ is the uncertainty, x_{min} is the minimum metabolite activity of 0.1 mM, the threshold concentration (Alexander, 1999), and x_{max} is the maximum metabolite activity of 1 M. The maximum free energy value occurs when the activity of the products is at the maximum value and that of the reactants is at the minimum value. Conversely, the estimated minimum free energy is evaluated when the activity of the products and reactants is at the minimum and maximum values, respectively.

Based on the calculation of the minimum and maximum free energy change, the reactions in the UM-BBD were separated into three distinct classes: (1) favorable, (2) conditionally favorable, and (3) unfavorable (Fig. 3). The first class of reactions has $\Delta_r G'_{\text{est,max}} \leq 0$ and includes 346 (25%) reactions. Under the range of metabolite activities investigated, these reactions are favorable in the direction written. The second classification has $\Delta_r G'_{\text{est,min}} < 0$ and includes 543 (49%) reactions that are conditionally favorable, where the directionality of these reactions depends on the species' concentrations. An example of this type of reaction is NADP-dependent alcohol dehydrogenase:

where $\Delta_r G'^{\circ}_{\text{est}}$ is 24.2 kJ/mol. The Gibbs free energy of this reaction can range from -55.8 to 104.2 kJ/mol. When $\Delta_r G'_{\text{est,min}}$ was used as a measure of feasibility, 187 reactions that were unfavorable based on $\Delta_r G'^{\circ}_{\text{est}}$ became thermodynamically favorable.

The third class of reactions has $\Delta_r G'_{\text{est,min}} > 0$, including 13 (1%) reactions. Given the range of metabolite activities studied here, these reactions are thermodynamically unfavorable as written. These reactions are of particular interest because although they are estimated to be thermodynamically unfavorable, they are still observed experimentally. We have reviewed the conditions under which these reactions were observed to rectify this contradiction. The hydroxychromene-carboxylate isomerase enzyme is utilized in the degradation of many polycyclic aromatic hydrocarbons, including methylquinoline, naphthalene, naphthoic acid, and phenanthrene. In the elucidation of the biodegradation pathway for naphthalene, the optimum pH for this enzyme was found to be 10 (Eaton and Chapman, 1992), where the reaction occurred entirely in the forward direction. At this pH, the predominant ion of the product of the isomerase reaction changes such that the reaction becomes thermodynamically favorable (the predominant ion was estimated using MarvinBeans pK_a estimation software, v. 4.1.1, Budapest, Hungary). Since pH 7 is more likely to occur in the environment, it is reasonable to expect that this reaction will proceed in the reverse

direction. This explains the discrepancy between the experimentally observed transformation and the thermodynamic estimate for several reactions (r0337, r0502, r0765, r0782, and r0815). In the case of three reactions (r0051, r0872, and r1104), the product has been hypothesized but not explicitly identified (Fournier et al., 2002; Schach et al., 1993; van Herwijnen et al., 2003). In our examination of the literature, we were unable to identify the reason for the inconsistency between experimental observations and thermodynamic estimates for the remaining five reactions.

Based on the minimum and maximum free energy estimates, nearly all (98%) of the reactions for which the free energy could be estimated are classified as thermodynamically favorable. These results are not surprising given the types of compounds included in the database—compounds known to be degraded in the environment via biodegradation reactions and pathways that have been observed experimentally. Thus, the UM-BBD is inherently biased. However, we propose that it is still attractive to use estimation methods to calculate the free energy of biodegradation reactions, particularly in the prediction of biodegradation pathways. Thermodynamics provides a way to compare possible pathways for a particular compound in order to determine which pathway generates the most energy for the organism.

Application to the Pathway Prediction System

To explore the connection between thermodynamics and prediction of attractive pathways, the estimated thermodynamic properties were applied to the PPS. The PPS is a user-guided system that uses metabolic rules involving chemical functional groups to predict the degradation of xenobiotics, where a rule is created if there is an example of the metabolism in the UM-BBD, or if it is known to occur in the environment. The rules consist of two parts: (1) the functional group undergoing biotransformation is identified and (2) the biotransformation occurs, transforming the molecule containing the functional group into a product using rule matching (Hou et al., 2003). More recently, the rules were prioritized by biodegradation experts who ranked them according to the likelihood with which the transformation would occur under standard aerobic conditions (Ellis and Wackett, 2006).

Here, the group contribution method was used to estimate the thermodynamic feasibility of the PPS biotransformation rules. There are 235 biotransformation rules and 903 of the reactions catalogued in the UM-BBD have been assigned to one or more rules. The thermodynamic properties of 747 of these reactions (83%) can be estimated using the group contribution method. Some of these reactions (34) are associated with more than one rule. Therefore, we have estimated the Gibbs free energy change of 781 biotransformations. Comparison of the thermodynamic feasibility with the PPS ranking of the rules shows there is no correlation between the likelihood that the biotransformation would occur as ranked by experts and the estimated thermodynamic feasibility of the reaction (Fig. 4). This is consistent with the fact that many factors influence the biodegradation of xenobiotic compounds, including environmental conditions, substrate characteristics and concentration (Boopathy, 2000). Thus, the absolute value of the Gibbs free energy of reaction alone cannot be used as the only measure of the likelihood of a reaction for a diverse set of xenobiotics.

For a given xenobiotic, however, thermodynamics may provide guidance about which pathway is more likely when alternatives exist. To explore this, the thermodynamic estimates were used to compare possible biodegradation pathways for a given xenobiotic. The PPS was used to predict the biodegradation pathways of three sample compounds, 3-chloroacrylic acid, 4-chlorobenzoate, and 2-butenenitrile, a compound found in the CAS Registry, but not catalogued in the UM-BBD. These are representative compounds for various classes of xenobiotic compounds, including chlorinated aliphatics, chlorinated aromatics, and nitrogen-containing compounds. The pathway was cutoff once a compound was formed for which the intermediary metabolism is known, where the compound is readily degraded, as indicated by the PPS (Fig. 5). We have compared predicted biodegradation routes of the same length to determine the relationship between the feasibility of each reaction step and the cumulative free energy of the pathway. Figure 6a compares various six-step pathways of 3-chloroacrylic acid, each leading to different endpoints. At the third reaction step, there is a choice of three alternative reactions. Considering the thermodynamic properties of these reactions, the obvious choice would be reaction 3c; however, this reaction is the least likely to occur amongst the three choices based on PPS rankings. If instead, we consider the cumulative free energy of the pathway, it is more thermodynamically favorable to proceed with reaction 3b, which is also the reaction that is most likely to occur. Similar results were found for three- and four-reaction pathways for 3-chloroacrylic acid (results not shown), as well as pathways for 4-chlorobenzoate, and 2-butenenitrile (Fig. 6b and c). This analysis reveals that when given the choice between alternate reactions, it is not always in the best interest of the organism (energetically) to choose the most thermodynamically favorable reaction. Rather, it is important to consider the thermodynamic topology of the reactions in the context of the entire biodegradation pathway. This is analogous to performing a depth-first search of the possible biodegradation routes rather than a breadth-first search. That is, given a starting reaction, the most energetically favorable pathway would be found by traversing the network of biodegradation reactions as far as possible through each pathway.

Conclusions

Group contribution has been shown to be a viable tool to provide a priori estimates of the thermodynamic properties of known and novel biodegradation reactions. This is not an estimate of the rate or extent of biodegradation, but rather the energetics of such reactions. The application of the group contribution method to estimate the energetic feasibility of reactions has been demonstrated on existing biodegradation reactions and pathways. This analysis can be extended to characterize the feasibility of predicted biodegradation pathways involving enzymes acting on new substrates. The thermodynamic framework alone is not able to predict biodegradation; however, when combined with reaction pathway prediction tools, it provides estimates of the energetic feasibility of alternative reactions, thereby guiding pathway prediction. Thermodynamic analysis can also guide metabolic engineering efforts by identifying unfavorable reactions that should not be implemented. It must be noted, however, that thermodynamics alone does not mean a particular biodegradation pathway should be implemented into a host organism. Additional information is required to determine the cellular feasibility of the reactions and pathways, that is, how implementation

of these reactions into a host organism will influence cell growth. Tools such as metabolic flux analysis (MFA) indicate how implementation of the novel pathway will influence the existing metabolic network of the organism. The thermodynamic estimates can be coupled to MFA to generate thermodynamically feasible flux profiles in order to study the feasibility of pathways under physiological conditions (Henry et al., 2007).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Contract grant sponsor: US Department of Energy

Contract grant sponsor: DuPont

Contract grant sponsor: Swiss Institute of Bioinformatics

This work is supported by the US Department of Energy, Genomes to Life Program. V.H. is supported by the Swiss Institute of Technology (EPFL) and the Swiss National Science Foundation. Additionally, S.D.F. received fellowships from the NIH Biotechnology Training Grant and NSF Graduate Research Fellowship Program. The authors thank Dr. Lynda Ellis and Dr. Lawrence Wackett for making the contents of the UM-BBD readily accessible. The authors also thank Dr. Christopher Henry for his input and making group contribution values available prior to publication and an anonymous reviewer whose comments substantially shaped this manuscript.

References

- Alberty RA. Calculation of standard transformed formation properties of biochemical reactants and standard apparent reduction potentials of half reactions. *Arch Biochem Biophys.* 1998; 358(1):25–39. [PubMed: 9750161]
- Alberty, RA. Thermodynamics of biochemical reactions. Hoboken: John Wiley & Sons; 2003.
- Alexander M. Biodegradation of organic chemicals. *Environ Sci Technol.* 1985; 18(2):106–111.
- Alexander, M. Biodegradation and bioremediation. San Diego: Academic Press, Inc; 1999. p. 302
- Benson, SW. Thermochemical kinetics: Methods for the estimation of thermochemical data and rate parameters. New York: John Wiley & Sons; 1976.
- Boethling RS, Alexander M. Effect of concentration of organic chemicals on their biodegradation by natural microbial communities. *Appl Environ Microbiol.* 1979; 37(6):1211–1216. [PubMed: 16345402]
- Boopathy R. Factors limiting bioremediation technologies. *Bioresour Technol.* 2000; 74(1):63–67.
- Bugg TDH, Winfield CJ. Enzymatic cleavage of aromatic rings: Mechanistic aspects of catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Nat Prod Rep.* 1998; 5:513–530.
- Chen W, Mulchandani A, Deshusses MA. Environmental biotechnology: Challenges and opportunities for chemical engineering. *AIChE J.* 2005; 51(3):690–695.
- Darvas, F. MetabolExpert, an expert system for predicting metabolism of substances. In: Kaiser, K., editor. QSAR in environmental toxicology. Riedel; Dordrecht: 1987. p. 71-81.
- Dolfing J. Energetics of anaerobic degradation pathways of chlorinated aliphatic compounds. *Microb Ecol.* 2000; 40:2–7. [PubMed: 10977871]
- Dolfing, J. Thermodynamic considerations for dehalogenation. In: Haggblom, MM., editor. Dehalogenation: Microbial processes and environmental applications. Secaucus, NJ: Kluwer Academic Publishers; 2003. p. 89-114.
- Dolfing J, Harrison BK. Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. *Environ Sci Technol.* 1992; 26:2213–2218.

- Dolfing J, Janssen DB. Estimates of Gibbs free energies of formation of chlorinated aliphatic compounds. *Biodegradation*. 1994; 5:21–28.
- Donnelly MI, Wolfe RS. The role of formylmethanofuran: Tetrahydromethanopterin formyltransferase in methanogenesis from carbon dioxide. *J Biol Chem*. 1986; 261(35):16653–16659. [PubMed: 3097011]
- Dua M, Singh A, Sethunathan N, Johri AK. Biotechnology and bioremediation: Successes and limitations. *Appl Microbiol Biotechnol*. 2002; 59:143–152. [PubMed: 12111139]
- Eaton RW, Chapman PJ. Bacterial metabolism of naphthalene: Construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions. *J Bacteriol*. 1992; 174(23):7542–7554. [PubMed: 1447127]
- Ellis LB, Wackett LP. The University of Minnesota Biocatalysis/Biodegradation Database: The first decade. *Nucleic Acids Res*. 2006; 34:D517–D521. [PubMed: 16381924]
- Erickson BD, Mondello FJ. Enhanced biodegradation of polychlorinated biphenyls after site-directed mutagenesis of a biphenyl dioxygenase gene. *Appl Environ Microbiol*. 1993; 59:3858–3862. [PubMed: 8285689]
- Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol*. 2007; 3:1–18.
- Fournier D, Halasz A, Spain J, Fiurasek P, Hawari J. Determination of key metabolites during biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine with *Rhodococcus* sp. strain DN22. *Appl Environ Microbiol*. 2002; 68(1):166–172. [PubMed: 11772623]
- Goldberg RN, Tewari YB, Bhat TN. Thermodynamics of enzyme-catalyzed reactions—A database for quantitative biochemistry. *Bioinformatics*. 2004; 20(16):2874–2877. [PubMed: 15145806]
- Gomez MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A. The environmental fate of organic pollutants through the global microbial metabolism. *Mol Syst Biol*. 2007; 3:1–11.
- Greene N, Judson PN, Langowski JJ, Marchant CA. Knowledge-based expert systems for toxicity and metabolism prediction: DEREK, StAR and METEOR. *SAR QSAR Environ Res*. 1999; 10:299–314. [PubMed: 10491855]
- Hatzimanikatis V, Li C, Ionita JA, Henry CS, Jankowski MD, Broadbelt LJ. Exploring the diversity of complex metabolic networks. *Bioinformatics*. 2005; 21(8):1603–1609. [PubMed: 15613400]
- Henry CS, Jankowski MD, Broadbelt LJ, Hatzimanikatis V. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys J*. 2006; 90:1453–1461. [PubMed: 16299075]
- Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-based metabolic flux analysis. *Biophys J*. 2007; 92:1792–1805. [PubMed: 17172310]
- Holmes DA, Harrison BK, Dolfing J. Estimation of Gibbs free energies of formation for polychlorinated biphenyls. *Environ Sci Technol*. 1993; 27:725–731.
- Hou BK, Wackett LP, Ellis LB. Microbial pathway prediction: A functional group approach. *J Chem Inf Comput Sci*. 2003; 43(3):1051–1057. [PubMed: 12767164]
- Jain RK, Kapur M, Labana S, Lal B, Sharma PM, Bhattacharya D, Thakur IS. Microbial diversity: Application of microorganisms for the biodegradation of xenobiotics. *Curr Sci*. 2005; 89:101–112.
- Jankowski MD, Henry CS, Broadbelt LJ, Hatzimanikatis V. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys J*. 2008; 95:1487–1499. [PubMed: 18645197]
- Jaworska J, Dimitrov S, Nikolova N, Mekenyan O. Probabilistic assessment of biodegradability based on metabolic pathways: Catabol system. *SAR QSAR Environ Res*. 2002; 13(2):307–323. [PubMed: 12071658]
- Klopman G, Dimayuga M, Talafous J. META. 1. A program for the evaluation of metabolic transformation of chemicals. *J Chem Inf Comput Sci*. 1994; 34(6):1320–1325. [PubMed: 7989397]
- Kovar K, Chaloupka V, Egli T. A threshold substrate concentration is required to initiate the degradation of 3-phenylpropionic acid in *Escherichia coli*. *Acta Biotechnol*. 2002; 22:285–298.
- Kummel A, Panke S, Heinemann M. Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol Syst Biol*. 2006; 2:2006. 0034. [PubMed: 16788595]

- Maskow T, von Stockar U. How reliable are thermodynamic feasibility statements of biochemical pathways? *Biotechnol Bioeng.* 2005; 92(2):223–2230. [PubMed: 15962336]
- Mavrovouniotis ML. Group contributions for estimating standard Gibbs energies of formation of biochemical compounds in aqueous solution. *Biotechnol Bioeng.* 1990; 36:1070–1082. [PubMed: 18595046]
- Mavrovouniotis ML. Estimation of standard Gibbs energy changes of biotransformations. *J Biol Chem.* 1991; 266:14440–14445. [PubMed: 1860851]
- Minkevich IG, Eroshin VK. Productivity and heat generation of fermentation under oxygen limitation. *Folia Microbiol.* 1973; 18:376–385. [PubMed: 4202199]
- Paul D, Pandey G, Pandey J, Jain RK. Accessing microbial diversity for bioremediation and environmental restoration. *Trends Biotechnol.* 2005; 23(3):135–142. [PubMed: 15734556]
- Pieper DH, Reineke W. Engineering bacteria for bioremediation. *Curr Opin Biotechnol.* 2000; 11:262–270. [PubMed: 10851148]
- Rapp P, Timmis KN. Degradation of chlorobenzenes at nanomolar concentrations by *Burkholderia* sp. strain PS14 in liquid cultures and in soil. *Appl Environ Microbiol.* 1999; 65(6):2547–2552. [PubMed: 10347041]
- Schach S, Schwarz G, Fetzner S, Lingens F. Microbial metabolism of quinoline and related compounds. XVII. Degradation of 3-methylquinoline by *Comamonas testosteroni* 63. *Biol Chem Hoppe Seyler.* 1993; 374(3):175–181. [PubMed: 8489738]
- Shelley MD, Autenrieth RL, Wild JR, Dale BE. Thermodynamic analysis of trinitrotoluene biodegradation and mineralization pathways. *Biotechnol Bioeng.* 1996; 50:198–205. [PubMed: 18624329]
- Sutherland TD, Horne I, Weir KM, Coppin CW, Williams MR, Selleck M, Russell RJ, Oakeshott JG. Enzymatic bioremediation: From enzyme discovery to applications. *Clin Exp Pharmacol Physiol.* 2004; 31:817–821. [PubMed: 15566400]
- Thauer RK. Biochemistry of methanogenesis: A tribute to Marjory Stephenson. *Microbiology.* 1998; 144:2377–2406. [PubMed: 9782487]
- Thauer RK, Jungermann K, Decker K. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev.* 1977; 41:100–180. [PubMed: 860983]
- van Herwijnen R, Springael D, Slot P, Govers HAJ, Parsons JR. Degradation of anthracene by *Mycobacterium* sp. strain LB501T proceeds via a novel pathway, through o-phthalic acid. *Appl Environ Microbiol.* 2003; 69(1):186–190. [PubMed: 12513994]
- VanBriessen JM. Thermodynamic yield predictions for biodegradation through oxygenase activation reactions. *Biodegradation.* 2001; 12:265–281. [PubMed: 11826909]
- Wagman, DD. NBS tables of chemical thermodynamic properties: Selected values for inorganic and C1 and C2 organic substances in SI units. New York; Washington, DC: American Chemical Society and American Institute of Physics; 1982.
- Watanabe K. Microorganisms relevant to bioremediation. *Curr Opin Biotechnol.* 2001; 12:237–241. [PubMed: 11404100]
- Yuan Z, VanBriessen JM. Yield prediction and stoichiometry of multi-step biodegradation reactions involving oxygenation. *Biotechnol Bioeng.* 2002; 80(1):100–113. [PubMed: 12209791]

Figure 1. Distribution of the Gibbs free energy change for the reactions in the UM-BBD. Gray bars indicate the number of reactions with a given $\Delta_r G'_{\text{est}}$; the black line denotes the cumulative percentage.

Figure 2.

Free energy landscape for the 4-chlorobiphenyl pathway. The cumulative free energy change of the pathway is given at each reaction step. The biodegradation pathway is shown: (1) biphenyl dioxygenase, (2) *cis*-2,3-dihydro-2,3-dihydroxybiphenyl dehydrogenase, (3) 2,3-dihydroxybiphenyl-1,2-dioxygenase, (4) 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase, (5) 2-hydroxypenta-2,4-dienoate hydratase, and (6) 4-hydroxy-2-oxovalerate aldolase. Dashed line indicates the intermediary metabolism is known.

Figure 3.

Classification of reactions based on $\Delta_r G'_{\text{est}}$. The black dots specify $\Delta_r G'_{\text{est}}$. Red error bars indicate the range of $\Delta_r G'_{\text{est}}$ values taking into account metabolite activity ranging from 0.1 mM to 1 M and uncertainty. Thermodynamically favorable reactions have maximum $\Delta_r G'_{\text{est}} \leq 0$; conditionally favorable reactions have minimum $\Delta_r G'_{\text{est}} < 0$; unfavorable reactions have minimum $\Delta_r G'_{\text{est}} > 0$.

Figure 4.

Comparison of the thermodynamic feasibility and the PPS ranking for biotransformation rules. The thermodynamic feasibility is based on the minimum and maximum free energy change of the reaction. The color bar indicates the PPS ranking, where each rule is prioritized according to the likelihood of occurring under standard aerobic conditions.

Figure 5.

Alternate biodegradation routes generated by the PPS. Three sample compounds were studied: **a**: 3-chloroacrylic acid, **b**: 4-chlorobenzoate, and **c**: 2-butenitrile. For each compound, only a subset of the pathways is shown. The endpoints of each pathway have known intermediary metabolism.

● Very likely ● Likely ● Neutral ● Unlikely ● Very unlikely

Figure 6.

Thermodynamic feasibility of alternate biodegradation routes. **a:** 3-chloroacrylic acid, **b:** 4-chlorobenzoate, and **c:** 2-butenitrile. The color bar indicates the PPS ranking and the cumulative free energy is given for each pathway shown. Alternative pathways are denoted by different symbols ($\circ$, $\square$, or $\diamond$) and the reaction numbers correspond with the biodegradation pathways shown in Figure 5.

Table I

Structural elements with no group contribution value.

Structural feature	Example compound	% of compounds
Carbon involved in three fused rings	Pyrene	5
NO ₂	Pentaerythritoltetranitrate	5
Chemical elements (As, B, Cr, Hg, Os, Se, Si, Sn, Tc, U)	Arsonacetate	4
Sulfinyl and sulfonyl groups	Dibenzothiophene-5-oxide	2
Halogens not attached to a benzene ring (excluding Cl)	Bromomaleylacetate	1
Miscellaneous (S=P, NO, CN)	N-Cyclohexylisocyanide	1
Small molecules	Perchlorate	<1

Table II

Types of reactions with no free energy estimate.

No. of reactions	Reason for no estimate	Example
229	Involve compounds that are not fully decomposed	<chem>Nc1cc(Br)c(O)c(Br)c1=O.O>>[O-]C(=O)c1cc(Br)c(O)c(Br)c1=O.[NH4+]</chem>
40	Not atom-balanced: missing intermediate reactions	<chem>O=C(CCc1ccccc1O)C(=O)[O-]>>O=C(O)c1ccccc1O</chem>
34	Missing structural information	<chem>Oc1ccccc1.O=P([O-])([O-])X>>Oc1ccccc1OP([O-])([O-])O.X</chem>