Accurate and Computationally Efficient Prediction of Thermochemical Properties of Biomolecules Using the Generalized Connectivity-Based Hierarchy

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JULY 2014

Impact Factor: 3.3 · DOI: 10.1021/jp505544y · Source: PubMed

3 AUTHORS, INCLUDING:

Arkajyoti Sengupta
Indiana University Bloomington
8 PUBLICATIONS **33** CITATIONS

SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

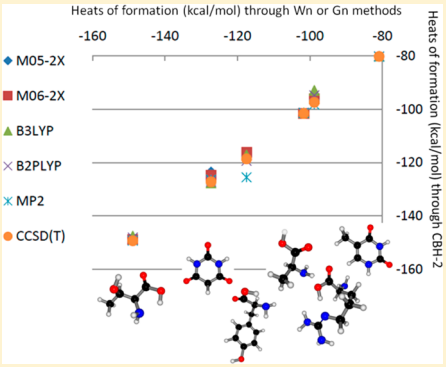
Accurate and Computationally Efficient Prediction of Thermochemical Properties of Biomolecules Using the Generalized Connectivity-Based Hierarchy

Arkajyoti Sengupta, Raghunath O. Ramabhadran, and Krishnan Raghavachari*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States

Supporting Information

ABSTRACT: In this study we have used the connectivity-based hierarchy (CBH) method to derive accurate heats of formation of a range of biomolecules, 18 amino acids and 10 barbituric acid/uracil derivatives. The hierarchy is based on the connectivity of the different atoms in a large molecule. It results in error-cancellation reaction schemes that are automated, general, and can be readily used for a broad range of organic molecules and biomolecules. Herein, we first locate stable conformational and tautomeric forms of these biomolecules using an accurate level of theory (viz. CCSD(T)/6-311++G(3df,2p)). Subsequently, the heats of formation of the amino acids are evaluated using the CBH-1 and CBH-2 schemes and routinely employed density functionals or wave function-based methods. The calculated heats of formation obtained herein using modest levels of theory and are in very good agreement with those obtained using more expensive W1-F12 and W2-F12 methods on amino acids and G3 results on barbituric acid derivatives. Overall, the present study (a) highlights the small effect of including multiple conformers in determining the heats of formation of biomolecules and (b) in concurrence with previous CBH studies, proves that use of the more effective error-cancelling isoatomic scheme (CBH-2) results in more accurate heats of formation with modestly sized basis sets along with common density functionals or wave function-based methods.



INTRODUCTION

Amino acids, uracils, and barbituric acid (along with its derivatives) are immensely relevant in biochemistry. As the building blocks of peptides and proteins, amino acids have been studied extensively to understand the structures of the polymeric proteins.[1,2] Similarly, uracil and its 5-methyl derivative (commonly known as thymine), being the primary components of nucleic acids (in RNA and DNA), have been studied to understand the structure and properties of nucleic acids.[3] Furthermore, barbituric acid and its derivatives find use in pharmaceutical applications as sedatives, hypnotics, anticonvulsants, and adjuncts in anesthesia.[4−6] Thus, a detailed knowledge of these monomers is important to understanding the three-dimensional structure and biochemistry in their polymeric forms.

The significant conformational flexibility and the existence of different tautomeric forms of these biomolecules necessitate a detailed structural and conformational analysis to accurately evaluate their properties such as vertical ionization energies,[7] gas-phase heats of formation,[8−10] and gas-phase acidities.[11−13] In particular, the presence of intramolecular hydrogen bonding interactions contributes to several low-lying energy conformers and different tautomeric forms.

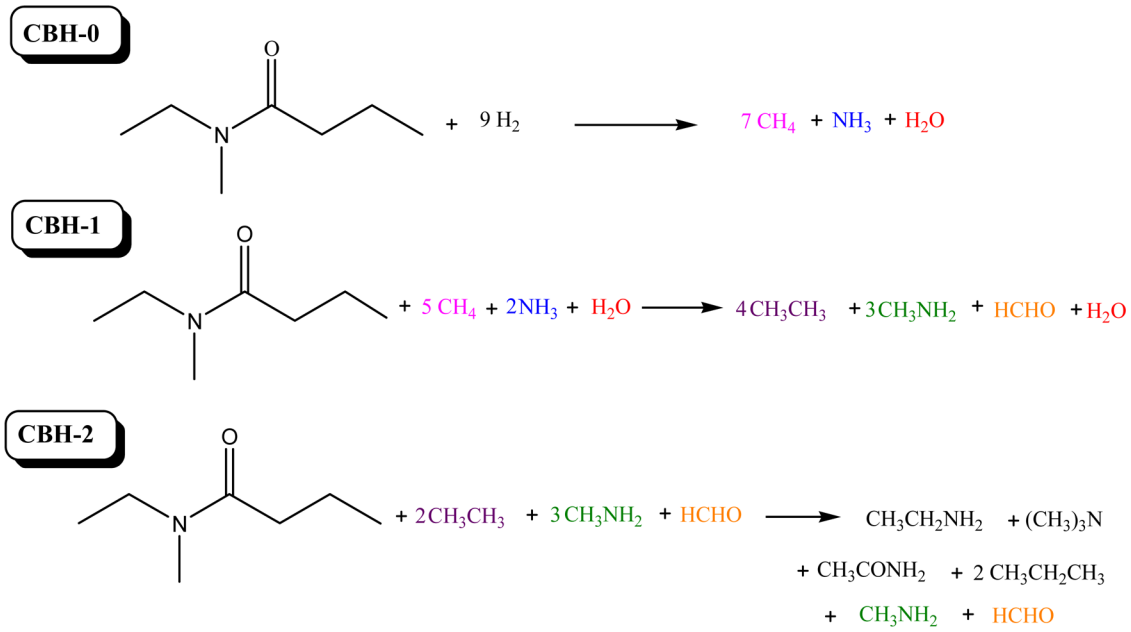
However, despite the importance of these biomolecules, reliable experimental thermochemical studies in the literature are scarce. Recent calorimetric studies by Roux and co-workers have reported the enthalpies of formation of alanine,[8] cysteine,[9] methionine,[10] and derivatives of uracil and barbituric acid in the gas phase.[14−18] Additionally, Sagadeev et al. have predicted the enthalpies of formation of the amino acids using a parametrized group additivity method.[19] However, while accurate composite schemes such as Gaussian-n (G_n) theory,[20−24] Weizmann-n (W_n) theory,[25−28] the ATOMIC protocol,[29−31] the HEAT protocol,[32−34] CCSDT/CBS,[35−38] and so forth can be used to derive heats of formation which are close to the experimentally observed values, the demanding computational expense of these methods has confined their use to relatively small molecules. Very recently, Karton et al. have calculated heats of formation of amino acids using W1-F12, W2-F12, and G4 methods.[39] The W1-F12 and W2-F12 calculations include scalar relativistic components, spin−orbit coupling, and a diagonal Born−Oppenheimer correction (DBOC). While these calculations offer accurate benchmark values, particularly in the absence of experimental heats of formation, such calculations can be highly demanding computationally. With increasing system size, such calculations can be computationally unachievable as reported in the cases of tyrosine and tryptophan by Karton et al.[39]

Alternatively, chemical reaction schemes[40,41] based on systematic error cancellation can be used to derive accurate heats of formation. For instance, Roux et al. and Stover et al. have used iosdesmic reaction schemes for barbituric acid and uracil

Received: June 4, 2014
Revised: July 24, 2014
Published: July 26, 2014

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

Scheme 1. Reactions Corresponding to the Different Rungs in the CBH Method for an Amide Molecule[a]



[a]The colored fragments represent their recursive presence on the different CBH rungs. The complete construction of the error cancellation scheme can be found in ref 43.

derivatives[14−18] and for amino acids,[12] respectively. They then performed G3 and G3MP2 calculations for the biomolecules involved in their reaction schemes, and thus obtained accurate heats of formations. Furthermore, during the preparation of this article, we came across the recent work of Dorofeeva et al., in which G4 calculations were used in conjunction with specific isodesmic reactions developed for each of the amino acids.[42] The reported heats of formation in their work are the averages of the values evaluated by the different reaction schemes for each of the amino acids. However, such reaction schemes are system-specific and require a manual and arduous matching of the bond and hybridization types. Thus, the lack of generalized reaction schemes that are system-independent had previously precluded a systematic development of a thermochemical hierarchy applicable to all organic molecules and biomolecules.[41]

The recently developed connectivity-based hierarchy (CBH) has provided an efficient solution to this problem and predicts accurate heats of formation for a diverse collection of organic molecules.[43] The structure-based hierarchy increasingly balances the bond types and hybridization of the various heavy (non-hydrogen) atoms in organic molecules. Accurate results are obtained with several different density functionals[44] as well as wave function-based methods, such as MP2 and CCSD(T), with modest basis sets.[45] Recently in an exploratory study, we applied CBH to two sulfur-containing amino acids, namely, cysteine and methionine. These sulfur-containing amino acids, with the maximum different number of heavy atoms among naturally occurring amino acids, were chosen to test the efficacy of our error-cancellation schemes.[46] The results are in excellent agreement with the recently reported more expensive W1-F12 and G4 calculations by Stover et al.,[12] Roux et al.,[9,10] Karton et al.,[39] and Dorofeeva et al.[42] In a further demonstration of CBH's applicability to biomolecules in this study, we use the isoatomic scheme (CBH-2)[47] to derive accurate enthalpies of formation for 28 different biologically relevant molecules containing multiple functional groups.

A succinct description of the construction of CBH is given here by the example of *N*-ethyl-*N*-methylbutanamide in Scheme 1.[48] CBH has alternating atom-centric and bond-centric rungs, called CBH-0, -1, -2, -3, and so forth. The lowest rung, called CBH-0, corresponds to what is traditionally known as the isogyric scheme,[40] where each heavy atom is preserved in its valence state by adding the required number of hydrogen atoms. The CBH-1 rung is formally the same as the popular isodesmic bond separation scheme.[40] On the CBH-2 rung (which is appropriately called the isoatomic scheme), the immediate chemical bonding environment of all heavy atoms is preserved.

In general, with the increase in the number of CBH rungs, the size of the fragments generated also increases. This results in the buildup of a larger part of the molecule, providing better error cancellation and resulting in accurate heats of formation. However, with an increase in the size of fragments, the experimental heats of formation of these larger fragments are not always known. Also, the computational expense of larger fragments (on higher rungs) increases. Thus, in conjunction with our previous results, for molecules of the size being considered in this study, we find that the isoatomic rung (CBH-2) is an effective scheme offering an excellent balance between cost and accuracy for the size of molecules studied herein and provides us fragments with experimentally known enthalpies of formation. Hence, while CBH-3 and other higher rungs can be appropriately defined, for larger molecules we limit our discussion in this study to the CBH-1 and CBH-2 rungs.

The first step in the derivation of thermodynamic properties using CBH (or other schemes) requires carrying out careful structural analysis to get correct molecular geometries and conformations. Prior to presenting our approach toward getting the right geometries and conformations, it is useful to provide a brief review of the current experimental and theoretical work in this area.

The conformational studies of the biomolecules considered in this work have been of interest to both experimental and

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

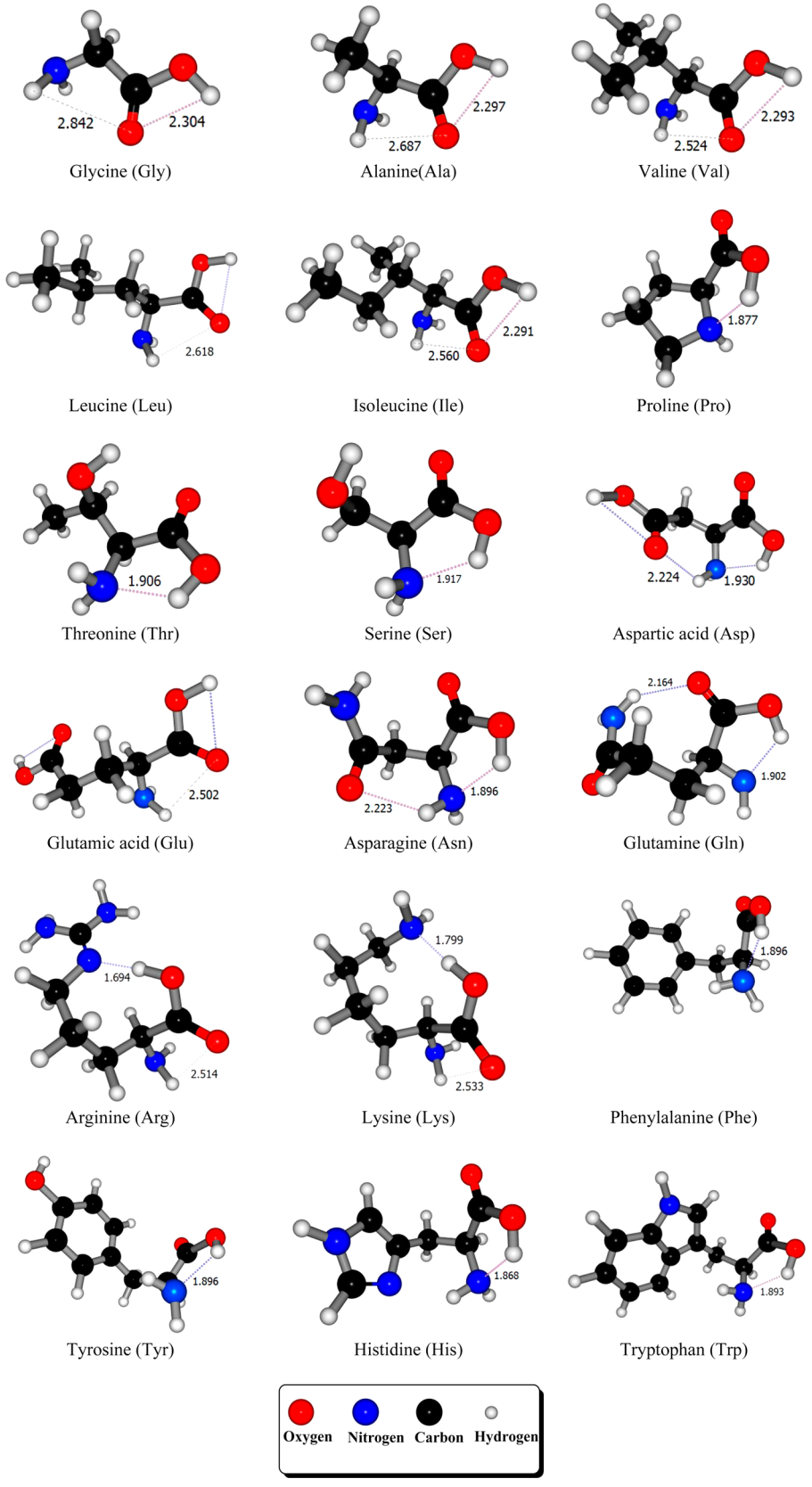


Figure 1. B3LYP/6-311++G(d,p)-optimized structures of the most stable conformers of the amino acids considered in the present study. The dominant hydrogen-bonding interactions occurring in the structures are marked in the structures. Distances between pairs of atoms are in angstroms.

theoretical chemists for decades.[49−51] Various experimental techniques have been developed to investigate the different stable conformers of glycine and alanine.[52−56] However, the vaporization of amino acids beyond alanine, via classical heating methods, results in decomposition and hence restricts gas-phase investigations. This resulted in a series of conformational

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

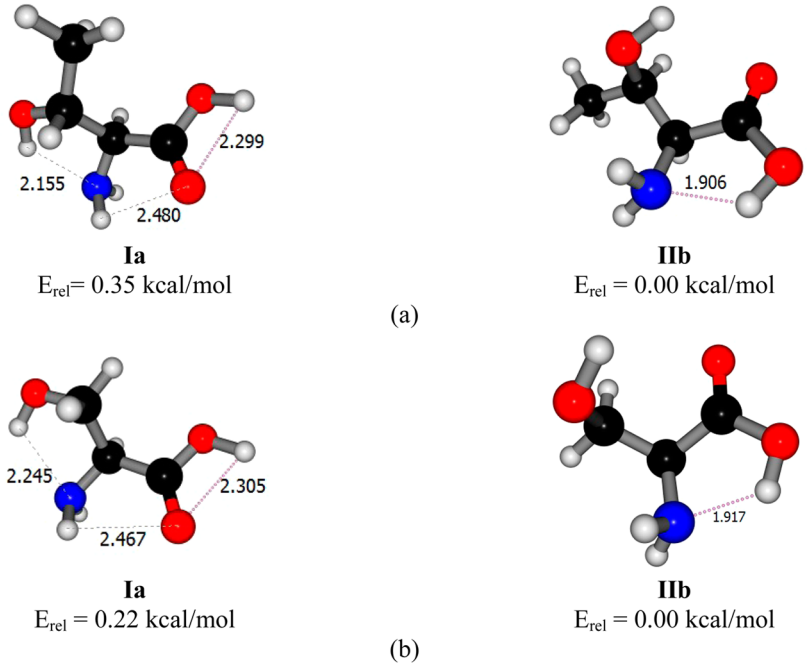


Figure 2. Optimized structures of two close-lying conformers of (a) threonine and (b) serine along with their E_{rel} at B3LYP/6-311++G(d,p).

analyses studied with the aid of computations.[57−59] Despite these advances, the global minima for many of the amino acids are still debatable due to the presence of several low-lying conformers. Moreover, the order of the low-lying conformers may depend on the level of theory used for optimization.[60−63] The comparison of experimental results with such studies where different levels of theory and basis sets are employed for optimization often helps in reaching a recommended level of theory to get the right conformational ordering. A recent experimental technique which combines laser ablation with Fourier transform microwave spectroscopy has been successful in observing multiple conformers of different amino acids.[64] The technique has been used to explore the conformations of over 15 amino acids including proline,[65] asparagine,[66] aspartic acid,[67] and glutamic acid.[68] The rotational constants of the observed conformers were found to be in agreement with the calculated values.

Similarly, barbituric acid contains four enolizable hydrogen atoms resulting in various tautomeric forms where the hydrogens are bound to either nitrogen, carbon, or oxygen atoms. Different ab initio and DFT studies on barbituric acid and its derivatives concur that the triketo form is substantially more stable than the other tautomeric forms, in agreement with the known experimental observations.[69−74] Thus, we performed our calculations on these previously reported lowest-energy structures.

COMPUTATIONAL DETAILS

All of the calculations have been performed using the Gaussian 09[75] suite of programs. Geometry optimizations were carried out at the B3LYP/6-311++G(d,p) level for the biomolecules. In order to obtain accurate energy ordering of the different low-lying energy conformers of the amino acids, single-point calculations at CCSD(T)/6-311++G(3df,2p) were performed. The CCSD(T)/6-311++G(3df,2p) energies were obtained using an additivity-based approach as shown in the following paragraph. The harmonic frequencies were scaled by a factor of 0.9854. The optimized geometries were all characterized as minima without any imaginary frequencies. Geometry optimizations were followed by single-point calculations with the 6-311++G(3df,2p) basis set using four different density functional methods—B3LYP,[76,77] M05-2X,[78] M06-2X,[79] and B2PLYP[80]—having different exchange and correlation energy functionals. Single-point calculations with the same basis set were also made at the MP2 and CCSD(T) levels to evaluate the performance of CBH for the biomolecules considered in this study using wave function-based methods.

The single-point energies obtained using these different methods were added to the thermal corrections to the enthalpies at 298.15 K to determine the heats of formation. Consistent with our idea of balancing the computational cost of CBH with the desired level of accuracy, computationally demanding CCSD(T)/6-311++G(3df,2p) calculations were avoided. Instead, we used an additivity-based approach to get the CCSD(T)/6-311++G(3df,2p) energies:

$$\begin{aligned}\mathrm{E(CCSD(T)/6\text{-}311{+}{+}G(3df,2p))} &= \mathrm{E(CCSD(T)/6\text{-}31{+}G(d,p))} + \mathrm{E(MP2/6\text{-}311{+}{+}G(3df,2p))} \\ &\quad - \mathrm{E(MP2/6\text{-}31{+}G(d,p))}\end{aligned}$$

RESULTS AND DISCUSSION

This section has been broadly divided into (1) structural analysis, to obtain the structures and low-energy conformations to obtain accurate thermochemistry, and (2) an evaluation of accurate heats of formation with CBH-1 and CBH-2 schemes using the structures obtained in this study.

Structural Analysis. Similar to our previous study,[46] here we restrict our analysis to the low-energy conformers and neglect conformers or tautomers having a population of less than 5%. This is based on the assumption that the conformers with a population of less than 5% contribute very little to the overall enthalpy of formation.

Amino Acids. The geometry optimizations of the various conformers of the 18 non-sulfur-containing amino acids have

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

been carried out at B3LYP/6-311++G(d,p). The different conformers of each amino acids were ordered according to the energies obtained from our calculations at CCSD(T)/6-311++G(3df,2p)//B3LYP/6-311++G(d,p). Since exhaustive conformational analyses of different amino acids have been carried out previously,[49−68] to avoid redundancy we just report the main findings in the article, and structures of the low-lying energy conformers are given in the Supporting Information. We have followed the nomenclature of the different conformers proposed by Alonso and co-workers in their studies,[65−68] a brief description of which is given in the Supporting Information.

The lowest-energy conformers for the different amino acids are shown in Figure 1. We observe for both glycine and alanine N−H···O═C interactions, resulting in conformer **Ia** being the lowest-energy conformer for both amino acids. The structures obtained in this study were compared with those in the previous studies and are in good agreement.[51,55,57,81,82] Only two other conformers of glycine and three other conformers of alanine were considered for further study as the remaining conformers had a population of less than 5%. With amino acids valine and leucine, the increase in the number of methyl groups and degrees of freedom led to an increase in the number of low-lying energy conformers.[83] Six of the 8 low-lying energy conformers of valine and 6 of the 13 different low-energy conformers of leucine were considered since the remaining conformers had a population of less than 5%. Conformer **Ia1** for both valine and leucine corresponds to the lowest-energy conformer shown in Figure 1, while the remaining low-energy conformers considered in the study are represented in Figure S1 in the Supporting Information. The lowest-energy conformer for valine has similar orientations of $-NH_2$ and −COOH groups as in glycine and alanine as shown in Figure 1. The lowest-energy conformers located for valine, leucine, and isoleucine are in agreement with the previous experimental and theoretical results.[84−87] Eight additional low-energy conformers have been considered in the study of the evaluation of the heat of formation of isoleucine. The four most stable conformers of proline having the maximum abundance have been considered in the study. Conformer **IIa** was identified as the global minimum and is in agreement with previous ab initio and experimental results.[88]

Serine and threonine are the two nonaromatic naturally occurring amino acids that have an −OH group. The addition of this −OH group results in multiple low-energy conformers with interesting intramolecular hydrogen-bonding possibilities. The occurrence of −OH···O═C, −OH···NH_2, and O═C−OH···NH_2 interactions, as shown in Figure 2, causes two conformers to lie close to each other on the energy scale. The conformational analysis of threonine gave **IIb** as the most stable conformer followed by conformer **Ia** (E_{rel} = 0.35 kcal/mol) in agreement with previous studies of Alonso et al.[89] For serine, we found that the lowest-energy conformer is **IIb** followed by conformer **Ia** (E_{rel} = 0.22 kcal/mol). Previous studies by Miao et al.[13] and Alonso et al.[90] reported conformer **Ia** to be the most stable conformer of serine. It is clear that the identification of the global minimum is a challenging problem in such cases where the energy differences are small. However, since all of the low-energy isomers are included in our analysis, accurate heats of formation can be obtained. Five low-energy conformers each of threonine and serine have been considered in the study.

For acidic and basic amino acids with additional −COOH or $-NH_2$ groups, we observe an increase in the number of O═C−OH···NH_2 and NH_2···O═C interactions that results in multiple low-energy conformers. The six most stable conformers each of aspartic acid and glutamic acid were considered for the study since the remaining ones had a population of less than 5%. For aspartic acid, conformer **IIb-I** was found to be the lowest-energy conformer followed by conformer **IIa-I** (E_{rel} = 0.39 kcal/mol). However, conformer **Ia-I**, for which we obtain E_{rel} = 0.77 kcal/mol, is reported to be the lowest-energy conformer by Alonso and co-workers.[67] Thus, we note that conformational analysis at a rigorous level of theory is important in identifing the most stable minima. The conformational analysis of glutamic acid resulted in conformer **Iagc1** as the most stable conformer, which is in agreement with previous calculations and experimental observations.[68]

Asparagine and glutamine are the two amino acids with $-CONH_2$ as the functional group in the side chain of the amino acids. In asparagine, the O═C−OH···NH_2 interactions resulted in the most stable conformer **IIa**. This result (in Figure 1) is in agreement with previous studies.[66,91] A similar orientation of glutamine with the O═C−OH···NH_2 interactions resulted in conformer **IIGGc1a**. The conformational analysis of glutamine has not been reported previously, and hence a systematic study was done on glutamine. Twenty-seven different minima were located for glutamine, but only the three most stable conformers of glutamine were considered for further study since the remaining conformers had a population of less than 5%.

The lowest-energy conformers for the two basic nonaromatic amino acids, lysine and arginine, were also identified, and the lowest-energy conformer and two other stable conformers were identified for lysine. The lowest-energy conformer is in agreement with the structure used in the recent study of Dorofeeva et al.[42] However, this result is in disagreement with B3LYP and MP2 optimizations by Boeckx et al.[60] The difference between the two low-energy conformers (Figure S1) using CCSD(T)/6-311++G(3df,2p)//B3LYP/6-311++G(d,p) calculations was found to be 0.25 kcal/mol. Thus, again we note that the identification of the global minimum is a challenging problem, and it is important to include all of the low-energy conformations. The presence of NH_2 and NH groups in arginine results in multiple stable conformers where intramolecular hydrogen-bonding interactions occur. Only three lower-energy conformers among the nine different minima were considered for the study since the remaining conformers had a population of less than 5%. The optimized lowest-energy conformer (shown in Figure 1) is in agreement with the CCSD(T) calculations based on MP2-optimized geometries by Ling et al.[63]

Aromatic amino acids phenylalanine, tyrosine, histidine, and tryptophan were also studied. Six different minima were located for phenylalanine, and conformer **IIa** (shown in Figure 1) was identified as the lowest-energy conformer. Conformer **IIa** was followed by **Ia** (E_{rel} = 0.51 kcal/mol) and **IIb** (E_{rel} = 0.60 kcal/mol). The remaining conformers had a population of less than 5% and were hence not considered here in the study. In tyrosine, the rotation about the phenolic −OH in addition to the $-NH_2$ and −COOH functional groups results in several conformers. Eight different low-energy conformers were located for tyrosine, and six of them were considered for further study. The lowest-energy conformer was found to be in agreement with the previous conformational studies on tyrosine.[92,93] Histidine can be present in two tautomeric forms, namely, the π and τ forms. A total of 25 different conformers possessing 2 different tautomeric forms were studied, and it was found that the most stable conformer had the τ tautomeric form in agreement with previous studies.[94] The O═C−OH···NH_2 and N_{His}···NH_2 interactions are present in the most stable conformers as shown in Figure 1.

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

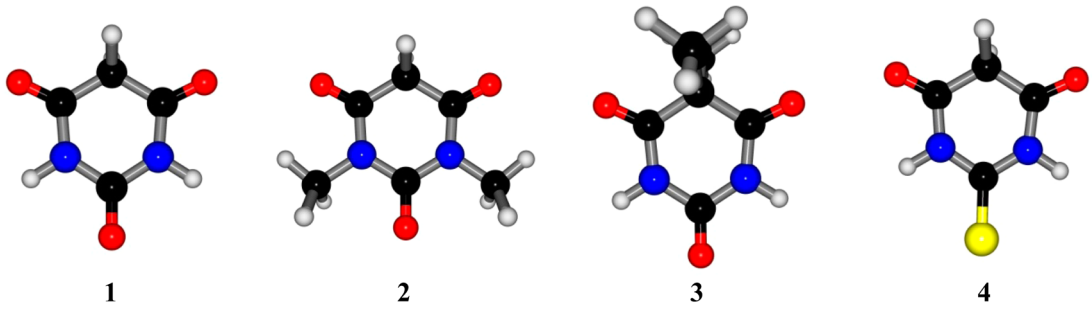


Figure 3. B3LYP/6-311++G(d,p)-optimized geometries of barbituric acid (**1**), 1,3-dimethylbarbituric acid (**2**), 5,5-dimethylbarbituric acid (**3**), and 2-thiobarbituric acid (**4**).

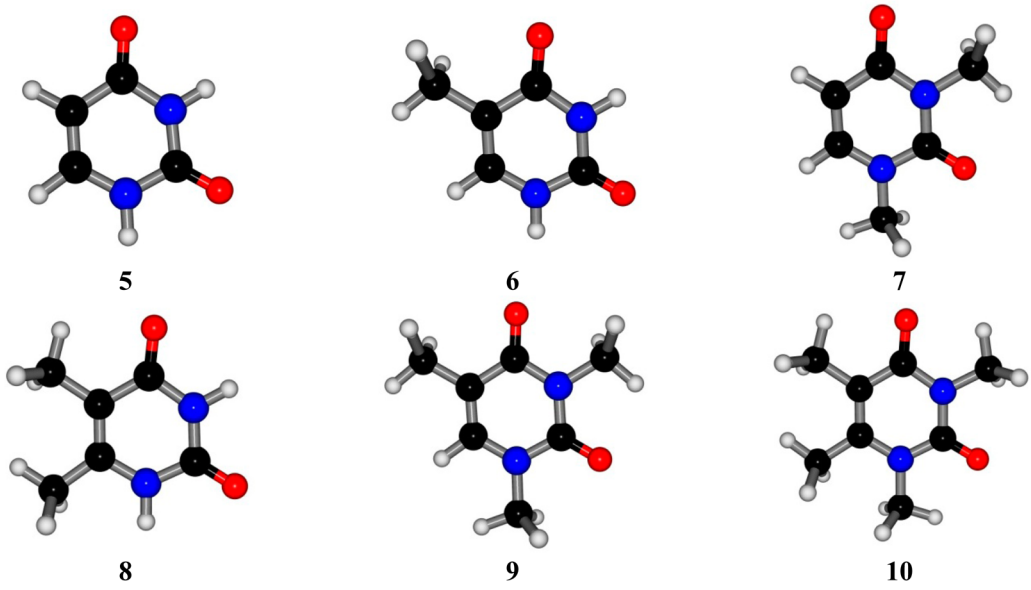


Figure 4. B3LYP/6-311++G(d,p)-optimized geometries of uracil (**5**), thymine (**6**), 1,3-dimethyluracil (**7**), 5,6-dimethyluracil (**8**), 1,3,5-trimethyluracil (**9**), and 1,3,5,6-tetramethyluracil (**10**).

Two conformers each of the π tautomeric form and τ tautomeric form were considered for the study since the remaining conformers had a population of less than 5%. Ten different conformers of tryptophan were located, of which four low-energy conformers were considered for further study.

The presence of the O═C−OH···NH_2 interaction results in **IIa** being the most stable conformer. It turns out that this interaction is an important interaction for all of the aromatic amino acids as the resulting conformer **IIa** is the most stable conformer for all four aromatic amino acids—phenylalanine, tyrosine, histidine, and tryptophan—as shown in Figure 1.

Barbituric Acid and Its Derivatives. The presence of the enolizable hydrogen atoms in barbituric acid and its derivatives results in various tautomeric forms and hence requires a detailed investigation. Figure 3 represents the B3LYP/6-311++G(d,p)-optimized geometries of barbituric acid and its derivatives (**1**−**4**). It was found that the triketo tautomer of barbituric acid (**1** in Figure 3) is the most stable form followed by the monohydroxyl tautomers as found in the previous studies.[69−71] The ΔG value for the monohydroxy tautomer was calculated to be 14.9 kcal/mol relative to the triketo form. This suggests that in the gas phase, barbituric acid will be present only in the triketo form. The optimized structure is in agreement with the MP2(Full)/6-31G(3df,2p)-optimized structure of barbituric acid.[14] It is fully planar and possesses C_{2v} symmetry as shown in Figure 3. 1,3-Dimethylbarbituric acid (**2**) contains one enolizable hydrogen atom and thus exists in two tautomeric forms. The keto form was calculated to be more stable than the enol form by 11.6 kcal/mol. This is in agreement with the calculations made by Bertolasi et al.[72] 5,5-Dimethylbarbituric acid (**3**) has two enolizable hydrogen atoms and can be present in four tautomeric forms. It was calculated that the keto form is the most stable form, followed by the monohydroxy tautomers. The 2,4-diketo-6-hydroxy and 4,6-diketo-2-hydroxy tautomers are less stable by 19.9 and 15.4 kcal/mol, respectively. Hence only the keto tautomer was considered for the study in the gas phase for species **1**−**3** shown in Figure 3. 2-Thiobarbituric acid (**4**) has 4 enolizable hydrogen atoms that lead to 10 tautomeric forms. Previous studies suggest that the 4,6-diketo-2-thione tautomer is the most stable one, and the high-energy differences from the other tautomers suggest that, in the gas phase, 2-thiobarbituric acid will be primarily present only in the 4,6-diketo-2-thione tautomeric form.[74] Hence only the optimized structure of the 4,6-diketo-2-thione is used for our gas-phase investigation on the heat of formation of derivative **4**.

Uracil and Its Derivatives. Uracil and five of its derivatives have been considered in this study. The optimized structures of the different uracil derivatives are shown in Figure 4. Uracil and its 5-methyl derivative (thymine) have been studied exhaustively to obtain the structures of the nucleosides in RNA and DNA, respectively. The optimized structures of **5** and **6** were in

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

agreement with their crystal structure determined by X-ray diffraction.[95,96] In both cases, the rings are planar and intermolecular interactions are present in the crystal structure. The optimized structures of 5,6-dimethyluracil, 1,3-dimethyluracil, 1,3,5-trimethyluracil, and 1,3,5,6-tetramethyluracil in this study were similar to the structures reported by Notario et al.[18]

Evaluation of Reaction Energies and Heats of Formation. The CBH-1 and CBH-2 schemes for all of the biomolecules considered in the study are given in Tables S1 and S2 of the Supporting Information. The reaction energies were first evaluated for the different biomolecules on the respective CBH-1 and CBH-2 rungs. These reaction energies and the available experimental heats of formation of the different fragments involved in CBH-1 and CBH-2 schemes of the various reaction schemes were then used to calculate the heats of formation of the various biomolecules.[97]

Reaction Energies. Different density functionals and wave function-based methods are used to calculate the reaction energies obtained with CBH. The reaction energies for the different amino acids at CBH-1 and CBH-2 are given in Tables S4 and S5 in the Supporting Information and are illustrated in Figure 5. The foremost observation is the large drop in the

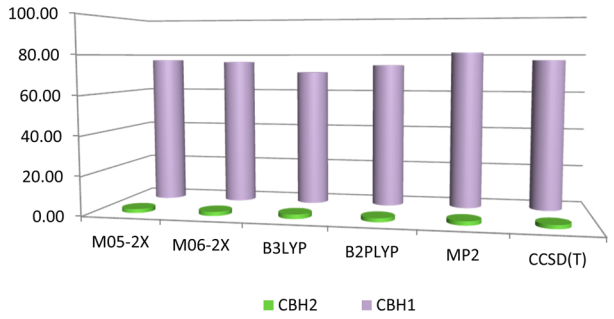


Figure 5. Pictorial representation of the variation of mean absolute reaction energies (*y* axis, kcal/mol) against the different CBH rungs used at different wave function-based as well as DFT methods. The 6-311++G(3df,2p) basis set was used throughout (extrapolated in the case of CCSD(T), see computational details).

reaction energies from CBH-1 to CBH-2 (Figure 5), which is consistent across the different levels of theory employed. This consistency in the calculated reaction energies at CBH-2 indicates that the calculated heats of formation would also be reliably obtained by different DFT or wave function-based levels of theory.[98]

Heats of Formation of the Lowest-Energy Structures of Amino Acids, Uracil, and Barbituric Acid Derivatives. The experimental heats of formation of the various fragments needed on the CBH-1 and CBH-2 rungs were obtained from the NIST Chemistry WebBook and Lias et al.'s thermochemical tables.[97] For methylenimine CH_2NH, the accurate value calculated using W2 theory by Oliveira et al.,[99] ΔH_f° = 21.1 kcal/mol, was used. In cases where the heats of formation of CBH-2 fragments were unknown, we used CCSD(T) calculations on CBH-1 to obtain the missing heats of formation. The CBH-1 reaction schemes for such fragments are provided in Table S8 in the Supporting Information.

The heats of formation for the amino acids evaluated at M06-2X and CCSD(T) levels of theory through CBH-1 and CBH-2 schemes are compared to previously reported experimental and calculated G3MP2 and W1-F12 heats of formation in Table 1. We observe that, despite the large reaction energies, the calculated heats of formation by the CBH-1 scheme are not significantly off. Moreover, for an accurate method such as CCSD(T), the results at CBH-1 and CBH-2 are similar for most of the amino acids. However, for some of them, namely, lysine, arginine, and proline, we observe that the CCSD(T) results differ significantly (~2 kcal/mol) across the CBH rungs. The present discrepancies are due to the presence of strong intramolecular interactions that are well described by larger fragments at CBH-2. Thus, in the presence of strong intramolecular interactions and/or ring strain, the use of CCSD(T) calculations at CBH-2 is recommended over CBH-1. Furthermore, the calculated values at CBH-2 are in very good agreement with the heats of formation evaluated by G3MP2 and W1-F12 calculations by Stover et al.[12] and Karton et al.,[39] respectively. Thus, in the following sections we focus on the more accurate CBH-2 results for the heats of formation and compare them with the available experimental and theoretical values.

The heats of formation of glycine using the CCSD(T) method at CBH-2 is within 0.5 kcal/mol of the experimental values and the value evaluated by Karton et al. through their W1-F12 calculations.[39] For six nonaromatic amino acids, the optimized lowest-energy conformers of the amino acids are in agreement with the structures on which the G3MP2 calculations were performed by Stover et al.[12] The CCSD(T) and M06-2X calculations at CBH-2 with these six amino acids gave mean absolute deviations of 1.5 and 1.1 kcal/mol from the ΔH_f° values evaluated by G3MP2 calculations.[12] Similarly, for the eight nonaromatic amino acids, where the lowest-energy conformers of the amino acids were in agreement with the structures used by Karton et al., the mean absolute deviations of evaluated ΔH_f° values of the different amino acids with CCSD(T) and M06-2X were calculated to be just 0.7 and 0.9 kcal/mol from the ΔH_f° values obtained by W1-F12 calculations.[39] Thus, remarkable accuracy is offered by the isoatomic scheme (CBH-2). In addition, the CBH-2 results are consistent across different levels of theory and are hence relatively independent of the levels of theory used as observed in our previous studies.[43,46]

Furthermore, the calculated ΔH_f° values for the different amino acids are also in agreement with the ones evaluated by G3MP2 and G4 calculations on the isodesmic schemes developed by Stover et al.[12] and Dorofeva et al.[42] While accurate calculations on such molecules having intramolecular hydrogen bonding and dispersion interactions can be performed with G3MP2 and G4 methods using isodesmic schemes, they still remain a challenge for the widely used DFTs. However, our CBH-2 calculations show that the M06-2X and M05-2X functionals can be used to obtain accurate heats of formation of such systems. With B3LYP, the hydrogen bonding interactions were underestimated (Table S5 in the Supporting Information), and hence the heats of formation evaluated by B3LYP calculations show larger differences from our CCSD(T) calculations or the W1-F12 calculations by Karton et al.[39] The B2PLYP results for most of the systems are within 2.0 kcal/mol of the values obtained by CCSD(T) calculations. However, in the future it would be useful to try such functionals in conjunction with DFT-D3,[100] which can improve the DFT results for noncovalent interactions.

The other important challenge in the present study comes from the unavailability of the experimental heats of formation with which to compare our results. Moreover, among the ones that are available there are discrepancies between the values given in the popular NIST Web site and the recent studies for glycine and alanine. We and others have pointed out such differences in

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

Table 1. Heats of Formation (kcal/mol) for the Lowest-Energy Conformers of the Amino Acids Calculated at M06-2X and CCSD(T),[a] with the values Compared to Calculated G3MP2[b] and Experimental Values

				CBH-1		CBH-2	
amino acid	expt[c]	W1-F12[g]	G3MP2 TAE[b]	M06-2X	CCSD(T)	M06-2X	CCSD(T)
Gly	−93.3 ± 1.1	−94.2	−91.9	−94.8	−94.1	−94.0	−93.8
	−94.2 ± 0.4[d]						
Ala	−99.1 ± 1.0	−101.7	−100.7	−101.7	−102.0	−101.6	−101.6
	−102.0 ± 0.7[e]						
	−101.6 ± 0.5[d]						
Pro	−87.5 ± 1.0	[h]	[k]	−89.9	−93.4	−89.4	−91.7
Thr	[f]	−148.9	−147.9	−148.2	−149.7	−148.8	−149.3
Ser	[f]	[h]	[k]	−139.5	−140.1	−139.6	−139.7
Leu	−116.4 ± 1.2	−118.8	−118.1	−117.1	−119.8	−117.9	−119.0
Ile	[f]	[h]	[k]	−116.6	−119.2	−117.3	−118.5
Val	−108.8 ± 1.0	−113.6	−112.9	−112.7	−114.6	−113.2	−113.9
Asp	[f]	[h]	[k]	−189.6	−190.0	−189.5	−190.2
Glu	[f]	−195.5	[k]	−194.6	−195.5	−194.7	−195.7
Asn	[f]	−146.5	−144.4	−146.4	−146.3	−146.6	−147.5
Gln	[f]	[h]	[k]	−151.0	−151.3	−151.3	−152.3
Arg	[f]	−98.8	[k]	−98.6	−99.2	−96.1	−97.4
Lys	[f]	[h]	[k]	−110.1	−112.4	−109.8	−110.8
Phe	[f]	−76.9	−77.1	−74.6	−76.8	−74.4	−76.6
Tyr	[f]	−117.5[j]	−118.9	−117.0	−118.9	−116.1	−118.6
His(τ)	[f]	−69.8	−65.8	−67.7	−70.0	−63.7	−69.3
His(π)	[f]	[i]	[k]	−67.9	−69.5	−63.9	−68.9
Trp	[f]	−58.4[j]	−59.6	−56.5	−60.4	−54.2	−59.7

[a]The geometries were obtained at the B3LYP/6-311++G(d,p) level of theory. The 6-311++G(3df,2p) basis set was used for single-point calculations. CCSD(T)/6-311++G(3df,2p) energies were obtained by extrapolation. [b]The calculated ΔH_f° values from the total atomization energies using G3MP2 calculations by Stover et al. in ref 12. [c]The heats of formation were taken from the NIST Chemistry WebBook[97b] unless otherwise indicated. [d]Experimental values from ref 101. [e]From ref 8. [f]Missing experimental values. [g]The ΔH_f° values were calculated with W1-F12 theory by Karton et al.[39] [h]The lowest-energy conformer of amino acids differed from the geometry on which the W1-F12 calculations were performed in ref 39. [i]ΔH_f° value of the His(π) tautomer was not calculated in ref 39. [j]W1-F12 atomization energies were not calculated; instead an average of values evaluated by G4, CBS-QB3, and G4(MP2)-6X levels were provided in ref 39. [k]The lowest-energy conformer of amino acids differed from the geometry on which G3MP2 calculations were made in ref 12.

the past.[39,42,46] Through our calculations and those of others, we recommend the use of −102.0 ± 0.7 kcal/mol reported by da Silva et al.[8] over −99.1 ± 1.0 kcal/mol reported in the popular NIST WebBook as an experimental reference for alanine. The chosen value is not only reported with less uncertainty but is also in agreement with G3 calculations carried out after considering multiple stable conformers of alanine.[8]

It was observed that for proline and valine the evaluated ΔH_f° values have large deviations with respect to the experimentally observed values. The CBH-2 calculations for these amino acids using the CCSD(T) method result in ΔH_f° values that are 4.0−5.0 kcal/mol lower than the experimentally observed values. The M06-2X calculations provide ΔH_f° values similar to CCSD(T) results and are in similar disagreement with the available experimental results. Moreover, our results are in accordance with values obtained by Stover et al.,[12] Dorofeeva et al.,[42] and Karton et al.[39] All of these independent calculations are in disagreement with the corresponding experimental results, and hence we can strongly recommend a revisiting of experiments for ΔH_f° values of proline and valine.

For some of the amino acids we found that the experimental heats of formation of the CBH-2 fragments for the parent molecules were not available with experimental uncertainty. For such a scenario we recommend CCSD(T) calculations with the CBH-1 scheme to internally calibrate and arrive at more accurate heats of formation for some of the fragments.[102] These calculations yielded a value of 13.23 kcal/mol as the heat of formation in the gas phase for vinylamine. Similar calculations were made for 2-aminopropene and *N*-methylmethanimine where the experimental heats of formation were unavailable. The CCSD(T) calculations resulted in a value of 3.6 kcal/mol for 2-aminopropene and 19.05 kcal/mol for *N*-methylmethanimine as heats of formation in the gas phase which are used in the present study.

The evaluated heats of formation for the aromatic amino acids at CBH-2 are in agreement with ΔH_f° values evaluated by G3MP2 calculations by Dixon and co-workers and W1-F12 calculations by Karton et al. (as seen in Table 1). Even the large reaction energies for histidine (τ) and tryptophan at CBH-2 with CCSD(T) resulted in just 0.9 kcal/mol as the mean absolute deviation from the ΔH_f° values evaluated by W1-F12 calculations by Karton et al.[39] However, we observe large differences among the different DFTs and CCSD(T) results. The large discrepancies suggests the deficiency of these DFTs including M06-2X for the aromatic molecules and thus recommend the use of evaluated ΔH_f° by CCSD(T) calculations for such conjugated systems. The MP2 calculations for the aromatic amino acids (Table S10 in Supporting Information) also resulted in greater deviations from the W1-F12 results, and hence we recommend the use of CCSD(T) to obtain accurate heats of formation of the aromatic amino acids. The deficiency in MP2 for other aromatic systems has been noted previously.[45]

The heats of formation of barbituric acid and uracil derivatives **1**−**10** were also evaluated using CBH-1 and CBH-2 schemes.

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

The heats of formation evaluated by M06-2X and CCSD(T) on the various CBH rungs are compared with the available experimental and calculated G3 values by Roux and co-workers in Table 2. It is readily observed that the CCSD(T) calculations

Table 2. Experimental and Calculated Heats of Formation ($\Delta H_f°$) of Barbituric Acid and Uracil Derivatives at CBH-2[a] against $\Delta H_f°$ Values Obtained through G3[b] Calculations

			CBH-2	
species	G3 TAE[b]	experimental	M06-2X	CCSD(T)
1	−127.3	−127.8 ± 0.4[c]	−124.7	−127.2
2	−132.0	−130.9 ± 0.5[d]	−126.5	−129.6
3	−143.3	−141.3 ± 0.5[e]	−139.3	−142.7
4	−67.0	−66.6 ± 0.6[f]	−64.5	−68.3
5	−71.6	−71.4 ± 0.3[g]	−68.2	−71.1
6	−80.9	−76.5 ± 1.0[h], −78.6 ± 1.0[i]	−77.0	−80.2
7	−76.5	−75.0 ± 0.4[j], −76.6 ± 0.5[g]	−70.0	−73.4
8	−90.3	−90.0 ± 0.6[k]	−86.4	−89.6
9	−85.7	−85.1 ± 0.7[k]	−78.7	−82.5
10	−92.4	−91.3 ± 0.7[k]	−85.4	−89.4

[a]The geometries were obtained at the B3LYP/6-311++G(d,p) level of theory. The 6-311++G(3df,2p) basis set was used for single-point calculations. CCSD(T)/6-311++G(3df,2p) energies were obtained by extrapolation. [b]The values were taken from refs14−18 where the G3 calculations were made using an atomization scheme. [c]Taken from ref 14. [d]Taken from ref 15. [e]Taken from ref 16. [f]Taken from ref 17. [g]Taken from ref 103. [h]Taken from ref 104. [i]Taken from ref 105. [j]Taken from ref 106. [k]Taken from ref 18.

at CBH-2 are in agreement with the $\Delta H_f°$ values obtained from G3 calculations by Roux et al.[14−18] The CCSD(T) calculations at CBH-2 yielded a value of −127.2 kcal/mol for the heat of formation for barbituric acid (**1**) against an experimental value of −127.8 kcal/mol.[14] As seen in the cases of the aromatic amino acids, irrespective of the high reaction energies of the barbituric acid derivatives, the heats of formation with CCSD(T) are in excellent agreement with the experimental values as well as G3 calculations using the atomization schemes and the isodesmic schemes devised by Roux et al.[14−17] Similar observations are made for uracil and its derivatives **5**−**10**, where the CCSD(T) results at CBH-2 are in agreement with the experimental values and the G3 calculations on total atomization schemes for the uracil derivatives by Notario et al.[18] Similar to aromatic amino acids, we observe the M06-2X and CCSD(T) results to differ consistently by more than 2 kcal/mol and thus recommend the CCSD(T) results over M06-2X in such scenarios. Our results suggest that the CBH methods offer an excellent alternative to many of the more expensive composite schemes.

Heats of Formation Taking Multiple Conformations into Account. Thus, far, we computed the heats of formation taking only the lowest-energy conformer into account. However, the wide conformational landscape exhibited by these amino acids (vide supra) suggests that for computing more accurate heats of formation, multiple conformations of the amino acids should be considered. Thus, in this section, the contributions of other low-lying energy conformers were included along with the lowest-energy conformer to evaluate the heats of formation at room temperature. Similar to our previous study,[46] we first calculated the individual enthalpies of formation for each conformation (ΔH_i) and then obtain the weighted enthalpy of formation (ΔH) as

$$\Delta H = \sum_{i=1}^{n} X_i \Delta H_i$$

where n is the number of conformations considered.[107] X_i refers to the weight of the ith conformer and is obtained by Boltzmann weighting the relative free energies (referenced to the lowest-energy conformer) as

$$X_i = \left\{ \frac{e^{-\Delta G_i/RT}}{\sum_n^{i=1} e^{-\Delta G_i/RT}} \right\}$$

We considered three stable conformers of glycine to evaluate an accurate heat of formation, which resulted in a value of −93.5 kcal/mol for CCSD(T) at CBH-2 against the experimental value of −93.3 kcal/mol reported in NIST. This value is closer to the experimental value than the one evaluated using just the lowest-energy conformer. A pictorial representation of the effect of inclusion of other stable conformers on the heats of formation of glycine and leucine is shown in Figure 6. The inclusion of other

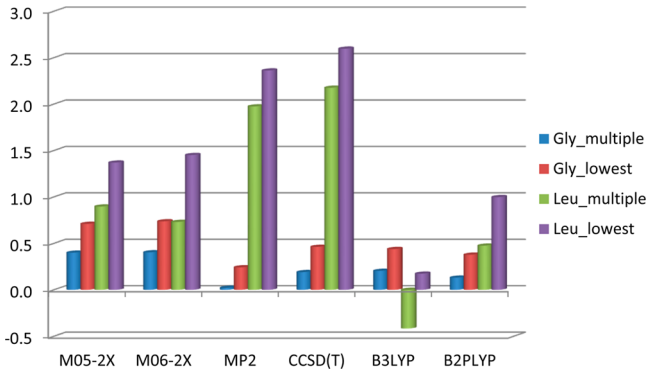


Figure 6. Values of experiment − theory (in kcal/mol) are compared for glycine and leucine when $\Delta H_f°$ is evaluated with the lowest-energy conformer (Gly_lowest and Leu_lowest) against the ones obtained with the inclusion of other low-lying energy conformers (Gly_multiple and Leu_multiple).

stable conformers slightly improves the computed heats of formation with all of the different methods used (as seen from Figure 6). Similar improvements upon including multiple conformations were also noted for the other amino acids. As expected, we observe an increase in $\Delta H_f°$ for proline from Table 1 to Table 3. The CCSD(T) calculations at CBH-2 resulted in $\Delta H_f°$ = −91.4 kcal/mol, which differs from the experimental value by 3.9 kcal/mol against 4.2 kcal/mol when only the lowest-energy conformer is considered. However, the calculated values are still in disagreement with the experimental value, and we recommend the calculated ones over the experimental value.

Similar to proline, in valine a gradual increase from a $\Delta H_f°$ value of −113.9 to −113.5 kcal/mol at CBH-2 using the CCSD(T) method is noted after including the other stable conformers. The presence of the −OH group in serine and threonine results in added intramolecular hydrogen bonding interactions (vide supra) and leads to many low-lying conformers. The inclusion of the other stable conformers results in a marginal increase in the $\Delta H_f°$ values by 0.3 kcal/mol each for threonine and serine at CBH-2.

The heat of formation for aspartic acid was calculated after including six low-energy conformers, and the $\Delta H_f°$ values were found to be −189.9 and −189.4 kcal/mol at CBH-2 with CCSD(T) and M06-2X, respectively. For glutamic acid, again we

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

Table 3. Enthalpies of Formation (kcal/mol) of the Amino Acids Calculated at M06-2X and CCSD(T),[a] Taking Multiple Stable Conformers into Account

			CBH-1		CBH-2	
amino acid	expt[b]	number of low-energy conformers considered[e]	M06-2X	CCSD(T)	M06-2X	CCSD(T)
Gly	−93.3 ± 1.1	3	−94.5	−93.8	−93.7	−93.5
	−94.2 ± 0.4[c]					
Ala	−99.1 ± 1.0	4	−101.3	−101.6	−101.2	−101.3
	−102.0 ± 0.7[d]					
	−101.6 ± 0.5[c]					
Pro	−87.5 ± 1.0	4	−89.6	−93.1	−89.1	−91.4
Thr	[f]	5	−147.8	−149.4	−148.4	−149.0
Ser	[f]	5	−139.2	−139.8	−139.4	−139.4
Leu	−116.4 ± 1.2	6	−116.4	−119.4	−117.1	−118.6
Ile	[f]	9	−116.1	−118.8	−116.9	−118.0
Val	−108.8 ± 1.0	6	−112.2	−114.1	−112.8	−113.5
Asp	[f]	6	−189.4	−189.6	−189.4	−189.9
Glu	[f]	6	−194.4	−195.3	−194.5	−195.4
Asn	[f]	1	−146.4	−146.3	−146.6	−147.5
Gln	[f]	3	−150.7	−151.1	−151.1	−152.1
Arg	[f]	3	−98.4	−98.9	−95.9	−97.1
Lys	[f]	3	−109.9	−112.2	−109.7	−110.6
Phe	[f]	3	−74.4	−76.6	−74.2	−76.3
Tyr	[f]	6	−116.7	−118.5	−115.7	−118.2
His(τ)	[f]	2	−67.6	−69.8	−63.6	−69.2
His(π)	[f]	2	−67.8	−69.4	−63.8	−68.8
Trp	[f]	4	−56.2	−60.1	−53.9	−59.4

[a]The geometries were obtained at the B3LYP/6-311++G(d,p) level of theory. The 6- 311++G(3df,2p) basis set was used throughout for single-point calculations. CCSD(T)/6-311++G(3df,2p) energies were obtained by extrapolation. [b]The heats of formation were taken from the NIST Chemistry WebBook[97b] unless indicated otherwise. [c]Experimental values from ref 101. [d]From ref 8. [e]The remaining conformers of the amino acid were found to have a population of less than 5%. [f]Missing experimental values.

considered six conformers as the remaining conformers had a population of less than 5%. The heats of formation of glutamic acid were calculated to be −195.4 and −194.5 kcal/mol at CBH-2 using CCSD(T) and M06-2X, respectively. The CBH-2 calculations using CCSD(T) after the inclusion of other stable conformers of glutamine, arginine, and lysine resulted in −152.1, −97.1, and −110.6 kcal/mol as their heats of formation.

The heats of formation for the aromatic amino acids were also calculated after the inclusion of other stable conformers. Six different conformations were considered for tyrosine, and the heat of formation of tyrosine was found to be −118.2 kcal/mol vs −118.6 kcal/mol when only the lowest-energy conformer was considered. Different conformations for each π and τ tautomer of histidine were considered for the evaluation of the heats of formation of histidine in the two different tautomeric forms present at room temperature. The CCSD(T) calculations at CBH-2 resulted in values of −68.8 and −69.2 kcal/mol as heats of formation for the π and τ tautomers of histidine after the inclusion of other stable conformers of the respective tautomeric forms.

Overall, the use of the isoatomic scheme (CBH-2) after the inclusion of the other low-lying energy conformers of amino acids results in accurate heats of formation of the different amino acids. A general observation for the different amino acids is that the inclusion of several low-lying energy conformers resulted in heats of formation within 0.6 kcal/mol from the values evaluated when only the lowest-energy conformer was considered. Thus, it should be important to note that the lowest-energy conformation is sufficient to achieve an accuracy of 1 kcal/mol for such biomolecules with a significant conformational flexibility as predicted in our earlier study.[46] Hence, as a benchmark we suggest that calculations with the lowest-energy conformer as in the present study as well by others[12,39,42] can be considered to be appropriate reference values for the heats of formation of the amino acids. However, as pointed out previously, the identification of the global minimum can be challenging in many cases, and it is preferable to include all of the low-energy conformations in such a thermodynamic analysis.

CONCLUSIONS

In this comprehensive study, the generalized connectivity-based hierarchy method[43] has been used on 28 bio-organic molecules to determine their heats of formation in the gas phase at room temperature. The set of biomolecules included 18 amino acids and 10 derivatives of uracil and barbituric acid. The present study is a significant advancement in the application[46] of the CBH method for natural organic molecules containing multiple functional groups and several low-lying conformers. The heats of formation for the biomolecules have been evaluated with CBH-1 (isodesmic bond separation scheme) and CBH-2 (isoatomic) schemes. Similar to previous studies, it is observed that the CBH-2 (isoatomic) scheme, preserving the atom connectivity in a molecule, resulted in more efficient error cancellation than the commonly used isodesmic-bond separation scheme (CBH-1). Conformational analyses of the different amino acids have been performed to obtain the lowest-energy conformer of the amino acids. Different hydrogen-bonding interactions were identified that resulted in several stable conformers. It was observed that the lowest-energy conformer of most amino acids is in agreement with the structures reported in previous studies. However, for serine, lysine, and aspartic acid our lowest-energy conformer is in disagreement with the

dx.doi.org/10.1021/jp505544y | J. Phys. Chem. B 2014, 118, 9631−9643

previous DFT and MP2 studies, though the energy differences between the conformations are small. Overall, we recommend the use of rigorous calculations (such as those employing CCSD(T) with a sufficiently large basis set) to derive accurate energies of the different conformations and accurate thermochemistry of the amino acids.

The heats of formation of barbituric acid and uracil derivatives evaluated by CBH-2 calculations are in agreement with the values obtained from G3 calculations by Roux et al.[14−17] for the different derivatives. Heats of formation of the different amino acids were evaluated (a) first with the lowest-energy conformers and (b) after including several low-lying conformers. In both the cases, the calculated ΔH_f° values were in agreement with the values obtained using G3MP2 and W1-F12 calculations by Stover et al. and Karton et al., respectively.[12,39] In agreement with the previous studies,[12,39,42] our calculated ΔH_f° values differ largely from the available experimental values for proline and valine, and revisions to the experimental values are suggested in these cases. Our calculations further show that the inclusion of other low-energy conformers has only a small effect, and thus including the effects of multiple conformers results in enthalpies of formation within 1 kcal/mol of the values obtained by considering only the lowest-energy conformer. Nevertheless, in light of the difficulties in the identification of the global minima, it is suggested that multiple stable conformers of molecules be taken into account to derive accurate thermochemical parameters.

Finally, this study successfully highlights the use of modest levels of theory that can be readily employed by both experts and nonexperts alike using the generalized CBH method. Thus, it is concluded that calculations with modest basis sets and regular density functionals or wave function-based methods with the CBH-2 scheme are an excellent alternative to the more expensive composite schemes for such biomolecules.

■ ASSOCIATED CONTENT

Supporting Information

Several tables of additional data, the CBH-1 and CBH-2 reaction schemes for all 28 biomolecules studied, and the optimized geometries and Cartesian coordinates of all of the conformers used. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kraghava@indiana.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge support from NSF grant CHE-1266154 at Indiana University.

■ REFERENCES

(1) Greenberg, D. M. *Amino Acids And Proteins: Theory, Methods, Application*; Blackwell Scientific: Oxford, 1951.

(2) Buxbaum, E. *Fundamentals of Protein Structure and Function*; Springer Science: New York, 2007.

(3) Neidle, S. *Principles of Nucleic Acid Structure*; Oxford Press: New York, 1999.

(4) Arzamastsev, A. P.; Luttseva, T. Y.; Sadchikova, N. P. Methods for the Analysis and Standardization of Drugs Belonging to the Class of Barbituric Acid Derivatives. *Pharm. Chem. J.* **2001**, *35*, 453−457.

(5) Levine, B. *Principles of Forensic Toxicology*, 2nd ed.; AACC Press: Washington, DC, 2003; p 173.

(6) Willow, M.; Johnston, G. A. R. Pharmacology of Barbiturates: Electrophysiological and Neurochemical Studies. *Int. Rev. Neurobiol.* **1983**, *24*, 15−49.

(7) Close, D. M. Calculated Vertical Ionization Energies of the Common α-Amino Acids in the Gas Phase and in Solution. *J. Phys. Chem. A* **2011**, *115*, 2900−2912.

(8) Ribeiro da Silva, M. A. V.; Ribeiro da Silva, M. D. D. M. C.; Santos, A. F. L. O. M.; Roux, M. V.; Foces, C. F.; Notario, R.; Mejía, R. G.; Juaristi, E. Experimental and Computational Thermochemical Study of α-Alanine (DL) and β-Alanine. *J. Phys. Chem. B* **2010**, *114*, 16471−16480.

(9) Roux, M. V.; Foces, C. F.; Notario, R.; Ribeiro da Silva, M. A. V.; Ribeiro da Silva, M. D. M. C.; Santos, A. F. L. O. M.; Juaristi, E. Experimental and Computational Thermochemical Study of Sulfur-Containing Amino Acids: L-Cysteine, L-Cystine, and L-Cysteine-Derived Radicals. S−S, S−H, and C−S Bond Dissociation Enthalpies. *J. Phys. Chem. B* **2010**, *114*, 10530−10540.

(10) Roux, M. V.; Notario, R.; Segura, M.; Chickos, J. S.; Liebman, J. F. The Enthalpy of Formation of Methionine Revisited. *J. Phys. Org. Chem.* **2012**, *25*, 916−924.

(11) Riffet, V.; Frison, G.; Bouchoux, G. Acid-Base Thermochemistry of Gaseous Oxygen and Sulfur Substituted Amino Acids (Ser, Thr, Cys, Met). *Phys. Chem. Chem. Phys.* **2011**, *13*, 18561−18580.

(12) Stover, M. L.; Jackson, V. E.; Matus, M. H.; Adams, M. A.; Cassady, C. J.; Dixon, D. A. Fundamental Thermochemical Properties of Amino Acids: Gas-Phase and Aqueous Acidities and Gas-Phase Heats of Formation. *J. Phys. Chem. B* **2012**, *116*, 2905−2916.

(13) Miao, R.; Jin, C.; Yang, G.; Hong, J.; Zhao, C.; Zhu, L. Comprehensive Density Functional Theory Study on Serine and Related Ions in Gas Phase: Conformations, Gas Phase Basicities and Acidities. *J. Phys. Chem. A* **2005**, *109*, 2340−2349.

(14) Roux, M. V.; Temprado, M.; Notario, R.; Foces, C. F.; Emelyanenko, V. N.; Verevkin, S. P. Structure-Energy Relationship in Barbituric Acid: A Calorimetric, Computational, and Crystallographic Study. *J. Phys. Chem. A* **2008**, *112*, 7455−7465.

(15) Roux, M. V.; Notario, R.; Foces, C. F.; Temprado, M.; Ros, F.; Emelyanenko, V. N.; Verevkin, S. P. Experimental and Computational Thermochemical Study of Barbituric Acids: Structure−Energy Relationship in 1,3-Dimethylbarbituric Acid. *J. Phys. Chem. A* **2011**, *115*, 3167−3173.

(16) Roux, M. V.; Notario, R.; Foces, C. F.; Temprado, M.; Ros, F.; Emelyanenko, V. N.; Verevkin, S. P. Experimental and Computational Thermochemical Study and Solid-Phase Structure of 5,5-Dimethylbarbituric Acid. *J. Phys. Chem. A* **2010**, *114*, 3583−3590.

(17) Roux, M. V.; Notario, R. Experimental and Computational Thermochemical Study of 2-Thiobarbituric Acid: Structure−Energy Relationship. *J. Phys. Chem. A* **2012**, *116*, 4639−4645.

(18) Notario, R.; Emelyanenko, V. N.; Roux, M. V.; Ros, F.; Verevkin, S. P.; Chickos, J. S.; Liebman, J. F. Thermochemistry of Uracils. Experimental and Computational Enthalpies of Formation of 5,6-Dimethyl-, 1,3,5-Trimethyl-, and 1,3,5,6-Tetramethyluracils. *J. Phys. Chem. A* **2013**, *117*, 244−251.

(19) Sagadeev, E. V.; Gimadeev, A. A.; Barabanov, V. P. The Enthalpies of Formation and Sublimation of Amino Acids and Peptides. *Russ. J. Phys. Chem. A* **2010**, *84*, 209−214.

(20) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gn theory. *WIREs Comput. Mol. Sci.* **2011**, *1*, 810−825.

(21) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. Gaussian-1 Theory: A General Procedure for Prediction of Molecular Energies. *J. Chem. Phys.* **1989**, *90*, 5622−5629.

(22) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 Theory for Molecular Energies of First- and Second-Row Compounds. *J. Chem. Phys.* **1991**, *94*, 7221−7230.

(23) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764−7776.

(24) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 Theory. *J. Chem. Phys.* **2007**, *126*, 84108−84112.
(25) Martin, J. M. L.; de Oliveira, G. Towards Standard Methods for Benchmark Quality *Ab Initio* Thermochemistry—W1 And W2 Theory. *J. Chem. Phys.* **1999**, *111*, 1843−1856.
(26) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. W3 theory: Robust Computational Thermochemistry in the kJ/mol Accuracy Range. *J. Chem. Phys.* **2004**, *120*, 4129−4141.
(27) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 Theory for Computational Thermochemistry: In Pursuit of Confident sub-kJ/mol Predictions. *J. Chem. Phys.* **2006**, *125*, 144108.
(28) Karton, A.; Martin, J. M. L. Explicitly Correlated W*n* Theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, *136*, 124114.
(29) Bakowies, D. Ab Initio Thermochemistry Using Optimal-Balance Models with Isodesmic Corrections: The Atomic Protocol. *J. Chem. Phys.* **2009**, *130*, 144113.
(30) Bakowies, D. Ab Initio Thermochemistry with High-Level Isodesmic Corrections: Validation of the ATOMIC Protocol for a Large Set of Compounds with First-Row Atoms (H, C, N, O, F). *J. Phys. Chem. A* **2009**, *113*, 11517−11534.
(31) Bakowies, D. Assessment of Density Functional Theory for Thermochemical Approaches Based on Bond Separation Reactions. *J. Phys. Chem. A* **2013**, *117*, 228−243.
(32) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. Heat: High Accuracy Extrapolated *Ab Initio* Thermochemistry. *J. Chem. Phys.* **2004**, *121*, 11599−11613.
(33) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. High-Accuracy Extrapolated *Ab Initio* Thermochemistry. II. Minor Improvements to the Protocol and a Vital Simplification. *J. Chem. Phys.* **2006**, *125*, 64108.
(34) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. High-Accuracy Extrapolated *Ab Initio* Thermochemistry. III. Additional Improvements and Overview. *J. Chem. Phys.* **2008**, *128*, 114111.
(35) Feller, D.; Dixon, D. A.; Peterson, K. A. Heats of Formation of Simple Boron Compounds. *J. Phys. Chem. A* **1998**, *102*, 7053−7059.
(36) Dixon, D. A.; Feller, D.; Sandrone, G. Heats of Formation of Simple Perfluorinated Carbon Compounds. *J. Phys. Chem. A* **1999**, *103*, 4744−4751.
(37) Feller, D.; Dixon, D. A. Theoretical Study of the Heats of Formation of Small Silicon-Containing Compounds. *J. Phys. Chem. A* **1999**, *103*, 6413−6419.
(38) Feller, D.; Dixon, D. A. Extended Benchmark Studies of Coupled Cluster Theory through Triple Excitations. *J. Chem. Phys.* **2001**, *115*, 3484−3496.
(39) Karton, A.; Yu, L. J.; Kesharwani, M. K.; Martin, J. M. L. Heats of Formation of the Amino Acids Re-Examined by Means of W1-F12 and W2-F12 Theories. *Theor. Chem. Acc.* **2014**, *133*, 1483−1497.
(40) Radom, L.; Lathan, W. A.; Hehre, W. J.; Pople, J. A. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. VIII. Geometries, Energies, and Polarities of C3 Hydrocarbons. *J. Am. Chem. Soc.* **1971**, *93*, 5339−5342.
(41) Wheeler, S. E.; Houk, K. N.; Schleyer, P. v. R.; Allen, W. D. A Hierarchy of Homodesmotic Reactions for Thermochemistry. *J. Am. Chem. Soc.* **2009**, *131*, 2547−2560.
(42) Dorofeeva, O. V.; Ryzhova, O. N. Gas-Phase Enthalpies of Formation and Enthalpies of Sublimation of Amino Acids Based on Isodesmic Reaction Calculations. *J. Phys. Chem. A* **2014**, *118*, 3490−3502.
(43) Ramabhadran, R. O.; Raghavachari, K. Theoretical Thermochemistry for Organic Molecules: Development of the Generalized Connectivity-Based Hierarchy. *J. Chem. Theory Comput.* **2011**, *7*, 2094−2103.
(44) Minnesota functionals M05-2X and M06-2X, along with BPW91, BMK, TPSSh, double hybrid B2PLYP, and commonly known B3LYP, were among the functionals tested in the work reported in ref 43.
(45) Ramabhadran, R. O.; Raghavachari, K. Connectivity-Based Hierarchy for Theoretical Thermochemistry: Assessment Using Wave Function-Based Methods. *J. Phys. Chem. A* **2012**, *116*, 7531−7537.
(46) Ramabhadran, R. O.; Sengupta, A.; Raghavachari, K. Application of the Generalized Connectivity-Based Hierarchy to Biomonomers: Enthalpies of Formation of Cysteine and Methionine. *J. Phys. Chem. A* **2013**, *117*, 4973−4980.
(47) The thermochemical reaction scheme obtained at the CBH-2 rung is called the isoatomic scheme because it preserves the immediate chemical environments of all of the atoms in an organic molecule. We therefore use the terms CBH-2 rung and isoatomic scheme interchangeably in this work. See refs 43 and 45 for more details.
(48) The detailed construction of the CBH can be found in refs 43 and 45.
(49) Vishveshwara, S.; Pople, J. A. Molecular Orbital Theory of the Electronic Structures of Organic Compounds. 32. Conformations of Glycine and Related Systems. *J. Am. Chem. Soc.* **1977**, *99*, 2422−2426.
(50) Suenram, R. D.; Lovas, F. J. Millimeter Wave Spectrum of Glycine. *J. Mol. Spectrosc.* **1978**, *72*, 372−381.
(51) Suellers, H. L.; Schafer, L. Investigations Concerning the Apparent Contradiction between the Microwave Structure and the Ab Initio Calculations of Glycine. *J. Am. Chem. Soc.* **1978**, *100*, 7728−7729.
(52) Brown, R. D.; Crofts, J. G.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. A Stark-Modulated Supersonic Nozzle Spectrometer for Millimetre-Wave Spectroscopy of Larger Molecules of Low Volatility. *J. Mol. Struct.* **1988**, *190*, 185−193.
(53) Schafer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. Theory Versus Experiment: The Case of Glycine. *J. Am. Chem. Soc.* **1980**, *102*, 6566−6568.
(54) Godfrey, P. D.; Firth, S.; Hatherley, L. D.; Brown, R. D.; Pierlot, A. P. Millimeter-Wave Spectroscopy of Biomolecules: Alanine. *J. Am. Chem. Soc.* **1993**, *115*, 9687−9691.
(55) Balabin, R. M. Conformational Equilibrium in Glycine: Experimental Jet-Cooled Raman Spectrum. *J. Phys. Chem. Lett.* **2010**, *1*, 20−23.
(56) Bazsó, G.; Najbauer, E. E.; Magyarfalvi, G.; Tarczay, G. Near-Infrared Laser Induced Conformational Change of Alanine in Low-Temperature Matrixes and the Tunneling Lifetime of its Conformer VI. *J. Phys. Chem. A* **2013**, *117*, 1952−1962.
(57) Jaeger, H. M.; Schafer, H. F.; Demaison, J.; Csaszar, A. G.; Allen, W. D. Lowest-Lying Conformers of Alanine: Pushing Theory to Ascertain Precise Energetics and Semiexperimental R_e Structures. *J. Chem. Theory Comput.* **2010**, *6*, 3066−3078.
(58) Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Császár, A. G.; Allen, W. D. Conformers of Gaseous Cysteine. *J. Chem. Theory Comput.* **2009**, *5*, 1511−1523.
(59) Allen, W. D.; Czinki, E.; Császár, A. G. Molecular Structure of Proline. *Chem.—Eur. J.* **2004**, *10*, 4512−4517.
(60) Boeckx, B.; Maes, G. Experimental and Theoretical Observation of Different Intramolecular H-bonds in Lysine Conformations. *J. Phys. Chem. B* **2012**, *116*, 12441−12449.
(61) Boeckx, B.; Maes, G. The Conformational Behavior and H-Bond Structure of Asparagine: A Theoretical and Experimental Matrix-Isolation FT-IR Study. *J. Biophys.Chem.* **2012**, *165*, 62−73.
(62) Schlund, S.; Müller, R.; Grabmann, C.; Engels, B. Conformational Analysis of Arginine in Gas Phase−A Strategy for Scanning the Potential Energy Surface Effectively. *J. Comput. Chem.* **2007**, *29*, 407−415.
(63) Ling, S.; Yu, W.; Huang, Z.; Lin, Z.; Haranczyk, M.; Gutowski, M. Gaseous Arginine Conformers and Their Unique Intramolecular Interactions. *J. Phys. Chem. A* **2006**, *110*, 12282−12291.
(64) Lesarri, A.; Mata, S.; López, J. C.; Alonso, J. L. A Laser-Ablation Molecular-Beam Fourier Transform Microwave Spectrometer: The Rotational Spectrum of Organic Solids. *Rev. Sci. Instrum.* **2003**, *74*, 4799−4805.
(65) Vaquero, V.; Cabezas, C.; Peña, I.; Perez, C.; López, J. C.; Alonso, J. L. Observation of Two New Conformers of Proline in the Gas Phase: A LA-MB-FTMW Study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4141−4144.

(66) Cabezas, C.; Varela, M.; Peña, I.; Mata, S.; López, J. C.; Alonso, J. L. The Conformational Locking of Asparagines. *Chem. Commun.* **2012**, *48*, 5934−5936.
(67) Sanz, M. E.; López, J. C.; Alonso, J. L. Six Conformers of Neutral Aspartic Acid Identified in the Gas Phase. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3573−3578.
(68) Mata, S.; Peña, I.; Sanz, M. E.; López, J. C.; Alonso, J. L. Preferred Conformers of Proteinogenic Glutamic Acid. *J. Am. Chem. Soc.* **2012**, *134*, 2305−2312.
(69) Raimondo, F.; Pieretti, A.; Gontrani, L.; Bencivenni, L. Hydrogen Bonding in Barbituric and 2-Thiobarbituric Acids: A Theoretical and FT-IR Study. *Chem. Phys.* **2001**, *271*, 293.
(70) Ralhan, S.; Ray, N. K. Density Functional Study of Barbituric Acid and its Tautomers. *J. Mol. Struct.: THEOCHEM* **2003**, *634*, 83.
(71) Daskalova, L. I.; Binev, I. Computational Study of Energies and Structures of 2,4,6-Pyrimidinetrione and its Anions. *Int. J. Quantum Chem.* **2006**, *106*, 1338−1345.
(72) Bertolasi, V.; Gilli, P.; Ferreti, V.; Gilli, G. Competition Between Hydrogen Bonding and Donor−Acceptor Interactions in Co-Crystals of 1,3-Dimethylbarbituric Acid with Aromatic Amines. *New. J. Chem.* **2001**, *25*, 408.
(73) Zuccarello, F.; Buemi, G.; Gandolfo, C.; Contino, A. Barbituric and Thiobarbituric Acids: A Conformational and Spectroscopic Study. *Spectrochim. Acta, Part A* **2003**, *59*, 139−151.
(74) Méndez, E.; Cerdá, M. F.; Gancheff, J. S.; Torres, J.; Kremer, C.; Castiglioni, J.; Kieninger, M.; Ventura, O. N. Tautomeric Forms of 2-Thiobarbituric Acid As Studied in the Solid, in Polar Solutions, and on Gold Nanoparticles. *J. Phys. Chem. C* **2007**, *111*, 3369−3383.
(75) Frisch, M. J.; Trucks, G. W.; Schlegel. H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalemani, G.; Barone, V.; Mennucci, B.; Pettersson, G. A.; et al. *Gaussian 09*, revision h08; Gaussian, Inc.; Wallington, CT, 2009.
(76) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(77) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(78) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364−382.
(79) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(80) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 34108.
(81) Császár, A. G. Conformers of Gaseous Glycine. *J. Am. Chem. Soc.* **1992**, *114*, 9568−9575.
(82) Hu, C. H.; Shen, M.; Schaefer, H. F. Glycine Conformational Analysis. *J. Am. Chem. Soc.* **1993**, *115*, 2923−2929.
(83) Rotations about C_α−C_β and C_β−C_γ bonds results in various low-energy conformers due to different types of intramolecular interactions shown in Scheme S2 in the Supporting Information.
(84) Lesarri, A.; Cocinero, E. J.; López, J. C.; Alonso, J. L. The Shape of Neutral Valine. *Angew. Chem., Int. Ed* **2004**, *43*, 605−610.
(85) Dokmaisrijan, S.; Lee, V. S.; Nimmanpipug, P. The Gas Phase Conformers and Vibrational Spectra of Valine, Leucine and Isoleucine: An Ab Initio Study. *J. Mol. Struct.: THEOCHEM* **2010**, *953*, 28−38.
(86) Cocinero, E. J.; Lesarri, A.; Grabow, J. U.; López, J. C.; Alonso, J. The Shape of Leucine in the Gas Phase. *ChemPhysChem* **2007**, *8*, 599−604.
(87) Lesarri, A.; Sánchez, R.; Cocinero, E. J.; López, J. C.; Alonso, J. L. Coded Amino Acids in the Gas Phase: The Shape of Isoleucine. *J. Am. Chem. Soc.* **2005**, *127*, 12952−12956.
(88) Lesarri, A.; Mata, S.; Cocinero, E. J.; Blanco, S.; López, J. C.; Alonso, J. L. The Structure of the Neutral Proline. *Angew. Chem., Int. Ed.* **2002**, *41*, 4673−4676.
(89) Alonso, J. L.; Peŕez, C.; Sanz, M. E.; López, J. C.; Blanco, S. Seven Conformers of L-Threonine in the Gas Phase: A LA-MB-FTMW Study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617−627.
(90) Blanco, S.; Sanz, M. E.; López, J. C.; Alonso, J. L. Revealing the Multiples Structures of Serine. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20183−20187.
(91) Chen, M.; Huang, Z.; Lin, Z. Ab Initio Studies of Gas Phase Asparagine Conformers. *J. Mol. Struct.: THEOCHEM* **2005**, *719*, 153−158.
(92) Ramaekersa, R.; Pajaka, J.; Rospenkb, M.; Maesa, G. Matrix-Isolation FT-IR Spectroscopic Study and Theoretical DFT(B3LYP)/6-31++ G** Calculations of the Vibrational and Conformational Properties of Tyrosine. *Spectrochim. Acta, A* **2005**, *61*, 1347−1356.
(93) Riziq, A. A.; Grace, L.; Crews, B.; Callahan, M. P.; Mourik, T. V.; de Vries, M. S. Conformational Structure of Tyrosine, Tyrosyl-glycine, and Tyrosyl-glycyl-glycine by Double Resonance Spectroscopy. *J. Phys. Chem. A* **2011**, *115*, 6077−6087.
(94) Tehrani, Z. A.; Tavasoli, E.; Fattahi, A. Conformational Behavior and Potential Energy Profile of Gaseous Histidine. *J. Mol. Struct.: THEOCHEM* **2010**, *960*, 73−85.
(95) Parry, G. S. The Crystal Structure of Uracil. *Acta Crystallogr.* **1954**, *7*, 313−320.
(96) Ozeki, K.; Sakabe, N.; Tanaka, J. The Crystal Structure of Thymine. *Acta Crystallogr., B* **1969**, *25*, 1038−1045.
(97) Experimental ΔH_f° values of molecules involved in the CBH reaction schemes were obtained from the following references for the evaluation of heats of formation of amino acids: (a) Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. *Gas Phase Ion and Neutral Thermochemistry*; American Chemical Society: Washington, DC, 1988. (b) *NIST Chemistry WebBook*; http://webbook.nist.gov/chemistry, accessed September 5, 2014.
(98) The enthalpies of formation evaluated with the remaining levels of theory are provided in Tables S10−S12 in the Supporting Information.
(99) De Oliveira, G.; Martin, J. M. L.; Silwal, I. K. C.; Liebman, J. F. Definitive Heat of Formation of Methylenimine, CH_2═NH, and of Methylenimmonium ion, $CH_2NH_2^+$, by Means of W2 Theory. *J. Comput. Chem.* **2001**, *22*, 1297−1305.
(100) Kruse, H.; Goerigk, L.; Grimme, S. Why the Standard B3LYP/6-31G* Model Chemistry Should Not Be Used in DFT Calculations of Molecular Thermochemistry: Understanding and Correcting the Problem. *J. Org. Chem.* **2012**, *77*, 10824−10834.
(101) Dorofeeva, O. V.; Ryzhova, O. N. Revision of Standard Molar Enthalpies of Formation of Glycine and L-Alanine in the Gaseous Phase on the Basis of Theoretical Calculations. *J. Chem. Thermodyn.* **2009**, *41*, 433−438.
(102) The CBH-1 reaction schemes for such fragments whose heats of formation were evaluated by CCSD(T) calculations are given in Table S8.
(103) Emel'yanenko, V. N.; Verevkin, S. P.; Achrainer, F.; Zipse, H. Book of Abstracts; 22nd International Conference on Chemical Thermodynamics and 67th Calorimetry Conference, August 5−10, 2012, Búzios, RJ, Brazil.
(104) Ribeiro da Silva, M. A. V.; Amaral, L. M. P. F.; Szterner, P. Thermochemical Study of 5-Methyluracil, 6-Methyluracil, and 5-Nitrouracil. *J. Chem. Thermodyn.* **2011**, *43*, 1924−1927.
(105) Nabavian, P. M.; Sabbah, R.; Chastel, R.; Laffite, M. Thermodynamique De Composes Azotes. II. Etude Thermochimique Des Acides Aminobenzoiques, De La Pyrimidine, De L'uracile Et De La Thymine. *J. Chim. Phys.* **1977**, *74*, 115−126.
(106) Imamura, A.; Takahashi, K.; Murata, S.; Sakiyama, M. Standard Enthalpies of Formation of Trimethyl Cyanurate, Malonamide, and 1,3-Dimethyluracil. *J. Chem. Thermodyn.* **1989**, *21*, 237−246.
(107) We have mentioned the number of conformers considered for the study of each amino acid in Table 3.