PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

PCCP

RSCPublishing

ARTICLE

Cite this: DOI: 10.1039/x0xx00000x
Received 00th January 2012,
Accepted 00th January 2012
DOI: 10.1039/x0xx00000x
www.rsc.org/

Molecular Thermodynamics of Metabolism: Quantum Thermochemical Calculations for Key Metabolites

N. Hadadi[a,b] , M. Ataman[a,b], V. Hatzimanikatis[a,b*], C. Panayiotou[c*]

[a]Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland
[b] Swiss Institute of Bioinformatics (SIB), CH-1015 Switzerland
[c]On leave of absence from the Department of Chemical Engineering, University of Thessaloniki, Greece

The present work is the first of a series of papers aiming at a coherent and unified development of the thermodynamics of metabolism and the rationalization of feasibility analysis of metabolic pathways.
The focus in this part is on high-level quantum chemical calculations of the thermochemical quantities of relatively heavy metabolites such as aminoacids/oligopeptides, nucleosides, saccharides and their derivatives in the ideal gas state. The results of this study will be combined with the corresponding hydration/solvation results in subsequent parts of this work in order to derive the desired thermochemical quantities in aqueous solutions. The above metabolites exist in a vast conformational/isomerization space including rotational conformers, tautomers or anomers exhibiting often multiple or cooperative intramolecular hydrogen bonding. We examine the challenges posed by these features for the reliable estimation of thermochemical quantities. We discuss conformer search, conformer distribution and averaging processes. We further consider neutral metabolites as well as protonated and deprotonated metabolites. In addition to the traditional presentation of gas-phase acidities, basicities and proton affinities, we also examine heats and free energies of ionic species. We obtain simple linear relations between the thermochemical quantities of ions and the formation quantities of their neutral counterparts. Furthermore, we compare our calculations with reliable experimental measurements and predictive calculations from the literature, when available. Finally, we discuss the next steps and perspectives for this work.

Keywords: Quantum thermochemistry; Heats of formation; Intramolecular hydrogen-bonding; Acidity; Basicity

* Corresponding authors:
cpanayio@cheng.auth.gr
vassily.hatzimanikatis@epfl.ch

Physical Chemistry Chemical Physics

I. Introduction

The energetics studies of metabolic pathways require a sufficiently accurate knowledge of the thermochemical quantities of involved substrates and metabolites[1]. Thermodynamics has been applied in the context of thermodynamic constraint-based analysis of genome-scale models of microorganisms for quantifying the feasible ranges for the Gibbs free energy change of a reaction[2-4] and for the systematic assessment of the degree of reversibility of metabolic reactions[5]. Numerous methods based on thermodynamic constraints and the laws of thermodynamics have also been applied in the study of the regulatory network of the cell[6-8]. Such studies require the knowledge of the standard Gibbs free energy change of the involved reaction in order to quantify the degree of thermodynamic favourability of the reactions in biological systems. The number of molecules involved in the biochemical reactions is huge and the currently available Gibbs free energies of formation covers only a small fraction of them[9, 10]. Thus, methods for the efficient use of this minimal information and the development of reliable predictive schemes of the above thermochemical quantities are highly desirable. In response to this need, our laboratory has designed a Group-Contribution method (GCM), which is a particularly useful tool for metabolic pathway analysis. Using this method, the standard Gibbs free energy of formation (and thus, the corresponding free energy change of their reactions) can be calculated with sufficient accuracy for a large percentage of dilute aqueous solutions[11].

GCMs are indispensable tools for the high throughput feasibility analysis of metabolites and pathways. However, since they are empirical methods, they have their limitations. The coverage and the quality of functional group contributions heavily depend on the existence of relevant experimental data and on the quality of the data used for their determination. Thus, the development of alternative non-empirical approaches, which could substitute experiment, is also very much anticipated not only for the practical calculations but also for the insight they provide. Such developments can also be used to enrich and improve GCMs or to help with consistency tests. With today's advances in *ab initio* and density functional theory (dft) calculations and the availability of computational resources with ever increasing speed, the development of the above mentioned alternative non-empirical approaches through quantum chemical calculations appear feasible. There are various routes for these calculations but, typically, two types of calculations are performed: gas-phase thermochemistry and solvation/hydration. In the present work, which is the first part of a series of papers on the thermodynamics of metabolism, the focus will be on gas-phase quantum thermochemical calculations for key metabolites of biological interest. Hydration will be dealt with in a subsequent paper.

Intrinsic properties of metabolites from gas-phase quantum chemical calculations form the basis for understanding their behaviour in more complex systems such as proteins, nucleic acids and polysaccharides. Remarkable progress is made for small to moderate-sized compounds with up to ten non-hydrogen atoms. However, the biochemical compounds in metabolic pathways such as aminoacids and oligopeptides, saccharides, nucleosides and their derivatives, often exceed this molecular size limit. For the heavier metabolites, we further need computational information since reliable experimental information for them is rather limited and often non-existing. Most of the compounds of interest in this work are solid / crystalline at ambient conditions with high melting points and very low vapour pressures and, thus, are elusive to direct gas-phase measurements. Most often, the standard heat of formation of these compounds in the gaseous state are obtained by combining the heats of formation in the pure solid state (typically, from heats of combustion) and the heat of sublimation (typically, by extrapolation from values obtained at significantly higher temperatures – if the compounds are not decomposed). Both of these typical experimental measurements, however, require extreme care as they are prone to significant experimental errors[12]. Therefore, for many of the heavy metabolites described above, the basic thermochemical quantities in the ideal gas state still resist an experimental determination with the acceptable "chemical accuracy" of 1 kcal/mol. Attempts to develop predictive group contribution methods based on experimental results for smaller molecules[13, 14] are not always successful for the above types of metabolites and require significant improvements[15]. Thus, today there is much interest in accurate theoretical calculations of thermodynamic quantities of these metabolites.

In a recent thorough review[16], the state of the art in quantum thermochemical calculations is compared with the corresponding group-contribution method (GCM) approach. Under certain conditions, the level of accuracy of current quantum thermochemical predictions in the ideal gas state compares with or overpasses the thermochemical accuracy of 1 kcal/mol[17, 18] for small to moderate sized molecules with 2 to 10 non-hydrogen atoms. The Gaussian – n (Gn) family of quantum chemical procedures[18, 19] achieves the above level of accuracy for moderate sized molecules but the accuracy decreases sharply for heavier molecules. This holds true for the predictions of absolute thermochemical quantities via the atomization energy differences. Under certain conditions, accurate predictions can be made with designing appropriate isodesmic reactions even without using high levels of theory and computation[17, 18]. In isodesmic reactions, bond types and groups are kept the same on the two reaction sides; therefore any flaws in theory and systematic errors will be mutually compensated. This makes the isodesmic reaction approach a computationally efficient way to calculate the enthalpies and Gibbs free energies of formation for relatively large molecules with tens or more heavy atoms[19, 20].

The key prerequisite in an isodesmic reaction approach is the availability of accurate thermochemical information on all other reactants and products besides the studied molecule. A lot of work has been done in this respect for the development of extensive thermochemical databases[21-23 24-26].

Metabolic reactions take place in aqueous environments where metabolites often exist as ions. Significant advances have been made in the field of gas-phase ion thermochemistry, both experimentally and computationally[27-34]. In a recent review[35], standard values for gas-phase basicities and proton affinities were recommended for the 20 (protein) aminoacids. Gas phase acidities for these aminoacids were also reported in recent studies[34, 36]. Yet, heats and free energies of formation of ionic forms are rarely reported.

There have also been some studies on the gas-phase quantum thermochemistry of saccharides and nucleosides, but not as extensive as for aminoacids. Reliable experimental thermochemical data for these classes of metabolites, especially nucleosides, are rare. In the NIST Webbook of Chemistry[23] there is a noticeable scarcity of data even for neutral nucleosides. Also, the few reported data for saccharides[24, 25] are rather questionable, as we will discuss later. In contrast, gas- phase ion energetic data are available in the NIST Webbook for some of these metabolites. Experimental and theoretical gas-phase basicities and proton affinities are known for various protonation sites of glucopyranose[37]. Values for these properties were also reported for three protonation sides on adenosine[38]. On the other hand, there exists extensive work on the relative energies of the various conformers of saccharides with respect to their most stable anomer conformers or the sugar puckering[39-41]. The effect of sugar puckering on the energetics of nucleosides was also studied[42].

Most of the thermodynamic studies in the literature are confined to ambient conditions. Their extension to remote temperatures and pressures requires accurate knowledge of additional properties relevant to the interaction of the studied metabolites with their neighbouring molecules. Heat capacities, thermal expansivities and isothermal compressibilities are among the properties that must be known for such an extension. In addition, the effect of external conditions on solvation phenomena, primarily hydration phenomena, is of key importance in the evolution of biochemical reactions and processes at remote conditions. This includes metabolite dissociation, ligand-binding, protein folding, and ion distributions. In this regard, molecular thermodynamics, which combines quantum chemical calculations with classical mixture thermodynamics and equation-of-state approaches for extrapolation to high temperatures and pressures, can be used for understanding the metabolism of microorganisms, not only at ambient conditions but also at extreme conditions of temperature and pressure[43-52].

The present work is the first part of a series of publications aiming at the development of a cohesive approach for the reliable estimation of the basic thermochemical quantities of metabolites, biomolecules and associated biochemical reactions over an extended range of external conditions of temperature and pressure. In this paper the focus will be on quantum chemical calculations of enthalpies and free energies of formation in gas phase via isodesmic reactions of key metabolites, such as aminoacids/oligopeptides, oligosaccharides, and nucleotides /nucleosides, for which experimental information is difficult to obtain. The obtained quantum chemical calculations complement the available thermochemical compilations and form the basis for the expansion and testing of subsequent developments. In the second part of this work, we will use the information obtain from this work to apply GCM [11] for the calculation of thermodynamic properties in aqueous and gas phases at ambient condition. Solvation / hydration and extension to remote T, P conditions will be the subject of subsequent parts.

2. Methods

2.1. Computational Tools

The heavy metabolites of interest in this work are fairly flexible and exist in a vast conformational/isomerization space, which makes their detailed theoretical computations a challenging task[53,54].

For the conformer search, we used the COSMOconfX suite (Cosmologic GmbH, Germany) and the Conformer Analysis application of Spartan 14 suite (Wavefunction, USA) – a Monte Carlo / Molecular Mechanics algorithm. The first gross selection of the prevailing conformers was screened further down to a few conformers by performing energy calculations with progressively increasing basis set. We reinserted some conformers with extensive intramolecular hydrogen bonding, which were rejected by the above search algorithm into the pool for further calculations at a higher level. Whenever available, we inserted optimal geometries from literature into the pool at this stage. The prevailing conformers were subject to further geometry optimization at progressively higher levels until the most stable conformer was identified. We performed further calculations for the isodesmic reactions with the most stable conformer or with the few (less than five) prevailing conformers.

We did Quantum chemical calculations at the DFT-D3 level with Grimme's dispersion correction (D3 London dispersion correction) with Becke–Johnson damping [55] as implemented in TURBOMOLE suite[56] with the resolution of the identity RI-J approximation[57]. Geometry optimization and vibrational frequency analysis were done with the Becke-3-Lee-Yang-Parr (B-3LYP) 3- parameter hybrid functional with Becke's popular nonlocal exchange functional and Lee/Yang/Parr nonlocal correlation functional[58, 59] with the def2-TZVP (Karlsruhe segmented contracted triple-z valence quality plus polarization) basis set[56]. All geometry-optimization / frequency calculations were performed with TURBOMOLE v. 6.5 suites (Cosmologic GmbH, Germany). Using this geometry, we conducted single point energy (SPE) calculations at a higher level in two alternative ways: First, the generalized gradient approximation (GGA) B97-D density functional[60, 61] was used with the quadruple z-valence quality def2-QZVPD basis set as implemented in TURBOMOLE suite with the above Grimme's D3 dispersion correction. We did the SPE calculations for all compounds considered, neutral or ionic. Second, we used the Chai and Head-Gordon ωB97X-D long-range corrected hybrid density functional[62] with the 6-311++G(2df,2p) basis set, as implemented in Gaussian 09 and in Spartan 14 suites of programs. Calculations at the second level were done on selected metabolites. Higher level of theory and especially larger basis sets are useful in such calculations with intramolecularly hydrogen-bonded metabolites in order to exclude basis set incompleteness errors (BSIE) or basis set superposition errors (BSSE).

We performed vibrational frequency calculations in order to verify that the structures were minima and, also, to obtain the zero point vibrational energy (ZPE) and the thermal corrections to the enthalpy, H_v, and free energy, G_v. We calculated the latter quantity from the entropy change through the classical equation, $\Delta_f G = \Delta_f H - T\Delta_f S$. The entropy term in the heavy metabolites is dominated by their many low frequencies[63], which are usually poorly approximated in the quantum thermochemical calculations, therefore the "normal-mode" approximation may not be valid for heavy and flexible molecules.

2.2. Accounting for Conformers, Anomers and Tautomers

Detailed conformational analysis and complete exploration of conformational space is a tedious process[53], especially for heavy metabolites with hundreds or thousands of conformers in both their canonical (neutral) and ionic form. We did not perform such a detailed analysis and the globally most stable conformer was, probably, not identified in some cases. From this point of view, our calculations should be considered as setting an upper limit in the thermochemical quantities of isolated conformers.

In figure 1, we compared the most stable structures obtained for methionine, glutamic acid and valine with the corresponding structures proposed by Stover et al[34], and the most stable structure for open chain glucose with the corresponding structures proposed by Sameera et al.[40] The differences in the calculations are not significant in the case of aminoacids. What is worth pointing out is the different stabilizing factors in the two cases. As is clear from the figure, the second row structures (this work) exhibit intramolecular hydrogen bonding that includes the relatively strong OH--- NH bond. In contrast, the upper row structures do not exhibit such intramolecular hydrogen bonds and appear as less compact structures. Based on these properties we expect that the upper row structures prevail at higher temperatures while the lower row structures be dominant at lower temperatures. In the case of glucose chains, we observe that both structures are stabilized by four intramolecular hydrogen bonds of the same OH--- O type and the difference in the calculations is not negligible. The right structure (this work) is more stable by ca. 9 kJ/mol. The reason for this difference resides on the more cooperative character of hydrogen bonding in the right structure (four vs. three consecutive hydrogen bonds)[64] and the relatively open structure (free terminal OH).

From the thermodynamic point of view, the pertaining value for the metabolite is an average over the conformer population considered coexisting at equilibrium. This averaging is typically done by adopting the classical Boltzmann distribution equation[65]:

$$X_i = \frac{exp\left(-\frac{\Delta_f G_i}{RT}\right)}{\Sigma_j exp\left(-\frac{\Delta_f G_j}{RT}\right)} \quad (1)$$

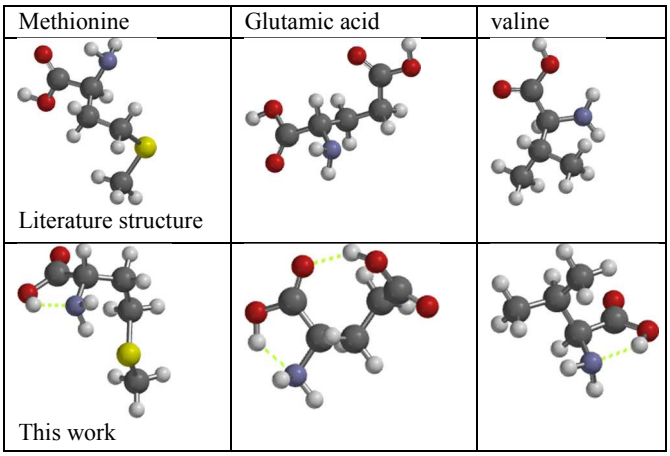
where X_i is the mole fraction of conformer i at equilibrium and $\Delta_f G_i$ is the Gibbs free energy of formation of conformer i. The summation is over all conformers of the studied molecule. The contribution of each conformer to the overall heat of formation of the studied metabolite is just the product of conformer's heat of formation with its mole fraction at equilibrium, and similarly for the Gibbs free energy of formation. Using equation (1), we observe that this contribution falls rapidly as the conformer's Gibbs free energy of formation departs from its lowest value in the conformer set, or, as the relative Gibbs free energy of formation departs from zero (Figure 2). Note that the relative Gibbs free energy corresponds to the difference of Gibbs free energy of formation of conformer i from the corresponding quantity of the global minimum or most stable conformer, i.e.

$$\Delta\Delta_f G_i = \Delta_f G_i - \Delta_f G_{min} \quad (2)$$

Thus, although the heat of formation of some conformers may depart by as much as 30 kJ/mol from the corresponding lowest value, the so-calculated average heat of formation differs by less than 4.0 kJ/mol from the lowest conformer's value.

The above conformer distribution analysis is valid as long as the conformer population is known sufficiently well. Omitting conformers near the global minimum may lead to significant errors. However, omitting all conformers above the usual cut-off of 1 kcal/mol in relative Gibbs free energy may also introduce a non-negligible error since the omitted conformers may be numerous. Another source of error is the fact that Gibbs free energies of formation are not usually known with high precision and the above averaging is then done with the overall energy.

A



B

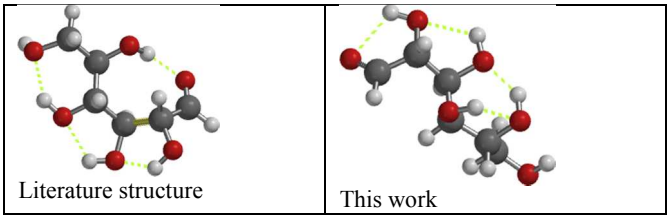


Figure 1: A: Comparison of most stable structures for methionine, glutamic acid and valine with the corresponding structures proposed by Stover et al.[34]. B: Comparison of the most stable structure of glucose chain with the corresponding literature structure[40].

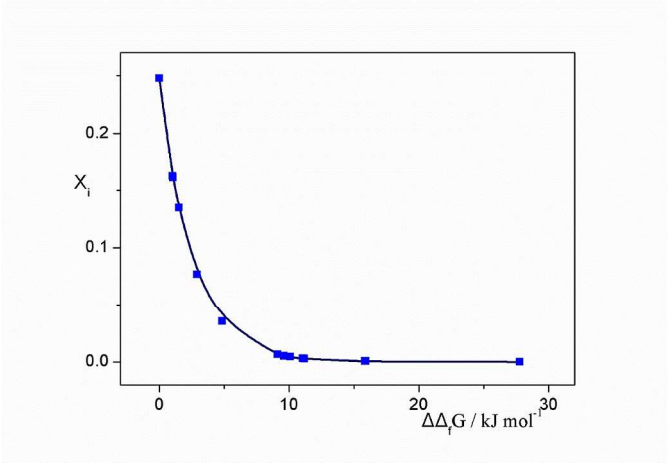


Figure 2: The mole fraction of conformers at equilibrium as a function of the corresponding relative Gibbs free energy of formation for aspartic acid.

In the case of conformers with extensive intramolecular hydrogen bonding, which is often the case with heavy metabolites, this error may not be negligible either.

The situation may be different in the case of tautomers. As figure 3 shows, for the calculation of the thermochemical quantities for inosine, as an example, we need the corresponding quantities for hypoxanthine. Hypoxanthine may be found in 7 tautomeric forms including the enol form of 6-hydroxypurine.

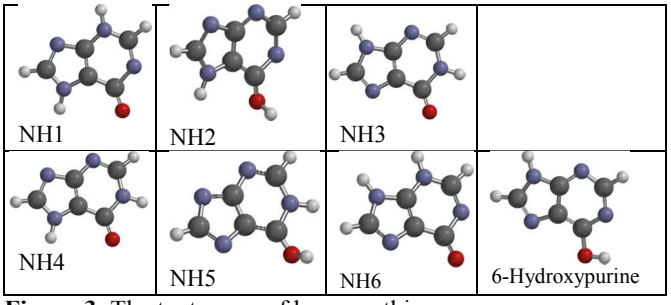


Figure 3: The tautomers of hypoxanthine

These structures correspond to a well-known isomer population for which equation (1) may be applied unequivocally. One may argue that some of these tautomers are not relevant since the ribose group occupies one site in inosine. However, regarding hypoxanthine itself, its average thermochemical quantities may still be obtained by applying equation (1), leading to a value of just 0.4 kJ/mol higher than the heat of formation of the most stable conformer, NH4. The situation in tautomers of nucleosides or histidine is different since some of them may be favored (stabilized) by strong intramolecular hydrogen bonds as shown in figure 4. The averaging results for guanine and cytosine are similar. Although one may infer from this discussion that the thermochemical quantities of the most stable tautomer are good approximations for the corresponding quantities for the metabolite, the knowledge of these quantities for the less stable tautomers are also important. We will come back to tautomers in a subsequent paper since the prevailing hydrated tautomer (lowest Gibbs free energy of formation in dilute aqueous solution) may not be the same with the one prevailing in the ideal gas state[66, 67]. The case of sugar anomers with open and cyclic structures is again different. The results through detailed sugar puckering analyses[40, 41] show that particular puckering geometries may be favoured in various glycobiology processes over the usual low energy equatorial

$4C_1$ conformation and, thus, a detailed knowledge of the puckering landscape is crucial for understanding these processes.

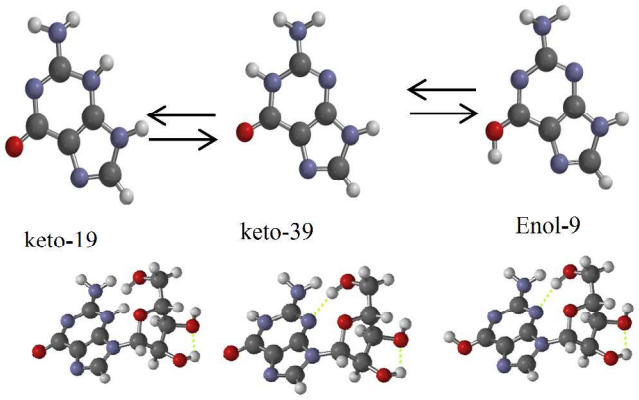


Figure 4: Upper part: The equilibrium between the keto-19 and keto -39 tautomers and enol-9 tautomer of guanine. Lower part: The corresponding tautomers of guanosine just under each guanine tautomer. Dashed lines show intramolecular hydrogen bonds.

Thus, it makes more sense to focus on the thermochemical quantities of the global minimum in each class of structures (open chain, furanoses, pyranoses) and use the above puckering analyses for the averaging via equation (1). However, considering the reported range and the distribution of enthalpy differences in the above puckering analyses, it is clear that the thermochemical quantities of sugars do not deviate significantly from the corresponding quantities of the global minimum in each case.

2.3. Heats and Gibbs Free-Energies of Formation in the Ideal Gas State of Key Metabolites

Among the heavy metabolites, the most studied one in the quantum thermochemistry literature are the aminoacids[34, 68]. Yet, the deviations of reported calculations prevent the establishment of a universally accepted database with reliable heats and Gibbs free energies of formation even for aminoacids. Reliable data for few aminoacids are reported[24, 25]. The lack of reliable data for the other two classes is more noticeable. The crucial factor is the availability of reliable data (such as in ATcT[24, 25]) for the lighter compounds that appear in the isodesmic reactions and this is important in the selection of the appropriate isodesmic reactions for each metabolite.

2.3.1. The Selection of Isodesmic Reactions

Apart from the standard requirement for the preservation of number and type of bonds on both sides of the isodesmic reaction, heavy metabolites may pose additional requirements as they may exhibit extensive intramolecular hydrogen bonding. One may wonder whether the reactants should account for the intramolecular interactions of the metabolites.

Glycerol is a good example for the case. Glycerol may exhibit two intramolecular hydrogen bonds, which further stabilize its conformers. Simple conformer search shows that conformer #1 in figure 5 is the most stable glycerol conformer. Two alternative isodesmic reactions for obtaining glycerol are the followings:

2 ethanol + methanol => glycerol + 2 methane (A)
or
2 ethylene glycol => glycerol + methanol (B)

In reaction B, we also have the option to choose between alternative reactant (ethylene glycol) conformers, as shown in figure 5.

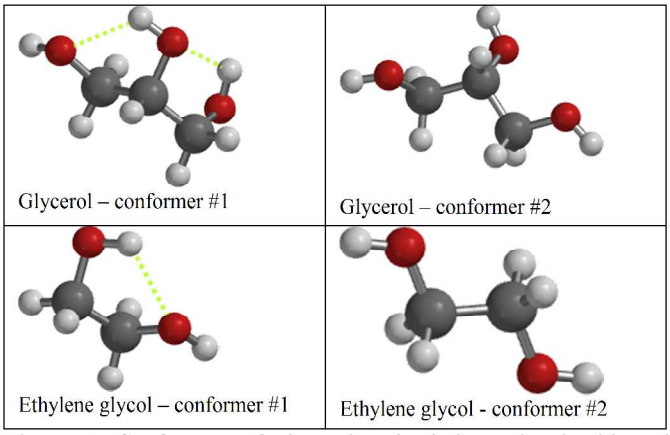


Figure 5: Conformers of glycerol and ethylene glycol with and without intramolecular hydrogen bonds (shown by dashed lines).

Table 1 summarizes the predicted thermochemical quantities with calculations at the DFT-D3/B3LYP/def2-TZVPD level for the various alternative reactions or reactant conformer combination schemes. As shown in this table, the isodesmic reaction A could lead to acceptable predictions. However, the reaction heats (and free energies) are not negligible. This may entail some error from the inefficient compensation of theory flaws from both reaction sides. As expected, the use of isodesmic reaction B with ethylene glycol conformer #2 leads to unacceptable results. In contrast, the intramolecularly hydrogen-bonded conformer #1 of ethylene glycol is the most appropriate reactant for isodesmic reaction B as shown by the predicted results in table 1, which also conforms to experiment. The heat of reaction in this case (-4.9 kJ/mol) is small and is probably indicative of a further stabilization of glycerol conformer #1 due to cooperativity of its dual intramolecular hydrogen bonding.

The glycerol example will be used as a guide for the selection of isodesmic reactions and conformer reactants in this work and especially in the case of saccharides. Thus, the isodesmic reaction for the formation of βD-fructopyranose from tetrahydropyran (THP) and ethylene glycol will be as follows:

$$THP + 3HOCH2CH2OH \Rightarrow \beta DFrp + CH3OH + 2C2H6$$

To be able to observe the hydrogen bonds, we show below the structural configuration of ethylene glycol and βDfructopyranose:

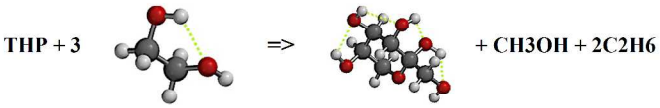


Figure 6: The isodesmic reaction for the formation of βDfructopyranose displaying the intramolecular hydrogen bonds (shown by dashed lines) on both sides.

By using this presentation for isodesmic reaction, we preserve types and numbers of covalent bonds and (most of the) intramolecular hydrogen bonds. The isodesmic reactions that have been considered in the present work are shown in Table 2.

Table 1: The predicted thermochemical quantities for glycerol based on quantum chemical calculations at the DFT-D3/B3LYP/def2-TZVPD level.

Glycerol conformer	Heat of reaction (kJ/mol)	Heat of formation (kJ/mol)		Free energy of formation (kJ/mol)	
		Predicted	Experim.	Predicted	Experim.
Isodesmic reaction A					
#1	-62.0	-583.7	-577.9[1], -582.8[2]	-445.8	-448[2]
#2	-48.5	-570.0		-435.6	
Isodesmic reaction B, ethylene glycol conformer #1					
#1	-4.9	-582.6	-577.9[1], -582.8[2]	-446.0	-448[2]
#2	+8.7	-569.1		-435.7	
Isodesmic reaction B, ethylene glycol conformer #2					
#1	-22.9	-617.7	-577.9[1], -582.8[2]	-462.3	-448[2]
#2	-9.3	-604.1		-452.0	

[1]Goos/ATcT [24, 25]; [2]DIPPR[26]

Table 2: The isodesmic reactions for neutral metabolites

Metabolite	Isodesmic Reaction
Aminoacids / dipeptides	
H2NCHRCOOH	H2NCHRCOOH (Gly) + CH3R => H2NCHRCOOH + CH4
Arginine (Arg)	Gly+guanidine+butylamine => Arg + CH4 + NH3
Histidine (His)	Gly+imidazole+propane => His+2 CH4
Proline (Pro)	pyrrolidine+acetic acid => Pro+ CH4
Tryptophan (Try)	Gly+indol+propane => Try+2 CH4
Alanylalanine (AlaAla)	2Ala +CH3NH2 => AlaAla + CH3OH + NH3
Alanylglycine (AlaGly)	Ala + Gly + CH3NH2 => AlaGly + CH3OH + NH3
Alanylphenylalanine (AlaPhe)	Ala + Phe + CH3NH2 => AlaPhe + CH3OH + NH3
Glycylglycine(GlyGly)	2Gly + CH3NH2 => GlyGly + CH3OH + NH3
Phenylalanylglycine (PheGly)	Gly + Phe + CH3NH2 => PheGly + CH3OH + NH3
Glycylvaline (GlyVal)	Gly + Val + CH3NH2 => GlyVal + CH3OH + NH3
Leucylglycine (LeuGly)	Leu + Gly + CH3NH2 => LeuGly + CH3OH + NH3
Saccharides	
αD-Glucose (αDGl)	3HOCH2CH2OH+tetrahydropyran(THP) => αDGl + CH3OH + 2C2H6
βD-Glucose (βDGl)	3HOCH2CH2OH + THP => βDGl + CH3OH + 2C2H6
Glucose – chain (GlCh)	3HOCH2CH2OH + C2H6 + CH3CHO => GlCh + CH3OH + 3CH4
2-deoxy-D-Glucose (2-deGl)	2HOCH2CH2OH + THP => 2-deGl + CH4 + C2H6
βD –Galactose (βDGal)	3HOCH2CH2OH + THP => βDGal + CH3OH + 2C2H6
βD –Mannose (βDMan)	3HOCH2CH2OH + THP => βDMan + CH3OH + 2C2H6
βD –Ribose (βDRib)	2HOCH2CH2OH + THF => βDRib + CH4 + C2H6
αD-Ribofuranose (αDRib)	2HOCH2CH2OH + THF => αDRib + CH4 + C2H6
2–deoxyD-ribose(2-deRib)	2HOCH2CH2OH + THF => 2-deRib + CH3OH + C2H6
Fructose-chain (FrCh)	3HOCH2CH2OH + CH3COCH3 => FrCh + CH3OH + 2CH4
βD-Fructopyranose (βDFrp)	3HOCH2CH2OH + THP => βDFrp + CH3OH + 2C2H6
βD-Fructofuranose (βDFr)	3HOCH2CH2OH + THF => βDFr + CH3CH2OH + 2CH4
βD–Xylofuranose (βDXyf)	3HOCH2CH2OH + THF => βDXyf + CH3CH2OH + CH3OH + C2H6
βD–Xylopyranose (βDXyp)	2HOCH2CH2OH + THP => βDXyp + 2 C2H6
αD–Xylopyranose (αDXyp)	2HOCH2CH2OH + THP => αDXyp + 2 C2H6
Sucrose	βDGl + βDFr + CH3COCH3 => Sucrose + 2 CH3OH
Maltose	2βDGl + CH3COCH3 => Maltose + 2 CH3OH
Nucleobases/Nucleosides	
6-Hydroxypurine	Adenine + CH3OH => 6-Hydroxypurine + CH3NH2
Adenosine	Adenine + βDRib + CH3NH2 => Adenosine + CH3OH + NH3
2-Deoxyadenosine	Adenine + 2-deRib + CH3NH2 => 2-Deoxyadenosine + CH3OH + NH3
2,3-Dideoxyadenosine	Adenosine + 2 CH4 => 2,3-Dideoxyadenosine + 2 CH3OH
Inosine	Hypoxanthine + βDRib + CH3NH2 => Inosine + CH3OH + NH3
Guanosine	Guanine + βDRib + CH3NH2 => Guanosine + CH3OH + NH3
Cytidine	Cytosine + βDRib + CH3NH2 => Cytidine + CH3OH + NH3
Thymidine	Thymine + βDRib + CH3NH2+ CH4 => Thymidine + 2CH3OH + NH3
Deoxythymidine	Thymine + THF+ CH3NH2+ CH3CH2OH => Deoxythymidine + 2CH4 + NH3
Uridine	Uracil+ βDRib + CH3NH2 => Uridine + CH3OH + NH3

2.3.2. Thermochemical Calculations for Neutral Metabolites

We calculated the heats and Gibbs free energies of the isodesmic reactions (Table 2) at 298.15 K and ambient pressure. We added ZPE (zero point vibrational energy) to SPE together with the thermal corrections for the enthalpy or the Gibbs free energy for each reactant or product. This results in their total enthalpy, H^0, and Gibbs free energy, G^0:

$$H^0 = SPE + \alpha ZPE + H_v \quad (3)$$

and

$$G^0 = SPE + \alpha ZPE + G_v \quad (4)$$

The scaling factor, α, has values close to 1. In this work it was set equal to 0.975. Detailed tables with this information are provided as supplementary material. The heat of a reaction

$\nu_1 R_1 + \nu_2 R_2 \Rightarrow \nu_3 R_3 + \nu_4 R_4$

can then be obtained in the usual manner as

$$\Delta_r H^0 = \nu_3 H^0_3 + \nu_4 H^0_4 - \nu_1 H^0_1 + \nu_2 H^0_2 \quad (5)$$

A similar approach can be used for the free energy of the same reaction, $\Delta_r G^0$. We can also derive the same result from the heats of formations of reactants and products, or:

$$\Delta_r H^0 = \nu_3 \Delta_f H^0_3 + \nu_4 \Delta_f H^0_4 - \nu_1 \Delta_f H^0_1 + \nu_2 \Delta_f H^0_2 \quad (6)$$

and similarly for $\Delta_r G^0$.

Calculation of the standard heats and free energies of formation of a metabolite requires accurate knowledge of the corresponding properties for all reactants and products of the isodesmic reaction. By replacing the known quantities we can use equation (6) in order to obtain the unknown heat of formation (similarly the unknown Gibbs free energy of formation) of the studied metabolite. The existing compilations[23-26] provide reliable data for these computations, which are summarized in Table 3 for the compounds that participate in our isodesmic reactions.

Table 3: Thermochemical data for the reactants of isodesmic reactions (table 2)

Compound	$\Delta_f H$ / kJ mol^{-1}	$\Delta_f G$ / kJ mol^{-1}
Methane	-74.52	-50.49
Ethane	-83.85	-31.92
Propane	-104.68	-24.39
n-Butane	-125.79	-16.7
isoButane	-134.18	-20.76
n-Pentane	-146.76	-8.813
isoPentane	-153.6	-14.05
Methanol	-200.94	-162.32
Ethanol	-234.95	-167.85
1-Propanol	-255.2	-159.9
2-Propanol	-272.7	-173.47
1-Butanol	-274.6	-150.3
Ethylene glycol	-387.5	-302.6
p-Ethylphenol	-144.05	-21.58
Acetaldehyde	-166.2	-132.8
Acetone	-215.7	-151.3
Acetic acid	-432.25	-374.6
Propanoic acid	-453.5	-366.7
Butanoic acid	-475.8	-360.0
Dimethyl ether	-183.94	-112.8
Methylamine	-19.38	32.07
Dimethylamine	-18.45	68.39
Propylamine	-70.1	41.7
Butylamine	-91.9	49.3
Pentylamine	-113.2	57.3
Ammonia	-45.57	-16.4
Benzene	82.88	129.6
Ethylbenzene	29.92	130.73
Tetrahydrofuran (THF)	-182.5	-79.69
Tetrahydropyran (THP)	-224.28	-80.37
Acetamide	-238.3	-159.53
Propanamide	-259.0	-151.05
Butanamide	-282.0	-132.13
Imidazole	131.5	192.75
Indole	156.6	237.3
Pyrrolidine	-3.6	114.7
Guanidine	27.95	95.94
Methyl propyl sulfide	-82.3	17.93
Adenine	225.7	348.67
Guanine	16	184.49
Cytosine	-69.5	50.24
Uracil	-301.5	-198.35
Thymine	-338	-192.27

2.3.3. Thermochemical Calculations for Ionic Metabolites

Gas phase acidities, gas phase basicities (GB) and proton affinities (PA) are important elements for understanding the acid-base behavior of our metabolites not only in the gas phase but also in aqueous solutions and numerous biological processes[66, 67]. The extensive experimental and theoretical research over several decades has led to reliable compilations of the above properties, especially for aminoacids[27, 29, 30, 34, 35]. For this case, one can use these experimental data to estimate heats and free energies of formation of protonated and deprotonated aminoacids if the corresponding quantities for the neutral species are available. On the other hand, one can perform the same level of quantum thermochemical calculations as described in section 2.1, and use appropriate isodesmic reactions in order to estimate the thermochemical quantities of the ionic metabolites. We have followed both approaches and we present the results in the next section. In order to clarify the computations, we first recall the definitions of gas phase acidity, GB, and proton affinity, PA.

Let reaction C be the protonation reaction of metabolite M :

$$M\,(g) + H^+\,(g) \Rightarrow MH^+\,(g) \quad (C)$$

and reaction D the deprotonation reaction for the same metabolite:

$$M\,(g) \Rightarrow [M\text{-}H]^-\,(g) + H^+\,(g) \quad (D)$$

The gas-phase basicity of metabolite M is equal to the negative of the standard Gibbs free energy change of reaction C, i.e.,

$$GB\,(M) = -\Delta_r G^0_C, \quad (7)$$

and its proton affinity is equal to the negative of the standard enthalpy change of the same reaction C, i.e.,

$$PA\,(M) = -\Delta_r H^0_C \quad (8)$$

The standard Gibbs free energy change of the deprotonation reaction D is the gas-phase acidity of metabolite M, i.e.,

$$\Delta G_{acid}\,(M) = \Delta_r G^0_D \quad (9)$$

and similarly for the enthalpy:

$$\Delta H_{acid}\,(M) = \Delta_r H^0_D \quad (10)$$

Calculations of gas phase acidities of aminoacids are typically done via the isodesmic reaction E

$$Aminoacid + CH3COO^- \Rightarrow [Aminoacid\ –H]^- + CH3COOH \quad (E)$$

and by anchoring the calculations to the known gas phase acidity for acetic acid[29, 33, 36]. Alternatively, one may use benzoic acid in the isodesmic reaction E or even glycine as there exist consensus values for it[29], namely ΔG_{acid} (Gly) = 1403.5 ± 1.5 kJ/mol and ΔH_{acid} (Gly) = 1433.5 ± 1.5 kJ/mol. Our calculations for ΔG_{acid} and ΔH_{acid} for glycine (1404.1 and 1434.4, respectively) fall within the consensus range and were used as a basis for the calculation of heats and free energies of formation of our deprotonated metabolites. For this purpose, however, we also need values for the heat and free energy of formation of deprotonated glycine itself. These quantities require

Physical Chemistry Chemical Physics

the knowledge of corresponding values for proton (H^+) or any other cation. The adopted standard thermochemical values for proton in this work, $\Delta_fG^0(g, H^+)$ = 1513 kJ/mol and $\Delta_fH^0(g, H^+)$ = 1533 kJ/mol, were recommended by Truhlar et al[69]. and were obtained by using the electron convention and the Fermi-Dirac statistics (not the Boltzmann statistics). Thus, the heat of formation of deprotonated glycine (cf. reaction D) is: Δ_fH (glycine-, g) = ΔH_{acid} (Gly) - $\Delta_fH^0(g, H+)$ + Δ_fH (glycine, g) = -491 kJ/mol and, similarly, Δ_fG (glycine-, g) = -409 kJ/mol. The thermochemical quantities of all other aminoacids can be obtained either directly as with glycine or through the isodesmic reaction

Aminoacid + glycine- => [Aminoacid – H]$^-$ + glycine (F)

We further obtain the thermochemical quantities for the protonated metabolites in the same way.

3. RESULTS AND DISCUSSION

We calculated the gas-phase standard heats and Gibbs free energies of formation of neutral metabolites and compared our estimations with available literature data, both, experimental and computational (table 4). In the case of aminoacids, our emphasis is on the comparison with the recent comprehensive calculations by Stover et al[34]. This is an important reference work since it reports heat of formation calculations with both the atomization energy route (predictions at the composite G3 – MP2 level) and the isodesmic reaction route. In addition, it claims to be "the first reliable set of these values in the gas phase" and thus constitutes a plausible testing benchmark for our own calculations. Table 4 also includes estimations by the composite T1 method (a modified G3-MP2 method) via Spartan 14 suite (Wavefunction, USA). Regarding GCM calculations for aminoacids and peptides, two representative works are those of Domalski and Hearing[14] and the more recent one by Sagadeev et al[70]. In general, our calculations for aminoacids fall within the range of literature values. They are also in agreement with the recent calculations by Stover et al[34]. On average, the discrepancy between the two sets of calculations via isodesmic reactions is less than 1 kcal/mol. Our calculations are on average more negative, most probably due to the fact that our reference value for glycine (-391.7 kJ/mol), which seems to be more widely accepted[15, 25, 68, 71], is lower than the value adopted by Stover et al[34] by almost 7 kJ/mol. Considering this different reference value, the two sets of calculations are well in accordance. However, our calculations still remain on average more negative than their G3-MP2 calculations. Stover et al. do not report Gibbs free energies of formation for comparison. Due to these discrepancies caused by the differences in the adopted reference values, we also report in Table 4 the heats and free energies of the isodesmic reactions that are not affected by such reference values.

Dipeptides offer new possibilities for intramolecular hydrogen bonding that were absent in their parent aminoacids, as shown in figure 7. The NH---N is not a very strong hydrogen bond but other types of stronger bonds may appear bringing in further stabilization of oligopeptides.

Table 4: Estimated standard thermochemical properties (ideal gas phase) of neutral metabolites from isodesmic reactions (Table 2). Geometry optimization and frequency calculations at DFT-D3/B3LYP/def2-TZPV(D) level and energy (SPE) calculations at DFT-D3/B97-d/def2-QZPVD level. Values in parentheses were calculated at the same geometry but with SPEs at the DFT/ωB97X-d/6-311++G(2df,2p) level. All quantities are given in kJmol^{-1}.

Metabolite	Δ_rG	Δ_fG	Δ_rH	Δ_fH	
	estimated				**literature**
Aminoacids					
Alanine	-16.4	-298.0 (-298.0)	-21.0	-421.9 (-422.0)	-415.9[a] , -421.3[a] -414.7[g], -435.5[c], -415.9[h], -419.4[b]
Arginine	-42.0	-129.9	-63.4	-398.9	-389.1[a] ,-395.8[a], -380.6[b]
Asparagine	-10.8	-411.4 (-409.0)	-35.9	-612.0 (-610.0)	-591.8[a] , -610[a], -593.8[b] -609.1[c], -590.5[h]
Aspartic acid	-5.5	-621.8 (-622.7)	-19.6	-790.2 (-791.1)	-786.6[a] ,-793.3[a], -804.4[c], -786.7[h], -787.8[b]
Cysteine	-18.3	-272.8	-33.3	-396.5	-378.2[a], -395[a], -378.1[h], -382.6[n], - 397.1[n]
Glutamic acid	-21.4	-631.0 (-633.6)	-26.1	-819.0 (-821.7)	-807[a] , -815.9[a], -825.0[c], -807.4[h], -810.7[b]
Glutamine	-14.1	-395.8	-26.9	-626.0	-629.7[c], -621.7[a] -611.2[h], -618.1[b]
Glycine		-300.1[e]		-391.7	*-391.7[e]* -392.1[f], -390.4[k]
Histidine	-56.9	-87.6	-75.9	-289.9	-221.8[h],-271.1[a] , -267.6[b] , -289.5[i]
Isoleucine	-16.0	-279.7 (-282.5)	-22.8	-493.5 (-496.3)	-486.6[a], -493.3[a], -487.1[b], -499.6[c], -486.8[h]
Leucine	-17.0	-280.6 (-282.7)	-24.5	-495.2 (-497.3)	-486.6 [a],-494.1[a], -489.7[b], -501.8[c], -486.8[f], -486.8[h], -487[5]
Lysine	-13.6	-205.9	-28.5	-458.8	-443.5[a] , -451.5[a], -443.4[h], -444.2[b]
Methionine	-16.2	-248.0 (-246.8)	-28.1	-427.8 (-426.4)	-412.1[a], -428.9[a], -426.5[b], -412.1[h]
Phenylalanine	-23.3	-142.2 (-139.0)	-36.3	-323.5 (-320.4)	-302.1[a] ,-322.6[a], -319.6[c], -312.9[f], -302[h], -318.4[b]
Proline	-12.4	-221.8	-29.8	-391.1	-373.2[a] , -387[a], -366.2[f,l], -373.3[h], -385.5[m]
Serine	-16.0	-433.5 (-433.6)	-29.5	-581.5 (-580.0)	-567.8 [a],-578.2[a], -567.8[h], -578.4[b]
Threonine	-16.5	-439.5 (-439.9)	-30.2	-620.0 (-620.3)	-603.8[a] ,-618.8[a], -620.9[c], -603.7[h], -616.0[b]

Tryptophan	-43.6	-29.8*	-58.0	-246.9	-217.2[a] ,-249.4[a], -215[h], -238.1[b]
Tyrosine	-17.2	-288.4 (-286.4)	-27.8	-489.0 (-487.0)	-482[a] , -490.4[a], -498.5[c], -481.9[h], -489.2[b]
Valine	-11.9	-282.2 (-283.4)	-21.3	-472.6 (-473.9)	-466.1[a] , -472.4[a], -481.2[c], -455.2[f], -466.1[h], -468.7[b], -455.2[j]
Dipeptides					
Alanylalanine	-15.1	-400.2	-18.4	-635.1	-648.3[c], -623.7[h]
Alanylglycine	-10.3	-397.6	-13.9	-600.4	-588.2[c]
Alanylphenylalanine	3.2	-226.1	4.8	-514.0	-534.9[b], -532.4[c] , - 509.7[h]
Glycylglycine	-11.0	-400.4	-15.6	-571.8	-528[c], -571.9[h]
Phenylalanylglycine	-15.1	-246.5	-10.0	-498.0	-500.9[b]
Glycylvaline	-9.8	-381.3 (-378.4)	-15.0	-652.2 (-649.3)	-633.8[c], -647.9[h]
Leucylglycine	-15.5	-385.4	-22.6	-682.4	-652.2[c]
Nucleobases / Nucleosides					
Hypoxanthine	-21.8	172.8	-21.5	38.0	
6-Hydroxypurine	19.5	194.6	15.4	59.5	72.6[b]
Adenosine	-21.2	-123.0 (-128.7)	-31.6	-481.6 (-487.4)	-472.7[b]
2-Deoxyadenosine	29.3	18.2 (20.3)	40.1	-315.2 (-313.3)	-288.1[b]
2,3-Dideoxyadenosine	60.9	161.6	77.0	-151.9	-139.9[b]
Inosine	-12.0	-310.5 (-305.8)	-22.0	-659.7 (-655.1)	-646.4[b]
Guanosine	-12.1	-298.9	-22.5	-682.2	-664.2[b]
Uridine	-5.6	-675.3	-14.2	-991.5	-989.5[b]
Thymidine	30.2	-521.6 (-523.2)	30.6	-856.8 (-858.5)	
Deoxythymidine	-83.4	-373.8	-111.4	-691.5	-684.2[b]
Saccharides					
Cytidine	-12.7	-433.7	-18.9	-764.2	-768.9[b]
αD-Glucose	-74.5	-837.5	-91.2	-1115.0	
βD-Glucose	-74.6	-836.6	-89.1	-1112.8	-1040.1[e] -1113.5[b]
Glucose – chain	-78.0	-836.7	-109.1	-1102.5	-1016.2[e] -1095.5[b]
2-Deoxy-D-Glucose	-79.6	-682.7	-97.3	-941.9	-949.6[b]
βD –Galactose	-71.0	-833.0	-87.2	-1110.9	-1111.5[b]
βD –Mannose	-72.2	-834.2	-88.1	-1111.8	
βD-Fructopyranose	-89.1	-851.1 (-856.3)	-106.8	-1130.5 (-1136.9)	-1139.4[b]
βD-fructose (βD-fructofuranose)	-216.2	-934.8	-146.0	-1112.5	-1039.3[e]
Fructose-chain	-32.8	-828.6 (-829.0)	-60.5	-1093.2 (-1095.3)	-1113.3[b]
βD –Ribose	-79.6	-682.1	-100.1	-902.9	
αD-Ribofuranose	-78.5	-681.0	-102.2	-905.0	-907.9[b]
2-deoxy-D-Ribofuranose	-56.9	-547.5	-71.1	-747.6	
βD-Xylofuranose	-59.1	-684.5 (-572.4)	-76.6	-907.5 (-911.1)	-918.5[b]
βD-Xylopyranose	-72.0	-693.7	-84.8	-920.1	
αD-Xylopyranose	-72.6	-694.3	-86.6	-921.9	
Sucrose	45.9	-1419.2	33.2	-1985.2	
Maltose	3.9	-1459.2	-1.3	-2013.2	-1838.4[e]

[a]Stover 2012[34] ; [b]T1 (G3-MP2) (Spartan14, Wavefunction); [c]Domalski and Hearing 1993[14]; [d]Dorofeeva 2010[72]; [e]Goos database[25]; [f]Pedley 1994[22]; [g]Ngauv Sabbah 1977[71]; [h]Sagadeev 2010[70], [i]Riffet [65], [j]NIST-JANAF [73], [k]Ngauv [71], [l]Sabbah [74]. [m]Contineanu[75]. [n]Ribeiro da Silva 2010[76].

Table 4 shows inconsistencies between our calculations and the literature values. Direct comparison with the GCM predictions[14] cannot be made since they have used an erroneous value for glycine (-375.3 kJ/mol). The consistency with the composite T1 method (Spartan 14 database – Wavefunction) is very good in the case of PheGly but poor in the case of AlaPhe. However, our calculations for this latter dipeptide are in good agreement with the recent estimation by Sagadeev[15].

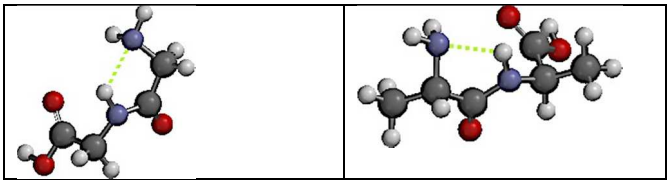


Figure 7: The NH---N intramolecular hydrogen bond in dipeptides GlyGly (left) and AlaAla (right).

The literature on saccharides and nucleosides is not as rich as that of aminoacids. The focus on saccharides is on the more stable furanose or pyranose closed ring structures[40, 41]. Their capacity to form multiple intramolecular hydrogen bonds, as shown in the previous section is contributing significantly to their stability and cannot be overlooked in selecting isodesmic reactions for their formation. There are very few literature data to compare with our calculations. Our calculations are consistent with the composite T1 predictions (Spartan14 dbase – Wavefunction) but we see noticeable differences with the reported ATcT values[25]. However, the discrepancies are too large (over 70 kJ/mol) to be attributed to erroneous calculations either via isodesmic reactions (our calculations) or via composite G3-MP2 method (Spartan 14). This is probably one of the very few cases to be reconsidered in ATcT compilation[25].

The nucleosides share features from nucleotides, especially tautomerism, and from the saccharide bD-ribose, especially the flexibility for intramolecular hydrogen bonding. As shown in figure 5, guanosine can form the strong OH –N hydrogen bonds with the guanine ring as well as the OH—O hydrogen bonds with ribose.

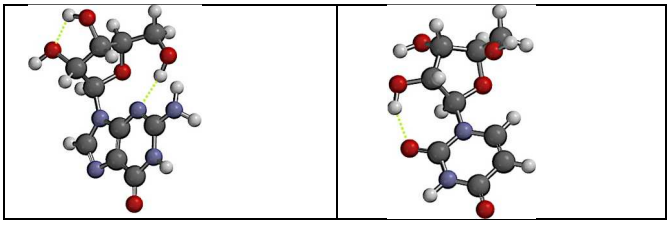


Figure 8. Hydrogen bonded conformers of guanosine (left) and uridine (right)

In uridine, the prevailing hydrogen bond is of the OH—O type with the carbonyl oxygen of the uracil ring. This hydrogen bonding capacity is a significant stabilizing factor for the nucleoside conformers and an adequate conformational search cannot overlook them.

For nucleosides, we made a thorough comparison in Table 4 with composite T1 calculations (Spartan 14 database – Wavefunction). The observed discrepancy between Spartan 14 and our estimations for adenosine, guanosine, inosine and derivatives is, in part, stemming from the fact that for adenine base, Spartan 14 database gives a heat of formation of 276.3 kJ/mol, which is significantly higher than the ATcT database value of 225.7 kJ/mol, while for guanine it gives a value of 44.09 kJ/mol, which is again higher than the ATcT value of 16 kJ/mol. The inconsistencies for the other nucleobases are significantly smaller as shown in the calculations.

In the majority of metabolites, the populations of species in the gas phase are dominated by the neutral as they are much more stable than the ionized structures. However, we also considered ionized structures, as this will contribute to our understanding of their hydration discussed in the third Part of this series of papers. Our main focus on ionic structures is on aminoacids for which we considered both protonation and deprotonation reactions. In Table 5 we report the gas-phase acidities (reaction enthalpies and Gibbs free energies) for aminoacids as well as the standard enthalpies and Gibbs free energies of formation of the deprotonated aminoacids. The agreement of our calculations with literature data is, in general, satisfactory. We can however observe noticeable differences such as for the case of ΔG_{acid} for methionine (18 kJ/mol). The formation quantities are all negative but there are no literature data for comparison.

Table 5: Estimated gas-phase acidities and standard thermochemical properties of ionic metabolites. Calculations as in Table 4. All quantities are given in kJmol^{-1}.

Metabolite	$\Delta_f G$	$\Delta_r G$ (ΔG_{acid})	$\Delta_r G$ lit	$\Delta_f H$	$\Delta_r H$ (ΔH_{acid})	$\Delta_r H$ lit
Aminoacids						
Alanine	-411	1400	1400[a], 1398.5[c]	-524	1431	1432[a,b] 1425±8.8
Arginine	-273	1356	1347[a], 1359.5[c]	-532	1388	1381[a], 1387[b] 1389±13
Asparagine	-572	1353	1354[a], 1359[c]	-760	1385	1386[a], 1384[b] 1388±13
Aspartic acid 1 Aspartate*	-830 -778	1305 1356	1316[b] 1364[a]	-991 -938	1332 1385	1349[a], 1345[b] 1394[a]
Cysteine	-411	1374	1369[a] 1364[c]	-525	1404	1399[a], 1396[b] 1393±13
Glutamic acid Glutamate*	-812 -795	1332 1349	1324[a] 1357[a]	-997 -979	1355 1373	1347[a], 1349[b] 1384[a]
Glutamine	-558	1351	1347[a], 1359[c]	-777	1382	1378[a,b] 1388±13
Glycine	-409	1404	1403[a], 1402[c]	-491	1434	1435[a], 1434[b] 1433[d]
Histidine	-252	1348	1345[a],	-447	1377	1376[a], 1374[b]

			1356[c]			1385±13
Isoleucine	-394	1399	1396[a], 1388.5[c]	-601	1426	1426[a,b] 1418±13
Leucine	-395	1398	1395[a], 1390[c]	-602	1427	1424[a], 1428[b] 1419±13
Lysine	-329	1390	1380[a] 1383[c]	-576	1416	1410[a], 1415[b] 1412±13
Methionine	-356	1403	1385[a], 1376[c]	-534	1427	1418[a], 1412[b] 1405±13
Phenylalanine	-268	1388	1384[a], 1379[c]	-433	1424	1416[a], 1417[b] 1408±13
Proline	-341	1394	1394[a] 1395[c]	-499	1425	1425[a], 1430[b] 1430±13
Serine	-583	1363	1363[a,c]	-722	1393	1392[a,b] 1392±13
Threonine	-592	1360	1359[a], 1360.5[c]	-763	1390	1390[a], 1397[b] 1390±13
Tryptophan	-155	1388	1390[a], 1380.5[c]	-358	1422	1423[a], 1422[b] 1410±13
Tyrosine	-415	1386	1382[a] 1378.5[c]	-602	1420	1415[a], 1419[b] 1408±13
Valine	-398	1391	1394[a], 1391[c]	-577	1427	1425[a], 1430[b] 1420±13

[a]Stover 2012[34] ; [b]Jones et al 2007[36] ;[c]O'Hair 1995[27], [d]Bouchoux 2011[35]. *Deprotonated side carboxyl

In Table 6 we report the gas-phase basicities and proton affinities for aminoacids, adenosine and βD-Glucose as well as the standard enthalpies and Gibbs free energies of formation of the protonated metabolites. The latter quantities were obtained from the GBs and PAs recommended in the literature and the corresponding quantities for the neutral metabolites from Table 4. We observe that our calculations conform with literature data in general. The obtained formation quantities are in most cases positive but no obvious conclusions can be drawn out of their values

Table 6: Estimated standard thermochemical quantities (ideal gas phase) of ionic metabolites from corresponding properties for neutral metabolites (Table 4) and their recommended basicities (GB) and proton affinities (PA)a. Values in parenthesis are our GB and PA calculations at the same level of theory as in Table 4. All quantities are given in kJmol^{-1}.

Metabolite	**GB[a]**	**$\Delta_f G^0$**	**PA[a]**	**$\Delta_f H^0$**
Alanine	868 (872)	347.0	902 (904)	209.0
Arginine Arginine 2*	1007 (954)	389.8 429.6	1046 (985)	99.7 149.1
Asparagine	905	196.6	942	-21.0
Aspartic acid	882	9.2	920	-177.2
Cysteine	870 (874)	370.2	903 (905)	233.5
Glutamic acid	908	-26.0	947	-233.0
Glutamine	935	182.2	975	-68.0
Glycine	854 (854)	358.9	887 (887)	254.3
Histidine	947	475.5	979	264.1
Isoleucine	885 (884)	348.3	919 (921)	120.5
Leucine	883 (885)	349.4	916 (920)	121.8
Lysine	952 (948)	355.1	994 (990)	80.2
Methionine	899 (910)	366.2	938 (945)	167.5
Phenylalanine	892 (896)	478.9	930 (929)	279.5
Proline	908	383.2	942	200.0
Serine	878 (876)	201.5 205.5	912 (908)	39.5
Threonine	886 (885)	187.5	919 (917)	-6.0
Tryptophan	909	574.2	945	341.1
Tyrosine	895	329.6	933	111.0
Valine	881 (887)	349.8	915 (920)	145.4
Adenosine	945[b]	445	979[b]	72.4
βD-Glucose	786.6[c]	-111.1	810.3[c]	-392.2

[a]Bouchoux 2012[35], [b]Bouchoux 2008[33], [c]Jebber 1996[37]. *Main chain –NH2

Table 7 summarizes the derived formation quantities, together with some formation quantities from the literature and shows a good consistency between our calculations and literature values. The formation quantities for all three types of species (neutral, protonated, deprotonated) follow similar trends. This is better visualized in figures 9 and 10. Figure 9 shows that the free energies of formation of protonated and deprotonated aminoacids vary linearly with the free energy of formation of the neutral aminoacids. The parameters of the straight lines are given in Table 8. The two lines in figure 9 are almost parallel with a slope close to one. The deviations from the straight line are mainly due to aminoacids having strong protonation sites other than the common -NH2 group in the main chain. As expected, the best example is arginine (Arg1 in figure 9) with the strong basic guanidine site. We also indicate in figure the free energy of formation of arginine protonated on the common –NH2 group (Arg2 in figure 9), which is now closer to the linear approximation.

Table 7: Thermochemical quantities of neutral, deprotonated, and protonated aminoacids. Literature values[35] are shown in parenthesis. All quantities are given in kJmol^{-1}.

Metabolite	$\Delta_f G^0$			$\Delta_f H^0$		
	neutral	deprotonated	protonated	neutral	deprotonated	protonated
Alanine	-298.0	-411	347.0	-421.9 (-423.1)	-524 (-528.6)	209.0 (209.5)
Arginine	-129.9	-273	389.8	-398.9	-532	99.7
Asparagine	-411.4	-572	196.6	-612.0	-760	-21.0
Aspartic acid	-621.8	-830	9.23	-790.2	-991	-177.2
Cysteine	-272.8	-411	370.2	-396.5	-525	233.5
Glutamic acid	-631.0	-812	-26.0	-819.0	-997	-233.0
Glutamine	-395.8	-558	182.2	-626.0	-777	-68.0
Glycine	-300.1	-409	358.9	-391.7 (-391.6)	-491 (-494.1)	254.3 (256.0)
Histidine	-87.6	-252	475.5	-289.9 (-289.5)[a]	-447 (-451.3)[a]	264.1 (265.1)[a]
Isoleucine	-279.7	-394	348.3	-493.5 (-492.1)	-601 (-606.3)	120.5 (123.9)
Leucine	-280.6	-395	349.4	-495.2 (-494.8)	-602 (-606.8)	121.8 (125.0)
Lysine	-205.9	-329	355.1	-458.8	-576	80.2
Methionine	-248.0	-356	366.2	-427.8	-534	167.5
Phenylalanine	-142.2	-268	478.9	-323.5	-433	279.5
Proline	-221.8	-341	383.2	-391.1 (-385.9)	-499 (-497.3)	200.0 (207.4)
Serine	-433.5	-583	201.5	-581.5	-722	39.5
Threonine	-439.5	-592	187.5	-620.0	-763	-6.0
Tryptophan	-29.8	-155	574.2	-246.9	-358	341.1
Tyrosine	-288.4	-415	329.6	-489.0	-602	111.0
Valine	-282.2	-398	349.8	-472.6 (-473.4)	-577 (-586.7)	145.4 (146.7)

[a]Bouchoux Riffet 2013[35, 65].

Similar comments can be made for the enthalpy of formation of protonated and deprotonated aminoacids (figure 10). Table 7 and figures 9 and 10 are useful for rationalizing acidities and basicities of aminoacids and their derivatives. It is also worth mentioning that one of the other protonated species (adenosine) has its formation quantities close to the straight lines in these figures. The other protonated species (βD-glucose), however, deviates from the lines to some extent. This diverse behavior requires some explanation.

Equation 7 may be rewritten as follows:

$$GB\ (M) = \Delta_f G^0(M) + \Delta_f G^0(H^+) - \Delta_f G^0(MH^+) \quad (7a)$$

Or

$$\Delta_f G^0(MH^+) = \Delta_f G^0(M) + \Delta_f G^0(H^+) - GB\ (M) \quad (11)$$

If R is the reference compound of the isodesmic reaction, which is often used to replace for the $\Delta_f G^0(H^+)$, equation (11) can be rewritten as:

$$\Delta_f G^0(RH^+) = \Delta_f G^0(R) + \Delta_f G^0(H^+) - GB\ (R) \quad (12)$$

Table 8: Parameters of linear fits in Figures 9 and 10

Protonated Aminoacids				
	Free energy of formation (Fig. 9)		Enthalpy of formation (Fig. 10)	
R^2	0.953		0.943	
Intercept	589.0	SD = 15.8	583.7	SD = 27.9
Slope	0.925	SD = 0.047	0.976	SD = 0.055
Deprotonated Aminoacids				
R^2	0.984		0.986	
Intercept	-105.3	SD = 10.8	-64.5	SD = 15.7
Slope	1.107	SD = 0.032	1.131	SD = 0.031

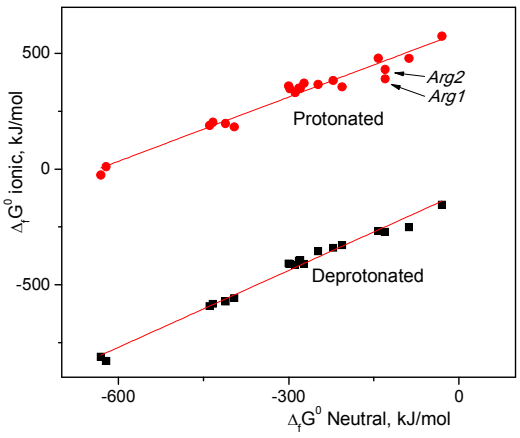


Figure 9: Formation free energy of protonated and deprotonated aminoacids vs. the formation free energy of their neutral counterparts.

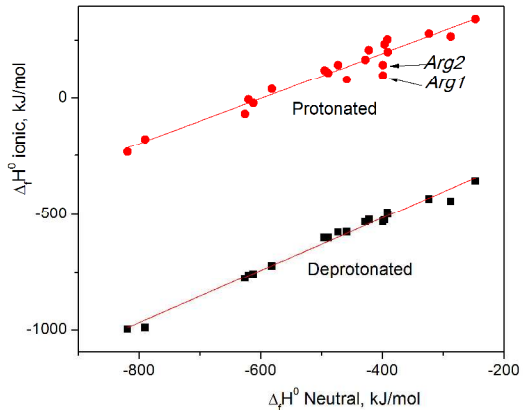


Figure 10: Formation enthalpies of protonated and deprotonated aminoacids vs. the formation enthalpy of their neutral counterparts.

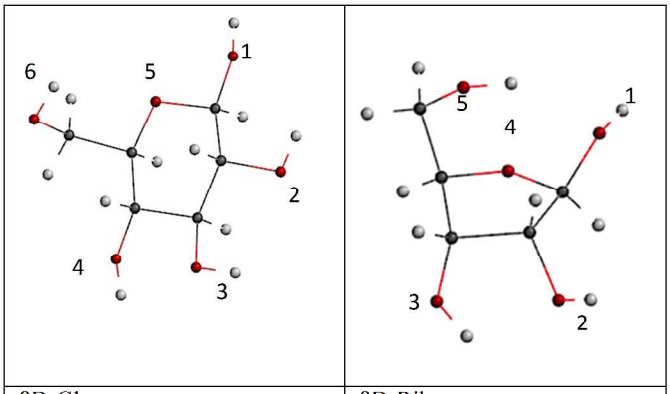
Replacing in equation (11), we obtain:
$\Delta_f G^0(MH^+) = \Delta_f G^0(M) + \Delta_f G^0(RH^+) - \Delta_f G^0(R) + GB\ (R) - GB\ (M)$ (13)

A linear relationship (figure 9) in the form of
$\Delta_f G^0(MH^+) = s\Delta_f G^0(M) + I = s\Delta_f G^0(M) + [\Delta_f G^0(RH^+) - s\Delta_f G^0(R)]$
or
$\Delta_f G^0(MH^+) = \Delta_f G^0(RH^+) + s[\Delta_f G^0(M) - \Delta_f G^0(R)]$ (14)
implies that:
$s\Delta_f G^0(M) = \Delta_f G^0(M) + s\Delta_f G^0(R) - \Delta_f G^0(R) + GB\ (R) - GB\ (M)$ (15)
which can be restated as:
$(s-1)\Delta_f G^0(M) = (s-1)\Delta_f G^0(R) + GB\ (R) - GB\ (M)$ (16)
i.e.,

$GB\ (M) = GB\ (R) + (s-1)[\Delta_f G^0(R) - \Delta_f G^0(M)]$ (17)

Equation 14 or equivalently equation 17 implies that the constant (intercept) term of the above linear relationship is dictated by the corresponding thermochemical quantities of the reference compound. Amines are often used to anchor gas phase basicities of aminoacids while for saccharides a more appropriate reference compound would be an alcohol. Proton affinities and gas-phase acidities can be treated in a similar way.

Table 9 demonstrates our calculations for the thermochemical quantities of protonated and deprotonated βD-glucose and βD-ribose along with data for reference compounds. The numbering of saccharide oxygens is done according to figure 11. The geometries of studied ionic glucose isomers are shown in figure 12.

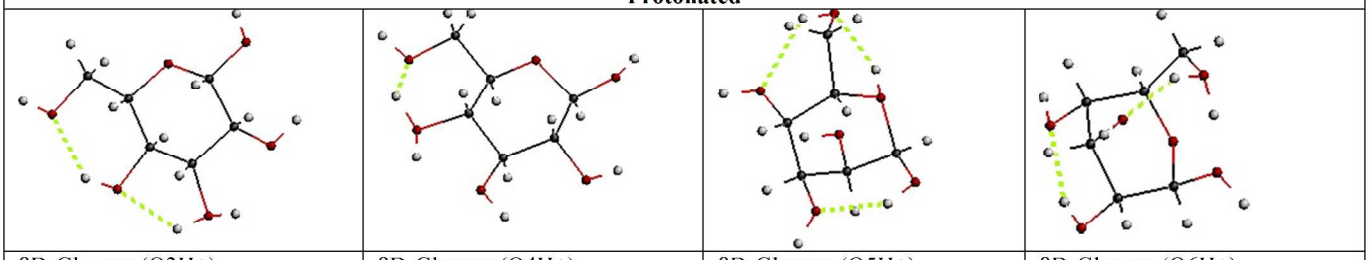


βD-Glucose	βD-Ribose

Figure11: Oxygen numbering in βD-glucose and βD-ribose

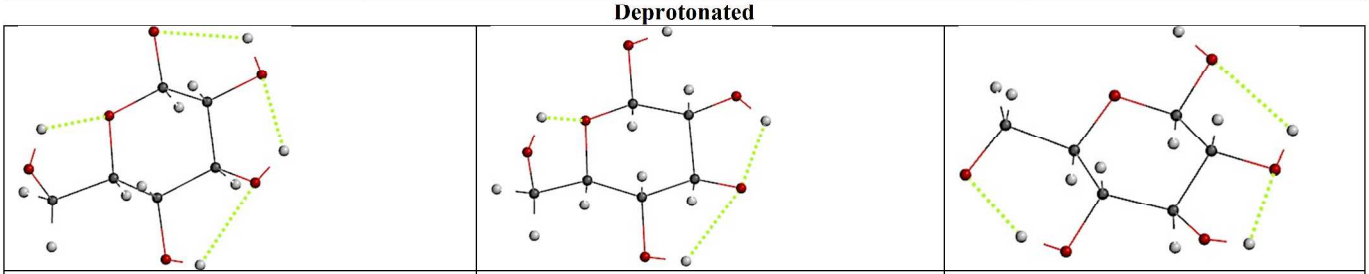
In all cases, the hydroxymethyl rotamers were allowed to take the most stable conformation regardless of the specific (TG, GT, GG) geometry of the corresponding neutral anomer. These more stable structures lead to higher GB and PA values compared to those reported by Jebber et al[37]. As observed in Table 9, the oxygens on the hydroxymethyl groups lead to the most stable protonated structures in both glucose and ribose. O_3 protonation in both sugars lead to the least stable structures. The ring oxygens appear to be relatively favored protonation sites in both sugars. On the other hand, O_1 is the most favored deprotonation site.

Protonated



βD-Glucose (O3H+)	βD-Glucose (O4H+)	βD-Glucose (O5H+)	βD-Glucose (O6H+)

Deprotonated



βD-Glucose (O1-H-)	βD-Glucose (O3-H-)	βD-Glucose (O6-H-)

Figure12: Geometries of the studied ions of βD-glucose anomer

The nucleobases possess sites of varying basicity / acidity as well. Data for adenine protonation is reported in Table 9 as an example. As shown, the stability of protonated adenines may differ by nearly 30 kJ/mol, depending on the (N) protonation site. Nucleosides, then, possess a multitude of protonation / deprotonation sites from their ribose and nucleobase constituents. We will come back to these features in a subsequent Part of this series of works where we will discuss them in relation to metabolite hydration. We may now return to the linearities of figures 9 and 10. Figures 13 and 14 are analogous to figures 9 and 10 with the difference of enhanced scales in order to accommodate the reference compounds as well. Parameters and statistics of the linear fittings are shown in Table 10. As can be observed, there is now much better linear fit in all cases. Classes of compounds anchored to a given reference fall on a straight line passing through reference compounds. Sugars, as an example, fall on the lines passing through alcohols. Aminoacids, amines, and nucleobases fall on nearly the same lines. The lines in each graph are almost parallel. This feature is particularly useful for a quick qualitative estimation of the formation quantities of ions or, equivalently, the gas-phase acidities and basicities from the formation quantities of neutral counterparts. We emphasize the point that these lines are for a qualitative estimation of gas phase acidities and basicities. As pointed out above, the very same neutral metabolite may exhibit different acid/base strength depending on its protonation/deprotonation sites. Thus, the observed scatter in the figures was expected. Nevertheless, the straight lines of figures 13 and 14 are useful tools in discussing gas-phase acidities and basicities.

Table 9: Standard thermochemical quantities (ideal gas) of ionic saccharides and auxiliary compounds. All quantities are given in $kJmol^{-1}$.

Deprotonated				
Compound	$\Delta_f G$	$\Delta_r G$ (ΔG_{acid})	$\Delta_f H$	$\Delta_r H$ (ΔH_{acid})
βD-glucose (O1-H-)	-937.1	1413.4	-1199.9	1440.9
βD-glucose (O3-H-)	-912.5	1438.0	-1174.9	1466.0
βD-glucose (O6-H-)	-894.8	1455.7	-1163.1	1483.8
βD-ribose (O1-H-)	-794.8	1400.3	-1009.3	1419.6
βD-ribose (O2-H-)	-766.1	1429.0	-973.9	1454.9
βD-ribose (O3-H-)	-761.6	1433.5	-973.8	1455.1
βD-ribose (O5-H-)	-771.5	1423.6	-991.1	1443.8
CH3OH	-110.3	1565[a]	-141.9	1592[a]
CH3CH2OH	-129.9	1551[a]	-189	1579[a]
CH3CHOHCH3	-116	1543[a]	-234.7	1571[a]
C5H11OH	-122	1537[a]	-267.7	1564[a]
CH3OCH2CH2OH	-297	1535[a]	-403	1564[a]
CH3COOH	-461.6	1426	-509.3	1456
Protonated				
	$\Delta_f G$	GB	$\Delta_f H$	PA
βD-glucose (O1H+)	-139.3	814.8	-420.3	838.5
βD-glucose (O2H+)	-61.9	737.4	-346.7	764.8
βD-glucose (O3H+)	-95.8	771.3	-385.0	803.1
βD-glucose (O4H+)	-130.2	805.7	-421.7	839.8
βD-glucose (O5H+)	-144.6	820.1	-434.6	852.7
βD-glucose (O6H+)	-156.8	832.4	-454.2	872.3
βD-ribose (O2H+)	12.8	818.1	-223.3	853.4
βD-ribose (O3H+)	58.3	772.6	-181.0	811.1
βD-ribose (O4H+)	31.4	799.5	-200.0	830.1
βD-ribose (O5H+)	12.8	818.1	-223.3	853.4
CH3OH	626.2	724.5[b]	577.8	754.3[b]
CH3CH2OH	599.2	746[b]	521.7	776.4[b]
CH3OCH3	636	764.2[b]	557.1	792[b]
CH3CHOHCH3	576.9	762.6[b]	467.3	793[b]
C6H5OH	694.1	786.3	631.9	817.3
CH3NH2	680.6	864.5[b]	614.6	899[b]
Pyridine	805.4	898.1[b]	743.1	930[b]
Adenine-N1	951.9	910.0 912.5[b]	816.7	942.0 942.8[b]
Adenine-N3	956.8	904.9	821.5	937.2 936[c]
Adenine-N7	979.2	882.5	845.7	912.9 908[c]
Guanine	769.9	927.6[b]	589.5	959.5[b]
Hypoxanthine	805.3	880.5[b]	658.7	912.3[b]
Thymine	470.7	850[b]	314.1	880.9[b]

[a]Lias et al. 1988[77] ; [b]Hunter and Lias 1998[29]; [c]Touboul et al 2008[38]

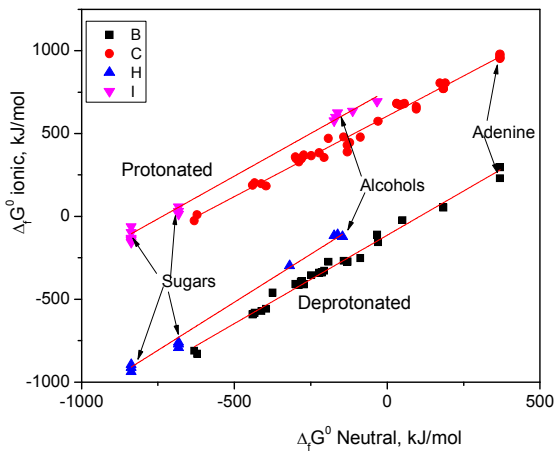


Figure 13: Formation free energies of protonated and deprotonated compounds (including aminoacids) vs. the formation enthalpy of their neutral counterparts. The lines are linear fits (parameters and statistics are shown in Table 10).

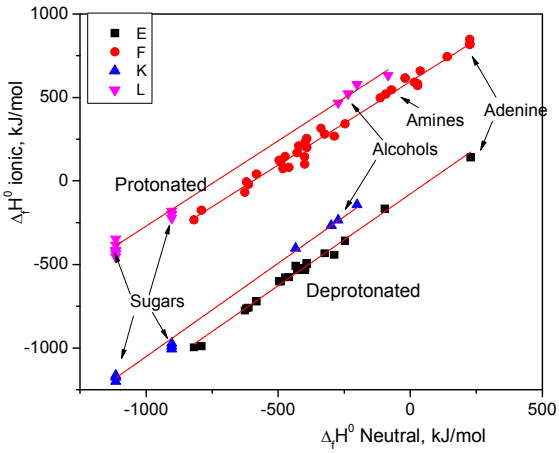


Figure 14: Formation enthalpies of protonated and deprotonated compounds vs. the formation enthalpy of their neutral counterparts. The lines are linear fits (parameters and statistics are shown in Table 10).

Table 10: Parameters of linear fits in Figures 13 and 14

Protonated Aminoacids and References				
	$\Delta_f G^0$ (Line C, Fig. 13)		$\Delta_f H^0$ (Line F, Fig. 14)	
R^2	0.983		0.983	
Intercept	605.4	SD = 6.1	594.6	SD = 9.48
Slope	0.971	SD = 0.022	0.999	SD = 0.023
Deprotonated Aminoacids and References				
	$\Delta_f G^0$ (Line B, Fig. 13)		$\Delta_f H^0$ (Line E, Fig. 14)	
R^2	0.988		0.989	
Intercept	-113.5	SD = 7.0	--77.8	SD = 11.3
Slope	1.067	SD = 0.022	1.098	SD = 0.024
Protonated Sugars and References				
	$\Delta_f G^0$ (Line I, Fig. 13)		$\Delta_f H^0$ (Line L, Fig. 14)	
R^2	0.993		0.996	
Intercept	759.0	SD = 13.7	750.9	SD = 17.8
Slope	1.031	SD = 0.038	1.021	SD = 0.033
Deprotonated Sugars and References				
	$\Delta_f G^0$ (Line H, Fig. 13)		$\Delta_f H^0$ (Line K, Fig. 14)	
R^2	0.988		0.992	
Intercept	62.4	SD = 25.7	65.9	SD = 27.3
Slope	1.160	SD = 0.042	1.117	SD = 0.034

Physical Chemistry Chemical Physics

4. CONCLUSIONS

The primary objective of the present work was the reliable estimation of the heats and Gibbs free energies of formation in the ideal gas state for the key metabolites, such as the aminoacids, oligosaccharides and nucleosides. The adopted route of quantum chemical calculations was through the study of the appropriate isodesmic reactions of metabolites. These isodesmic reactions are based on the thermochemical data for the other reactants and products and therefore, we used the most reliable values for these thermochemical quantities of the auxiliary reactants and products. This leads to a sufficiently reliable set of predicted standard thermochemical quantities of the studied metabolites. Despite the scarcity of available reliable experimental data and the dispersion of available theoretical calculations, our calculations are consistent with the more recent and reliable quantum thermochemical calculations in literature. In the case of aminoacids, thermochemical quantities were also estimated for their protonated and deprotonated ions. Furthermore, we provided a first comprehensive thermochemical table together with the accompanying figures encompassing neutral, protonated and deprotonated aminoacids, βD-glucose, βD-ribose and adenosine including all 20 aminoacids. The observed linear trends may have broader implications that deserve further systematic studies. Calculations of the present work will be very useful in subsequent GCM and hydration works.

Acknowledgements

We want to thank Professor Stefan Grimme, Dr Warren Hehre and Dr Sean Ohlinger for the valuable comments and discussions. NH and MA were supported by the Swiss National Science Foundation. VH were supported by funding from Ecole Polytechnique Fédéral de Lausanne (EPFL) and SystemsX.ch

References

1. T. Geyer, *Current Opinion in Structural Biology*, 2013, **23**, 218-223.
2. N. D. Price, J. A. Papin, C. H. Schilling, and B. O. Palsson, *Trends Biotechnol.*, 2003, **21**, 162–169.
3. C. S. Henry, M. D. Jankowski, L. J. Broadbelt, and V. Hatzimanikatis., *Biophys. J.*, 2006, **90**, 1453-1461.
4. A. M. Feist, C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, and L. J. B. P. D. Karp, V. Hatzimanikatis, and B. Ø. Palsson, *Mol. Syst. Biol.*, 2007, **3**, 1–18.
5. A. Kummel, S. Panke, and M. Heinemann, *BMC Bioinformatics*, 2006, **7**, 1–12.
6. D. A. Beard, and H. Qian., *Am. J. Physiol. Endocrinol. Metab.*, 2005, **288**, E633–E644.
7. A. Kummel, S. Panke, and M. Heinemann, *Mol. Syst. Biol.*, 2006, **2**, 1–10.
8. C. S. Henry, L. J. Broadbelt and V. Hatzimanikatis, *Biophysical journal*, 2007, **92**, 1792-1805.
9. R. N. Goldberg, Y. B. Tewari, and T. N. Bhat, *Bioinformatics*, 2004, **20**, 2874–2877.
10. A. Jinich, D. Rappoport, I. Dunn, B. Sanchez-Lengeling, R. Olivares-Amaya, E. Noor, A. B. Even and A. Aspuru-Guzik, *Scientific reports*, 2014, **4**, 7022.
11. M. D. Jankowski, Henry, C.S., Broadbelt, L.J., and Hatzimanikatis, V., *Biophys J*, 2008, **95**, 1487-1499.
12. Y. D. Cai, T. Huang, L. Chen and B. Niu, *Biomed Research International*, 2013.
13. N. Cohen, *J Phys Chem Ref Data*, 1996, **25**, 1411-1481.
14. E. S. H. Domalski, E. D., *J Phys Chem Ref Data*, 1993, **22**, 805-1159.
15. E. V. Sagadeev, A. A. Gimadeev, D. V. Chachkov and V. P. Barabanov, *Russian Journal of General Chemistry*, 2009, **79**, 1490-1493.
16. V. van Speybroeck, R. Gani and R. J. Meier, *Chemical Society Reviews*, 2010, **39**, 1764-1779.
17. L. A. R. Curtiss, P. C.; Raghavachari, K., *J. Chem. Phys.*, 2001, **114**, 108–117.
18. K. C. Raghavachari, L. A., ed., *Quantum-Chemical Methods for Accurate Theoretical Thermochemistry*, Kluwer, New York, 2002.
19. L. A. Curtiss, P. C. Redfern and K. Raghavachari, *Chemical Physics Letters*, 2010, **499**, 168-172.
20. O. V. Dorofeeva, O. N. Ryzhova and M. A. Suntsova, *J Phys Chem A*, 2013, **117**, 6835-6845.
21. L. V. V. Gurvich, I. V.; Alcock, C. B., *Thermodynamic Properties of IndiVidual Substances*, Hemisphere, New York, 1989 (vol. I); 1990 (vol. II).
22. B. Pedley, ed., *Thermochemical Data and Structures of Organic Compounds*, TRC, College Station, TX, 1994.
23. P. J. Linstrom and W. G. Mallard, *J Chem Eng Data*, 2001, **46**, 1059-1063.
24. B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszevski, S. J. Bittner, S. G. Nijsure, K. A. Amin, M. Minkoff and A. F. Wagner, *J Phys Chem A*, 2004, **108**, 9979-9997.
25. A. B. a. B. R. Elke Goos, 2014.
26. G. H. Thomson, *J Chem Eng Data*, 2000, **45**, 145-145.
27. S. Gronert and R. A. J. Ohair, *J Am Chem Soc*, 1995, **117**, 2071-2081.
28. S. Gronert, D. C. Simpson and K. M. Conner, *J Am Soc Mass Spectr*, 2009, **20**, 2116-2123.
29. L. S. Hunter EP, *J Phys Chem Ref Data* 1998, **27**, 413–656.
30. A. G. Harrison, *Mass Spectrom Rev*, 1997, **16**, 201-217.
31. I. S. Hahn and C. Wesdemiotis, *Int J Mass Spectrom*, 2003, **222**, 465-479.
32. L. Drahos, C. Peltz and K. Vekey, *J Mass Spectrom*, 2004, **39**, 1016-1024.
33. G. Bouchoux, *Mass Spectrom Rev*, 2007, **26**, 775-835.
34. M. L. Stover, V. E. Jackson, M. H. Matus, M. A. Adams, C. J. Cassady and D. A. Dixon, *J Phys Chem B*, 2012, **116**, 2905-2916.
35. G. Bouchoux, *Mass Spectrom Rev*, 2012, **31**, 391-435.
36. C. M. Jones, M. Bernier, E. Carson, K. E. Colyer, R. Metz, A. Pawlow, E. D. Wischow, I. Webb, E. J. Andriole and J. C. Poutsma, *Int J Mass Spectrom*, 2007, **267**, 54-62.
37. K. A. Jebber, K. Zhang, C. J. Cassady and A. ChungPhillips, *J Am Chem Soc*, 1996, **118**, 10515-10524.
38. D. Touboul, G. Bouchoux and R. Zenobi, *J Phys Chem B*, 2008, **112**, 11716-11725.
39. A. Chung-Phillips and Y. Y. Chen, *J Phys Chem A*, 1999, **103**, 953-964.
40. W. M. C. Sameera and D. A. Pantazis, *J Chem Theory Comput*, 2012, **8**, 2630-2645.
41. H. B. Mayes, L. J. Broadbelt and G. T. Beckham, *J Am Chem Soc*, 2014, **136**, 1008-1022.
42. A. Hocquet, N. Leulliot and M. Ghomi, *J Phys Chem B*, 2000, **104**, 4560-4568.
43. M. L. Mavrovouniotis, *J. Biol. Chem.*, 1991, **266**, 14440–14445.
44. J. P. Amend and H. C. Helgeson, *Geochimica Et Cosmochimica Acta*, 1997, **61**, 11-46.
45. A. V. Plyasunov and E. L. Shock, *Geochimica Et Cosmochimica Acta*, 2000, **64**, 439-468.
46. R. A. Alberty, *Thermodynamics of Biochemical Reactions.*, Massachusetts Institute of Technology Press, Cambridge, MA., 2003.
47. J. C. M. Scholten, J. C. Murrell, and D. P. Kelly, *Arch. Microbiol*, 2003, **179**, 135–144.
48. M. Tanaka, Okuno, Y., Yamada, T., Goto, S., Uemura, S. and Kanehisa, M., *Genome Inform*, 2003, **14**, 370–371.
49. C. Li, J. A. Ionita, C. S. Henry, M. D. Jankowski, V. Hatzimanikatis, and a. L. J. Broadbelt, *Chem. Eng. Sci*, 2004, **59**, 5051–5060.
50. Y. B. Tewari and R. N. Goldberg, *Journal of Chemical Thermodynamics*, 2005, **37**, 720-728.
51. D. E. a. H. LaRowe, H. C., *Geochimica Cosmochimica Acta* 2006, **70**, 4680–4724.
52. A. a. D. Picard, I., *Biophys. Chem.*, 2013, **183**, 30–41.
53. J. Rak, P. Skurski, J. Simons and M. Gutowski, *J Am Chem Soc*, 2001, **123**, 11695-11707.
54. S. L. Ling, W. B. Yu, Z. J. Huang, Z. J. Lin, M. Haranczyk and M. Gutowski, *J Phys Chem A*, 2006, **110**, 12282-12291.
55. S. Grimme, S. Ehrlich and L. Goerigk, *J Comput Chem*, 2011, **32**, 1456-1465.
56. M. Von Arnim and R. Ahlrichs, *J Comput Chem*, 1998, **19**, 1746-1757.
57. K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor Chem Acc*, 1997, **97**, 119-124.
58. C. T. Lee, W. T. Yang and R. G. Parr, *Phys Rev B*, 1988, **37**, 785-789.
59. A. D. Becke, *J Chem Phys*, 1993, **98**, 5648-5652.

60. A. D. Becke, *J Chem Phys*, 1997, **107**, 8554-8560.
61. S. Grimme, *J Comput Chem*, 2006, **27**, 1787-1799.
62. J. D. Chai and M. Head-Gordon, *Phys Chem Chem Phys*, 2008, **10**, 6615-6620.
63. J. B. O. a. J. Boerio-Goates, *Chemical Thermodynamics: Principles and Applications*, 2000.
64. D. Missopolinou and C. Panayiotou, *J Phys Chem A*, 1998, **102**, 3574-3581.
65. V. Riffet and G. Bouchoux, *Phys Chem Chem Phys*, 2013, **15**, 6097-6106.
66. S. A. B. Trygubenko, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slavicek, P.; Hobza, P., *Phys. Chem. Chem. Phys.* , 2002, **4**, 4192–4203.
67. M. R. Hanus, F.; Kabelac, M.; Kubar, T.; Bogdan, T. V.; Trygubenko, S. A.; Hobza, P., *J. Am. Chem. Soc.*, 2003, **125**, 7678–7688.
68. O. V. Dorofeeva and O. N. Ryzhova, *Journal of Chemical Thermodynamics*, 2009, **41**, 433-438.
69. I. M. Alecu, J. J. Zheng, Y. Zhao and D. G. Truhlar, *J Chem Theory Comput*, 2010, **6**, 2872-2887.
70. E. V. Sagadeev, A. A. Gimadeev and V. P. Barabanov, *Russian Journal of Physical Chemistry A*, 2010, **84**, 209-214.
71. S. N. Ngauv, R. Sabbah and M. Laffitte, *Thermochimica Acta*, 1977, **20**, 371-380.
72. O. V. Dorofeeva, I. N. Kolesnikova, I. I. Marochkin and O. N. Ryzhova, *Structural Chemistry*, 2011, **22**, 1303-1314.
73. M. W. J. Chase, *NIST-JANAF Thermochemical Tables, 4th ed.*, American Institute of Physics, Woodbury, NY, 1998.
74. R. Sabbah and M. Laffitte, *Bulletin De La Societe Chimique De France Partie I-Physicochimie Des Systemes Liquides Electrochimie Catalyse Genie Chimique*, 1978, I50-I52.
75. I. Contineanu, A. Neacsu, R. Zgirian, S. Tanasescu and S. Perisanu, *Thermochim Acta*, 2012, **537**, 31-35.
76. A. F. Santos and M. A. Ribeiro da Silva, *The journal of physical chemistry. A*, 2009, **113**, 9741-9750.
77. S. G. Lias, J. E. Bartmess, J. F. Liebman, J. L. Holmes, R. D. Levin and W. G. Mallard, *J Phys Chem Ref Data*, 1988, **17**, 1-861.