Unbiased disentanglement of conformational baths with the help of microwave spectroscopy, quantum chemistry and artificial intelligence: the puzzling case of homocysteine

Cite as: J. Chem. Phys. (in press) (2022); https://doi.org/10.1063/5.0102841
Submitted: 13 June 2022 • Accepted: 15 July 2022 • Accepted Manuscript Online: 15 July 2022

Iker León, Marco Fusè, Elena R. Alonso, et al.

Unbiased disentanglement of conformational baths with the help of microwave spectroscopy, quantum chemistry and artificial intelligence: the puzzling case of homocysteine

Iker León,[1, a)] Marco Fusè,[2, a)] Elena R. Alonso,[1] Santiago Mata,[1] Giordano Mancini,[2] Cristina Puzzarini,[3] José L. Alonso*,[1] and Vincenzo Barone*[2]

[1)]*Grupo de Espectroscopia Molecular (GEM), Edificio Quifima, Laboratorios de Espectroscopia y Bioespectroscopia Parque Cientifico UVa, Universidad de Valladolid, 47005 Valladolid (Spain).*

[2)]*SMART Laboratory, Scuola Normale Superiore, piazza dei Cavalieri 7, 56126 Pisa, Italy*

[3)]*Dipartimento di "Chimica Giacomo Ciamician", University of Bologna, via F. Selmi 2, 40126, Bologna, Italy*

(*Electronic mail: vincenzo.barone@sns.it)

(*Electronic mail: jlalonso@qf.uva.es)

(Dated: 13 July 2022)

An integrated experimental-computational strategy for the accurate characterization of the conformational landscape of flexible biomolecule building blocks is proposed. This is based on the combination of rotational spectroscopy with quantum-chemical computations guided by artificial intelligence tools. The first step of the strategy is the conformer search and relative stability evaluation performed by means of an evolutionary algorithm. In this step, last generation semiempirical methods are exploited together with hybrid and double-hybrid density functionals. Next, the barriers ruling the interconversion between the low-lying conformers are evaluated in order to unravel possible fast relaxation paths. The relative stabilities and spectroscopic parameters of the "surviving" conformers are then refined using state-of-the-art composite schemes. The reliability of the computational procedure is further improved by the inclusion of vibrational and thermal effects. The final step of the strategy is the comparison between experiment and theory without any *ad hoc* adjustment, which allows an unbiased assignment of the spectroscopic features in terms of different conformers and their spectroscopic parameters. The proposed approach has been tested and validated for homocysteine, a highly flexible non-proteinogenic α-amino acid. The synergism of the integrated strategy allowed the characterization of five conformers stabilized by bifurcated N-$H_2\cdots$O=C hydrogen bonds, together with an additional conformer involving a more conventional HN$\cdots$H-O hydrogen bond. The stability order estimated from the experimental intensities as well as the number and type of conformers observed in the gas phase are in full agreement with the theoretical predictions. Analogously, a good match has been found for the spectroscopic parameters.

I. INTRODUCTION

Gas-phase studies of biomolecule building blocks provide vital information about their intrinsic stereo-electronic properties: the absence of any solvent or, more generally, non-innocent environment allows extracting all the structural and spectroscopic characteristics without any strong external perturbation[1–3]. This is the most natural framework for accurate quantum-chemical (QC) computations and high-resolution molecular spectroscopy, but a number of difficulties have to be faced for medium-sized flexible molecules. From the theoretical point of view, the effective exploration of flat potential energy surfaces (PESs) and the characterization of their stationary points become prohibitive using the local optimization methods implemented in the current QC software. Furthermore, the size of the systems and the number of different structures to be investigated prevent the systematic use of very accurate, but very expensive state-of-the-art QC methodologies[4]. In this connection, the ongoing development of accurate yet effective composite schemes[5] and powerful PES exploration methods based on artificial intelligence (AI)[6] are paving the route toward robust approaches, also exploitable by non specialists. From the experimental point of view, owing to recent developments and refinements[7–9], the supersonic-jet expansion (SE) technique[10] possibly coupled to laser ablation (LA)[2,11] has been extending the field of application of high-resolution molecular spectroscopy[12,13]. Indeed, LA allows the gas-phase study of compounds with high melting points, whereas the relevant cooling of the translational, rotational and vibrational degrees of freedom of the species in the SE of gases increases the intensity of the spectral signals. This also leads to a considerable simplification of the spectrum because of the increased population of low-lying states, thus allowing the investigation of the most stable conformers. Therefore, microwave (MW) spectroscopy combined with LA and SE is one of the few techniques that allow detailed structural, energetic and spectroscopic characterizations of flexible, medium-sized biomolecules. However, a potential drawback of SE related to its intrinsic cooling process, is the number and nature of conformers that can be actually investigated. A very illustrative example of conformer inter-conversion in SEs was offered by Miller et al.,[14] who described how the original conformer distribution could change during the cooling process through either vertical collapse or conformer inter-conversion. The latter process leads to a bi-

[a)]Contributed equally to this work

ased picture of the system thermochemistry because of "missing" conformers[15–17]. Nevertheless, the degree to which interconversion occurs during the cooling associated to SE can be fully explained by the intra-molecular PES of the species under investigation. This peculiar behavior must be considered when studying flexible systems characterized by rugged conformational landscapes of high dimensionality[18–20].

Although several studies of biomolecule building blocks have been performed (see e.g. refs. 16–18,21–25), the computational characterization of PESs is usually restricted to more or less systematic searches of energy minima employing geometries (when not also energies) of limited accuracy. These approaches do not allow the *a priori* prediction of the spectroscopic outcome, but only its *a posteriori* interpretation in terms of agreement between experimental and computed spectroscopic parameters for a reduced number of conformers, which are not necessarily the most stable ones. In this latter case, the *a posteriori* analysis might lead to misleading interpretation, with any unbiased disentanglement of the conformational landscape becoming impossible. Based on these premises, we have decided to develop and validate a fully integrated strategy combining artificial intelligence, state-of-the-art quantum chemistry and microwave spectroscopy for the quantitative description of the PES of flexible systems in the gas phase in terms of both energy minima and interconversion paths.

With the aim of illustrating the proposed strategy at work, we will make specific reference to isolated amino acids, whose conformational behavior is ruled by backbone (ϕ' = LP–N–C$^{\alpha}$–C', ψ = N–C$^{\alpha}$–C'–OH and ω = C$^{\alpha}$–C'–O–H, with LP denoting the nitrogen lone-pair) and side-chain (σ) torsional angles (see Figure 1). We prefer to use σ labels in place of the more customary χ ones in order to avoid any ambiguity with the elements of the quadrupole coupling tensor (vide infra). Furthermore, the presence of two hydrogen atoms and the consequent non planarity of the NH_2 moiety suggest the use of the ϕ' dihedral angle in place of its standard ϕ = H–N–C$^{\alpha}$–C' + 180° counterpart (with $\phi' = \phi + 120°$).

The most stable structures of amino acid backbones involve the formation of hydrogen bonds between the COOH and NH_2 moieties. According to the intramolecular interactions established and the values of the torsional angles introduced above, conformers can be classified as I (bifurcated $NH_2\cdots$O=C hydrogen bond, $\phi' \approx 180°$, $\psi \approx 180°$, $\omega \approx 180°$), II (N$\cdots$H(O) hydrogen bond, $\phi' \approx 0°$, $\psi \approx 0°$, $\omega \approx 0°$), or III (bifurcated $NH_2\cdots$OH hydrogen bond, $\phi' \approx 180°$, $\psi \approx 0°$, $\omega \approx 180°$)[2,26]. Noted is that the type I hydrogen bridges can be formed for both $\omega \approx 0°$ and $\omega \approx 180°$, with the latter value being intrinsically preferred by the carboxylic moiety. The standard g, t and g$^-$ labels are employed to denote the staggered conformers for each dihedral angle in flexible side chains ($\sigma_i \approx 60°, 180°$ or $-60°$, respectively).

In view of the discussion above, we have selected as test case homocysteine (see Figure 1), a non-proteinogenic α-amino acid that plays a vital role as an intermediate in generating other amino acids and biomolecules.[27] Its structure is similar to that of methionine, but the presence of the SH moiety in place of the SCH_3 group (see Figure 1) might give rise to additional backbone / side-chain hydrogen bonds. At the same time, the inclusion of an additional methylene group in the side-chain of homocysteine with respect to cysteine (see Figure 1) increases the challenge because, on the one hand, a larger number of low-energy conformers is expected and, on the other hand, additional relaxation pathways are possible.

II. RESULTS AND DISCUSSION

The first step of our computational strategy is to find and characterize all low-energy minima, with special reference to their relative stability. The PESs of flexible systems are typical examples of rugged high dimensional surfaces, where convex optimization methods undergo premature convergence and approaches based on systematic scans become extremely expensive due to the combinatorial explosion (at a 30° resolution, a complete sampling of the conformational space of homocysteine would involve 12^6 points). Therefore, we have further refined a two-step algorithm rooted in the Island Model evolutionary algorithm[6] and driven by the PROXIMA library[28], which has proven to be extremely effective for several flexible molecules (details are provided in the Supporting Information, SI)[29].

The initial exploration phase involved three replicas, each consisting of about 4000 constrained geometry optimizations, employing a low-cost (GFN2-XTB[30]) semi-empirical method. Then, a first reduction to about 1000 structures was obtained by applying a threshold of 1250 cm^{-1} (15 kJ mol^{-1}) with respect to the absolute energy minimum. These candidates were compared with each other in terms of the root-mean-square deviations of heavy atom positions and rotational constant. The 300 structures surviving after this selection were further reduced to 22 by clustering procedures (see SI for details). Subsequently, full geometry optimizations at the PW6B95/jul-cc-pV(D+d)Z level[31–33], also including Grimme's D3BJ dispersion correction[34] (hereafter, this level of the theory is referred to as PW6), led to the relaxation of 6 structures to adjacent energy minima, thus leaving 16 conformers spanning an energy range of about 1000 cm^{-1} (11 kJ mol^{-1}). Their geometries were further refined at the rev-DSD-PBEP86-D3/jun-cc-pV(T+d)Z level (hereafter rDSD)[32,33,35]. The number of conformers was then increased to 17 after additional exploration of staggered orientations around the σ_3 dihedral angle (required to analyze possible relaxation paths, *vide infra*) at the rDSD level, which showed that the Ig$^-$g$^-$t structure is an energy minimum at this level of theory. Other computational methods (PW6B95, B3LYP-D3[34,36] and Møller-Plesset perturbation theory to the second order, MP2[37], all in conjunction with the jun-cc-pV(T+d)Z basis set[32,33], hereafter denoted as PW6TZ, B3 and MP2, respectively) were also considered for comparison purposes (see details in the SI). This composite strategy allows for strongly reducing the number of costly geometry optimizations by hybrid and, especially, double-hybrid functionals without any loss of accuracy in the final results. In any case, we were left with 17 conformers lying (at the rDSD level) within 700 cm^{-1} (about 8 kJ mol^{-1}) above the absolute en-

Accepted to J. Chem. Phys. 10.1063/5.0102841

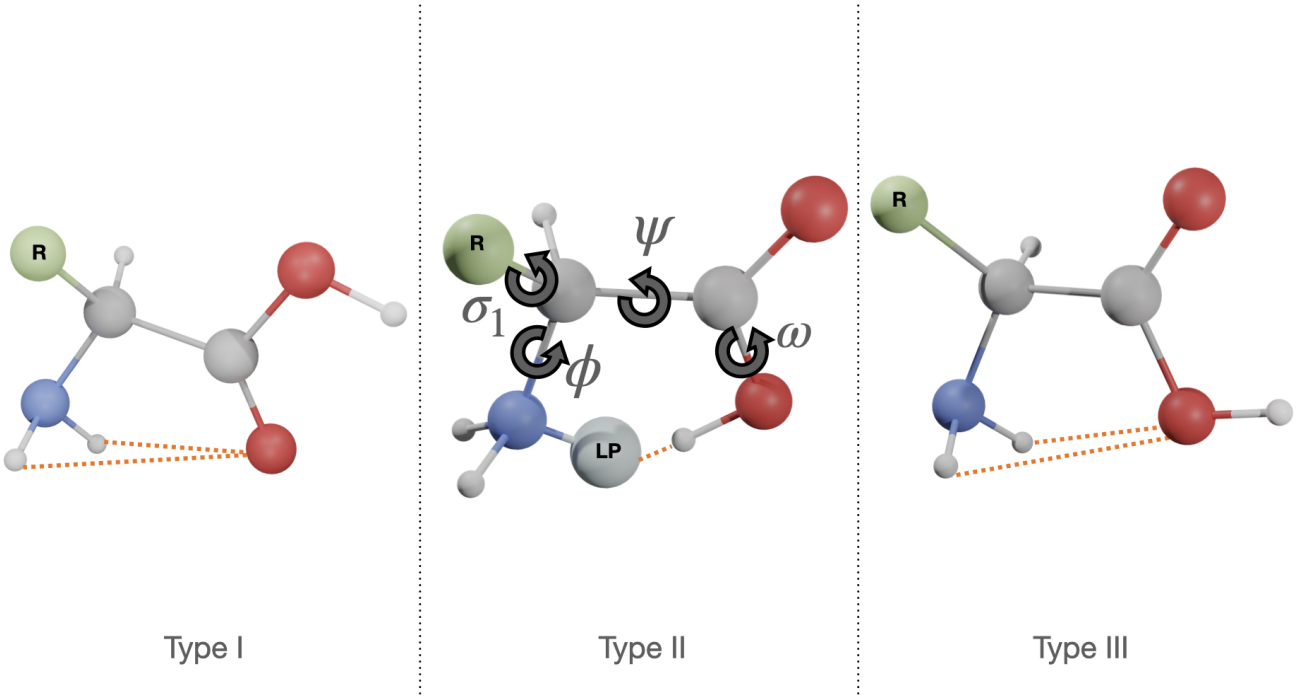


FIG. 1. Structure and main dihedral angles of a generic amino acid, with R denoting the side-chain (CH_2SH, CH_2-CH_2-SH and CH_2-CH_2-SCH_3 in the case of cysteine, homocysteine, and methionine, respectively). See main text for further details.

ergy minimum (see Tables S2 and S3 of the SI), thus pointing out the significantly larger flexibility of homocysteine with respect to cysteine[6,18,25]. Interestingly, computational levels often employed in this type of studies (i.e. B3 and/or MP2) strongly alter the rDSD trends (see Figure 2 and Table S2 of the SI). To give an example, at the rDSD level, the most stable conformer is $Ig^-g^-g^-$, whereas it is $IItgg^-$ at the MP2 level and $IIg^-g^-g^-$ at the PW6, PW6TZ and B3 levels. Even the number of stable minima lying below a given energy threshold was found to strongly depend on the level of theory. This situation prompted us to refine the relative energies by performing very accurate single-point energy computations on top of the rDSD geometries. These improved energy evaluations exploited the so-called CBS+CV-F12 composite scheme, which involves calculations using explicitly correlated (F12) techniques[38]. In details, electronic energies were computed using the CCSD(T)-F12b(3C/FIX) method[39–41] (hereafter simply referred to as CCSD(T)-F12), within the frozen-core (fc) approximation, in conjunction with the cc-pV(n+d)Z-F12 (with n=D,T) orbital basis sets[42] and the corresponding auxiliary basis sets. Extrapolation to the complete basis set (CBS) limit was then carried out by exploiting the standard n^{-3} two-point formula[43] and the core-valence (CV) correlation contribution was incorporated as difference between all-electron (ae) and fc CCSD(T)-F12 calculations, both in the cc-pCV(D+d)Z-F12 basis set[44]. Further details are given in the SI. The agreement between the energies obtained at this level and the corresponding rDSD values (see Figure 2 and Table S2 of the SI) is indeed remarkable with mean and maximum errors being 12.4 and 38.1 cm^{-1} (0.15 and 0.46 kJ mol^{-1}), respectively. This finding gives further support to the use of the rDSD model for the evaluation of structures and spectroscopic properties. Noted is that the errors are about one third of those shown by the MP2 and B3 counterparts, and only rDSD does not present any significant stability inversion with respect to the CBS+CV-F12 order.

The next step of the proposed strategy is to proceed from electronic energy differences to the corresponding enthalpies at 0 K (ΔH_0°) and Gibbs free energies at room temperature (ΔG°), with rDSD harmonic frequencies being employed to compute zero point energies (ZPEs) and vibrational partition functions (see Table S4 of the SI). This step leads to significant changes in the trend issuing from relative electronic energies (see Figure 2 and Tables S2, S4 of the SI) and, in particular, to the destabilization of all the conformers showing type II hydrogen bridges. Indeed, only one of them ($IIg^-g^-g^-$) still lies within 600 cm^{-1} above the most stable (i.e. lowest free energy) conformer. Since the Ig^-gg^- conformer is also slightly destabilized, we are left with 12 conformers possibly detectable in MW experiments (see Figure 3 and Table S2 of the SI).

To allow the comparison with experiment, together with geometries (straightforwardly providing equilibrium rotational constants) and harmonic frequencies, also dipole moments, nuclear quadrupole couplings and quartic centrifugal distortion constants were computed at the rDSD level (see Tables

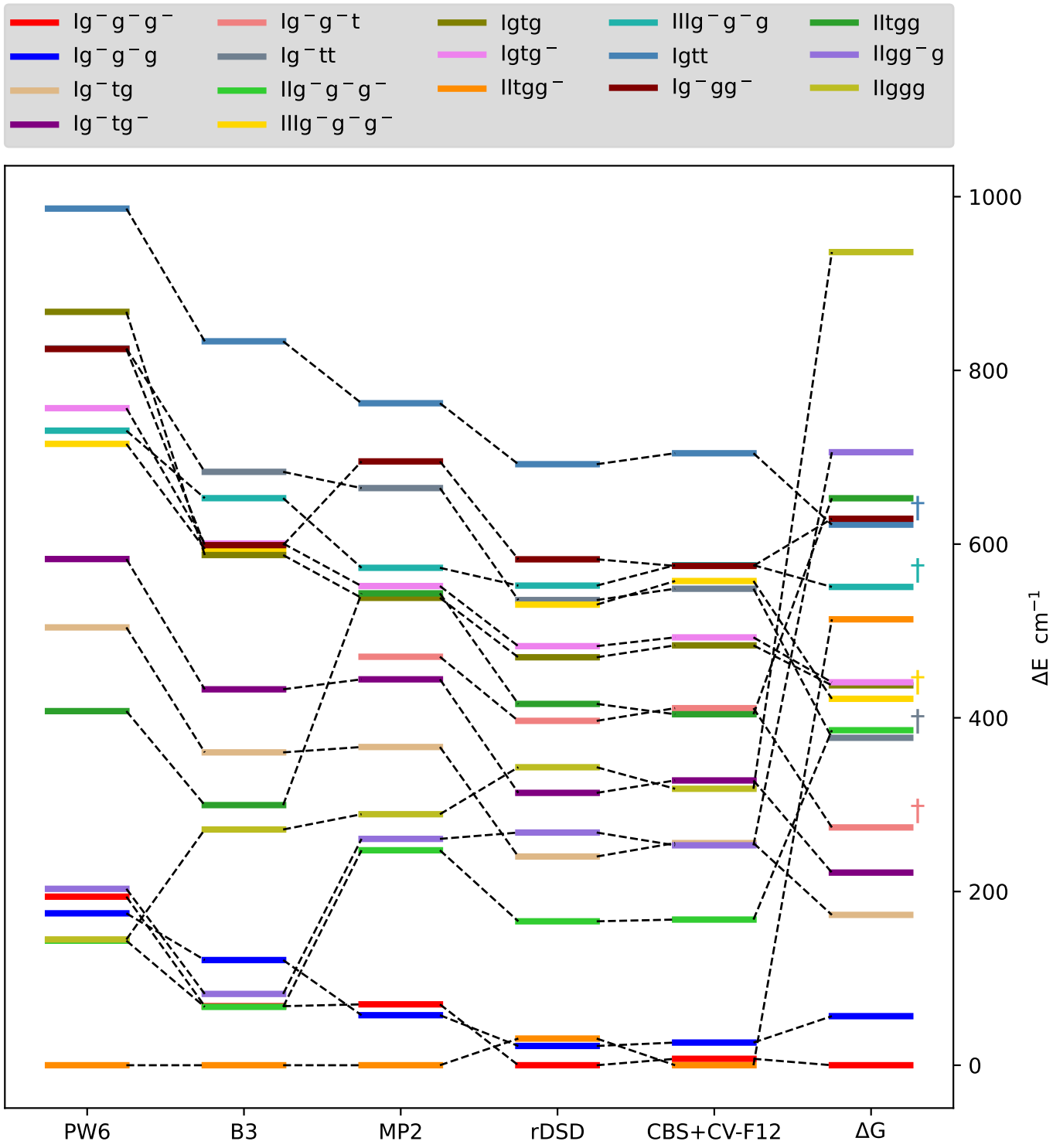


FIG. 2. Comparison between the relative stability of low-lying cysteine conformers obtained at different levels of theory. Conformers marked with a dagger relax to more stable conformers. ΔG values are obtained from CBS+CV-F12 electronic energies (see text).

S4, S5 and S6 in the SI), whereas the anharmonic contributions, required for going from equilibrium to effective rotational constants[45,46], were computed at the PW6 level (see Table S5 in the SI). More accurate structures, and thus improved equilibrium rotational constants, were obtained by correcting the rDSD geometrical parameters with the so-called linear regression approach (LRA)[45]. Within the latter, systematic errors affecting bond lengths and valence angles are corrected based on linear regressions, whose parameters were derived from a large database of accurate semi-experimental equilib-

rium geometries[46] (see Table S5 of the SI).

As mentioned above, conformational relaxation can take place under the experimental conditions whenever the energy barriers ruling the interconversion are sufficiently low, with an upper limit of about 400 cm^{-1} (4.8 kJ mol^{-1}) being usually employed for discriminating in rotational spectroscopy of amino acids and related compounds[15–17]. With the aim of understanding which conformational relaxations can take place, we performed relaxed torsional scans at the rDSD level in order to obtain preliminary information on low-energy interconversion paths. Next, after full geometry optimizations, the second-derivatives of the energy with respect to Cartesian coordinates (Hessian matrices) of the most significant transition states were computed at the rDSD level. Some of the torsional scans performed are graphically represented in Figure 4, whereas the rDSD dihedral angles and imaginary frequencies of the most significant transition states together with their relative energies (both rDSD and CBS+CV-F12) are given in Tables S3 and S7 of the SI.

The results show that rotation of the SH group (σ_3 dihedral angle) from the less stable trans arrangement to the preferred gauche ones is always ruled by a negligible energy barrier, so that the Ig$^-$tt conformer can easily relax to either Ig$^-$tg or Ig$^-$tg$^-$. In the same way, the Ig$^-$g$^-$t conformer can easily relax to Ig$^-$g$^-$g$^-$, whereas a barrier of more than 650 cm^{-1} separates the Ig$^-$tg$^-$ conformer from Ig$^-$tg, thus allowing the detection of both forms in the gas-phase mixture. The same applies to the conversion of Ig$^-$g$^-$g into Ig$^-$g$^-$g$^-$, which requires to overcome a barrier of about 600 cm^{-1}, largely sufficient to trap both structures during the expansion, thus allowing their detection. Conversion of type III into type I conformers can occur via rotation around the ψ dihedral angle. Contrary to the case of cysteine[18], in homocysteine this process is ruled by low energy barriers (below 100 cm^{-1}) due to the lack of any constraint from the side chain, because the SH moiety is not involved (unlike cysteine) in any significant hydrogen bond with other polar groups. As a consequence, the IIIg$^-$g$^-$g$^-$ and IIIg$^-$g$^-$g conformers relax to Ig$^-$g$^-$g$^-$ and Ig$^-$g$^-$g, respectively. Conversion of type II structures into other conformers involves concerted rotation around at least two dihedral angles, with the corresponding minimum energy paths being ruled by energy barriers sufficiently high to trap the IIg$^-$tg$^-$ conformer, thus allowing its detection. The ZPE-corrected energies of some minima lie above or very close to those of the corresponding transitions states (365.9 vs. 314.6 cm^{-1} for Ig$^-$g$^-$t; 636.6 vs. 550.0 cm^{-1} for IIIg$^-$g$^-$g$^-$, and 638.4 vs. 618.7 cm^{-1} for IIIg$^-$g$^-$g). As a consequence, these isomers cannot be trapped even at very low temperatures, whereas gas-phase vibrational studies[47] (but not MW experiments) could possibly characterize the Ig$^-$g$^-$t conformer, whose ZPE-corrected energy lies marginally below that of the pertinent TS (505.4 vs. 593.5 cm^{-1}). The final picture of the observable low-lying conformers of homocysteine is summarized in Figure 3.

Since the barrier height provides crucial information on whether a conformer can be observed or not, it is important to point out that, in the present case, the QC approaches usually employed in conformational studies (e.g., B3 and other hybrid functionals or MP2) tend to fail in locating challenging low-energy minima and/or reproducing the correct stability order of the others (also including the most stable one). On the contrary, the double-hybrid rDSD model performs remarkably well in this connection and in correctly evaluating the energy barriers ruling interconversion paths (see Figure 2 and Tables S2 and S7 of the SI). As a matter of fact, the rDSD level provides for transition states mean and maximum absolute errors, with respect to the CBS+CV-F12 reference (15.1 and 37.7 cm^{-1}, i.e., 0.18 and 0.45 kJ mol^{-1}), that are comparable to those noted for minima (12.4 and 38.1 cm^{-1}, i.e., 0.15 and 0.46 kJ mol^{-1}) and to those obtained at the CCSD(T)-F12/cc-pV(D+d)Z-F12[40,42] level, hereafter referred to as DZ-F12 (15.1 and 28.5 cm^{-1}, i.e., 0.18 and 0.34 kJ mol^{-1}).

According to both rDSD and CBS+CV-F12 results, the most populated conformer of homocysteine is Ig$^-$g$^-$g$^-$, which is stabilized by a bifurcated hydrogen bond between the two aminic hydrogen atoms and the carboxylic group, with the latter in a cis conformation ($\omega \approx 180^\circ$). The close similarity of Ig$^-$tg$^-$ and Ig$^-$tg conformers (see Tables S2 and S3 of the SI) prompted us to reoptimize their geometry (together with that of the most stable Ig$^-$g$^-$g$^-$ conformer) at the very reliable DZ-F12 level, also including CV contributions at the MP2-F12/cc-pCV(D+d)Z-F12 level (hereafter DZ-F12+CV). The results, collected in Tables S3 and S5 of the SI, show that the rDSD and DZ-F12+CV dihedral angles are very similar to each other and that the LRA corrections applied to rDSD bond lengths and valence angles provide rotational constants very close to the DZ-F12+CV counterparts. Furthermore, vibrational corrections to rotational constants have been computed for two representative conformers (Ig$^-$g$^-$g$^-$ and IIg$^-$g$^-$g$^-$) at the rDSD level. It is noteworthy that the results are very similar to the PW6 counterparts for the type-II structure, whereas a larger correction is obtained for the type-I structure (see Table S5 of the SI). In order to take this effect into account in the comparison with experimental data (vide infra), the vibrational corrections of all the other type-I conformers have been scaled by the ratio between rDSD and PW6 results for the Ig$^-$g$^-$g$^-$ prototype.

In general terms, homocysteine follows the trend observed for all but one of the aliphatic amino acids studied till now, namely that type I conformers are more populated than the type II counterparts.[18] Type III structures have been detected in MW experiments only in the presence of conjugative effects[48] or when polar side chains can form additional hydrogen bridges with the backbone[2,18,21,49]; in all these cases, they are the least populated forms. The inhability of the homocysteine side chain to form strong hydrogen bonds with the backbone is confirmed by the comparable values for the relative electronic energies of the structures of type II and III with respect to the absolute energy minimum (of type I) obtained for homocysteine (e.g., 167.7 and 557.5 cm^{-1} for IIg$^-$g$^-$g$^-$ and IIIg$^-$g$^-$g$^-$ with respect to Ig$^-$g$^-$g$^-$) and glycine (191.4 and 621.9 cm^{-1} for II and III with respect to I at the CBS+CV level, albeit using slightly different composite schemes[50]).

In conclusion, eight conformers are expected to be detectable for homocysteine, with six structures of type I and two conformers of type II. However, the population of the

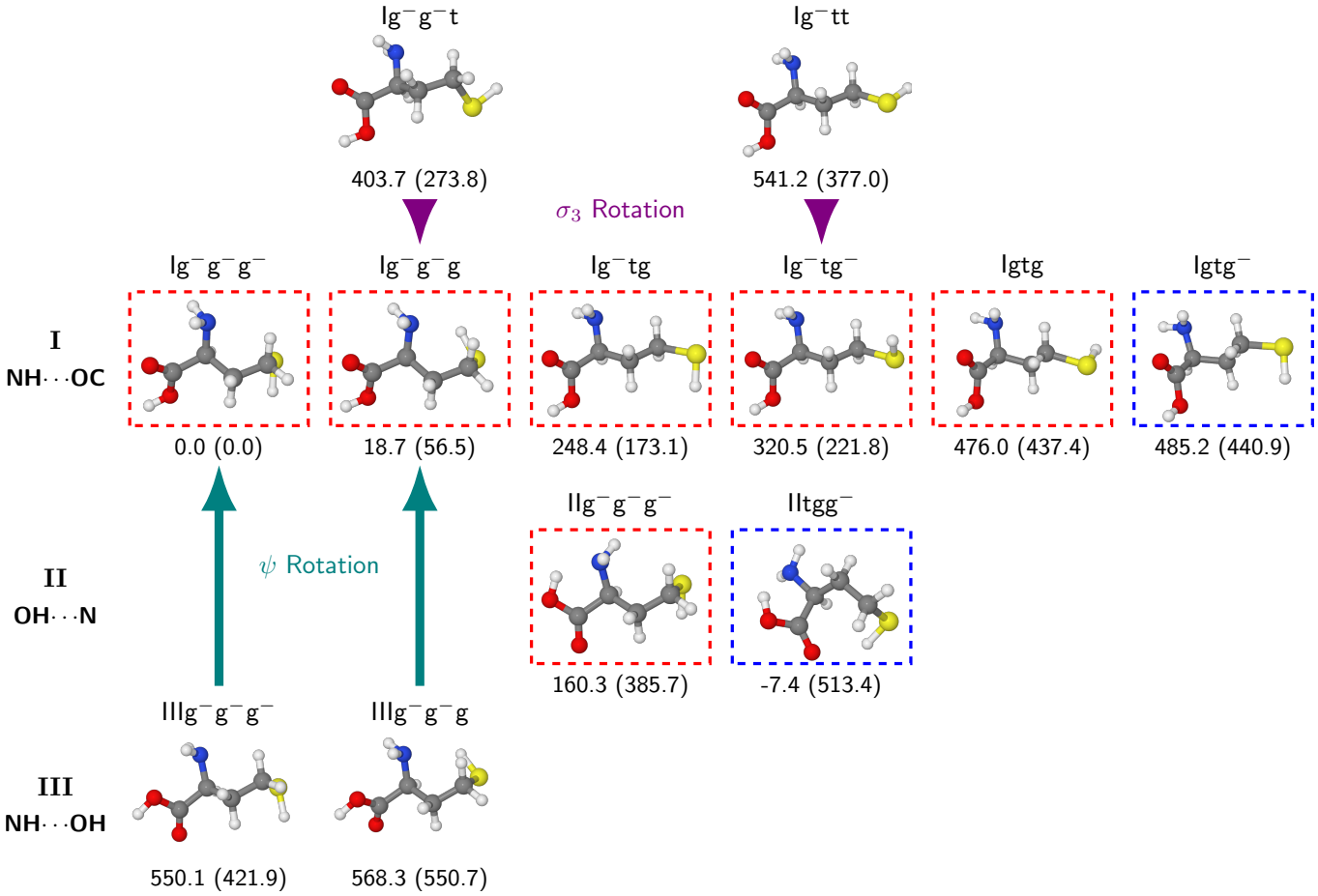


FIG. 3. Predicted low-energy conformers of homocysteine. The conformers not involved in effective relaxation paths are highlighted by red dotted frames, with the arrows indicating the preferred isomerization paths. Blue frames highlight structures not involved in relaxation paths and not experimentally observed. Relative CBS+CV-F12 electronic energies and (in parentheses) Gibbs free energies at room temperature in cm^{-1} are also reported.

IItgg$^-$ conformer could be too low to allow its unequivocal detection and the very similar spectroscopic parameters of the Igtg and Igtg$^-$ conformers could hamper the disentanglement of the contributions of the latter one. The computed ΔG° (see Table S2 of the SI) lead to populations decreasing in the order Ig$^-$g$^-$g$^-$ > Ig$^-$g$^-$g > Ig$^-$tg > Ig$^-$tg$^-$ > IIg$^-$g$^-$g$^-$ > Igtg ≈ Igtg$^-$ > IItgg$^-$. The picture issued from our fully *a priori* predictions is thus very different from that of the parent cysteine amino acid, where six conformers were actually detected, with all backbone hydrogen bond-type structures (I, II and III) observed[18]. Noted is that not all the detected conformers of cysteine are among the six most stable ones predicted by accurate QC computations[25] and the lack of any relaxation path characterization does not allow further analysis.

The large number of conformers and the unusual sensitivity of the results to the computational level make homocysteine the perfect test for the integrated experimental-computational strategy purposely set up for biomolecule building blocks. Indeed, it allows us to stress the importance of combining the AI/QC analysis with experimental data in order to unambiguously elucidate the problem.

In view of its high melting point (m.p. 233C), neutral homocysteine molecules have been brought in the gas phase using the LA technique. The latter combined with a chirped-pulse Fourier transform MW spectrometer (CP-FTMW)[51,52] permitted the recording of the rotational spectrum of homocysteine in the 6.0-12.0 GHz frequency region, which is shown in Figure 5. Its analysis allowed us to characterize six rotamers, labeled as *L,M, N, O, P, Q*, whose rotational parameters are collected in Table I. For some conformers, the assignment straightforwardly proceeded owing to the very good agreement between experimental and computed rotational constants. In detail, rotamer *L* was easily assigned to the Ig$^-$g$^-$g$^-$ structure and rotamer *Q* to Igtg. Instead, the experimental rotational constants of rotamers *N* and *O* are close to each other and required a deeper analysis. For both conformers, we measured strong *a*-type transitions, but only the rotamer *N* shows moderate transitions of *b*- and *c*-types. On the contrary, rotamer *O* shows weak *c*-type transitions but no *b*-type transitions. These outcomes are in good agreement with the calculated electric dipole moments for the Ig$^-$tg and Ig$^-$tg$^-$ structures, respectively. Analogously, the rotamer *P*

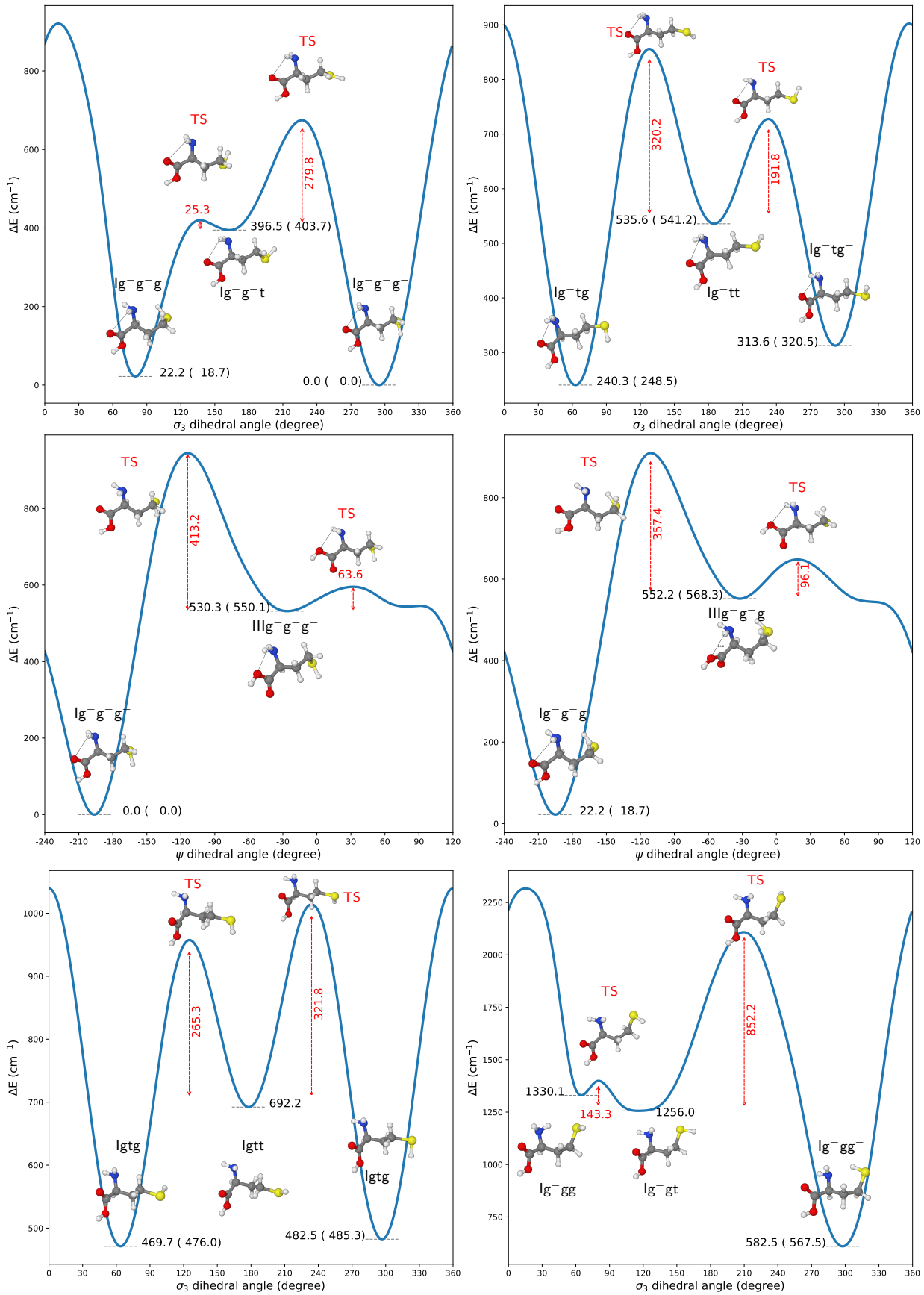


FIG. 4. Potential energy profiles (in cm^{-1}) along ψ and σ_3 dihedral angles computed at the rDSD and, in parentheses, CBS+CV levels.

TABLE I. Experimental spectroscopic constants for the six observed rotamers of homocysteine, together with the computed relative energies and spectroscopic parameters of the six most stable energy minima.

Exp.[a]	L	M	N	O	P	Q
A_0	3132.1285(20)[b]	3279.338(69)	3448.9350(12)	3462.2693(36)	3260.8978(10)	2713.3895(54)
B_0	865.73549(25)	848.87390(20)	719.16486(21)	717.92241(56)	843.482212(98)	871.0027(28)
C_0	768.52872(25)	772.24934(12)	644.48207(20)	645.32222(35)	768.94684(12)	779.2561(36)
Δ_J	-0.0941(27)	Not determined.	-0.0000353(30)	Not determined	-0.0896(10)	Not determined
Δ_{JK}	-0.45(11)	Not determined	-0.000757(67)	Not determined	-1.128(48)	Not determined
χ_{aa}	-3.3067(83)	-4.0533(128)	-3.2402(35)	-4.1007(30)	-3.0631(43)	Not determined
χ_{bb}	1.8925(67)	2.5207(207)	1.9421(25)	2.1745(38)	2.0811(36)	Not determined
χ_{cc}	1.4142(67)	1.5327(207)	1.2981(25)	1.9261(38)	0.9820(36)	Not determined
μ_a	Observed	Observed	Observed	Observed	Observed	Observed
μ_b	Observed	Not observed	Observed	Not observed	Observed	Not observed
μ_c	Observed	Observed	Observed	Observed	Observed	Observed
Computed	$Ig^-g^-g^-$	Ig^-g^-g	Ig^-tg	Ig^-tg^-	$IIg^-g^-g^-$	Igtg
A_0^c	3127.5 (3128.3)	3305.5	3436.5 (3444.5)	3413.3 (3422.1)	3258.2	2713.4
B_0^c	863.0 (865.5)	846.9	718.9 (719.5)	712.6 (713.1)	840.6	861.5
C_0^c	766.6 (768.7)	776.7	645.4 (645.8)	643.2 (643.6)	766.8	777.0
Δ_J	-0.1003	-0.0791	-0.0417	-0.0417	-0.0739	-0.1297
Δ_{JK}	-0.501	-0.865	-0.507	-0.530	-1.347	-1.184
χ_{aa}	-3.68 (-3.64)	-4.35	-3.29	-3.33 (-3.30)	-2.99 (-3.33)	-0.98
χ_{bb}	1.88 (1.86)	2.51	1.92	1.89 (1.90)	1.94 (1.87)	0.02
χ_{cc}	1.80 (1.69)	1.84	1.37	1.44 (1.40)	1.45 (1.47)	0.96
μ_a	0.65	1.09	1.60	1.58	-3.18	1.63
μ_b	1.11	-0.28	-1.16	-0.20	2.72	-0.03
μ_c	0.89	-0.44	1.44	0.74	-1.71	-1.16
ΔE_{el}^{d}	0.0	26.1	255.8	327.9	167.7	483.5
$\Delta H_0^{\circ d,e}$	0.0	59.0	226.4	292.6	283.3	482.4
$\Delta G^{\circ d,f}$	0.0	63.8	180.5	229.0	385.6	444.9

[a] A_0, B_0, and C_0 are the rotational constants (in MHz); Δ_J and δ_{JK} are two quartic centrifugal distortion constants (in kHz); χ_{aa}, χ_{bb}, and χ_{cc} are the diagonal elements of the ^{14}N nuclear quadrupole coupling tensor (in MHz); μ_a, μ_b, and μ_c are the electric dipole moment components (in debye). [b] Standard errors are shown in parentheses in units of the last digits. [c] The rDSD structural parameters corrected using the LRA approach (see text) led to the equilibrium rotational constants, then corrected for the vibrational corrections. In parentheses, the effective rotational constants obtained from CBS+CV-F12 geometries with the same vibrational corrections. Quartic centrifugal distortion, nitrogen quadrupole coupling constants and dipole moment components are at the rDSD level, with the exception of the values in parentheses, which were obtained at the CCSD(T)/jun-cc-pV(T+d)Z level also including CV contributions at the MP2/aug-cc-pwC(T+d)Z level. [d] Electronic energy at the CBS+CV-F12 level. [e] harmonic rDSD ZPE correction. [f] rDSD thermal contributions.

can be distinguished from *M* because of *b*-type transitions, which are only possible for the former species. Thus, the rotamer *P* is assigned to $IIg^-g^-g^-$ and the rotamer *M* to Ig^-g^-g.

The presence of one ^{14}N nucleus, which has a non-vanishing quadrupole moment, allows for gaining further information on the conformer structures. The quadrupole moment interacts with the electric field gradient at the nucleus, thus leading to the so-called nuclear quadrupole coupling, which determines the splitting of the rotational energy levels and, as a consequence, the splitting of the rotational transitions (hyperfine structure).[53] In homocysteine, nitrogen quadrupole coupling constants provide accurate information on the orientation of the NH_2 group. For this reason, in a second step, some transitions assigned in the LA-CP-FTMW spectrum were measured using the LA-MB-FTMW spectrometer (MB standing for Molecular Beam)[8,54], which is characterized by higher sensitivity and resolution with respect to the former. All measured transitions and hyperfine components are collected in Tables S8 to S13 of the SI. Figure 5b shows an example of some transitions exhibiting the characteristic hyperfine structure. The quadrupole coupling constants obtained from the analysis of the hyperfine structures are collected in Table I. Their comparison with the corresponding computed values further confirms the conformational assignment based on rotational constants. This is particularly true for the *M* and *P* conformers, whose predicted quadrupole coupling constants are very similar to the experimental counterparts. Instead, discriminating between the *N* and *O* conformers is more challenging and requires a deeper analysis. Although they can be distinguished in terms of the observed selection rules and intensities as explained above, the quadrupole coupling constants of the Ig^-tg form reproduce very well the hyperfine structures of the conformer denoted as *N* in Figure 5, whereas the same does not apply to the comparison of the hyperfine structures of Ig^-tg^- with that of the conformer *O*.

Since the error affecting the computed A_0 effective rotational constant of the O rotamer (49 MHz) is about 5 times

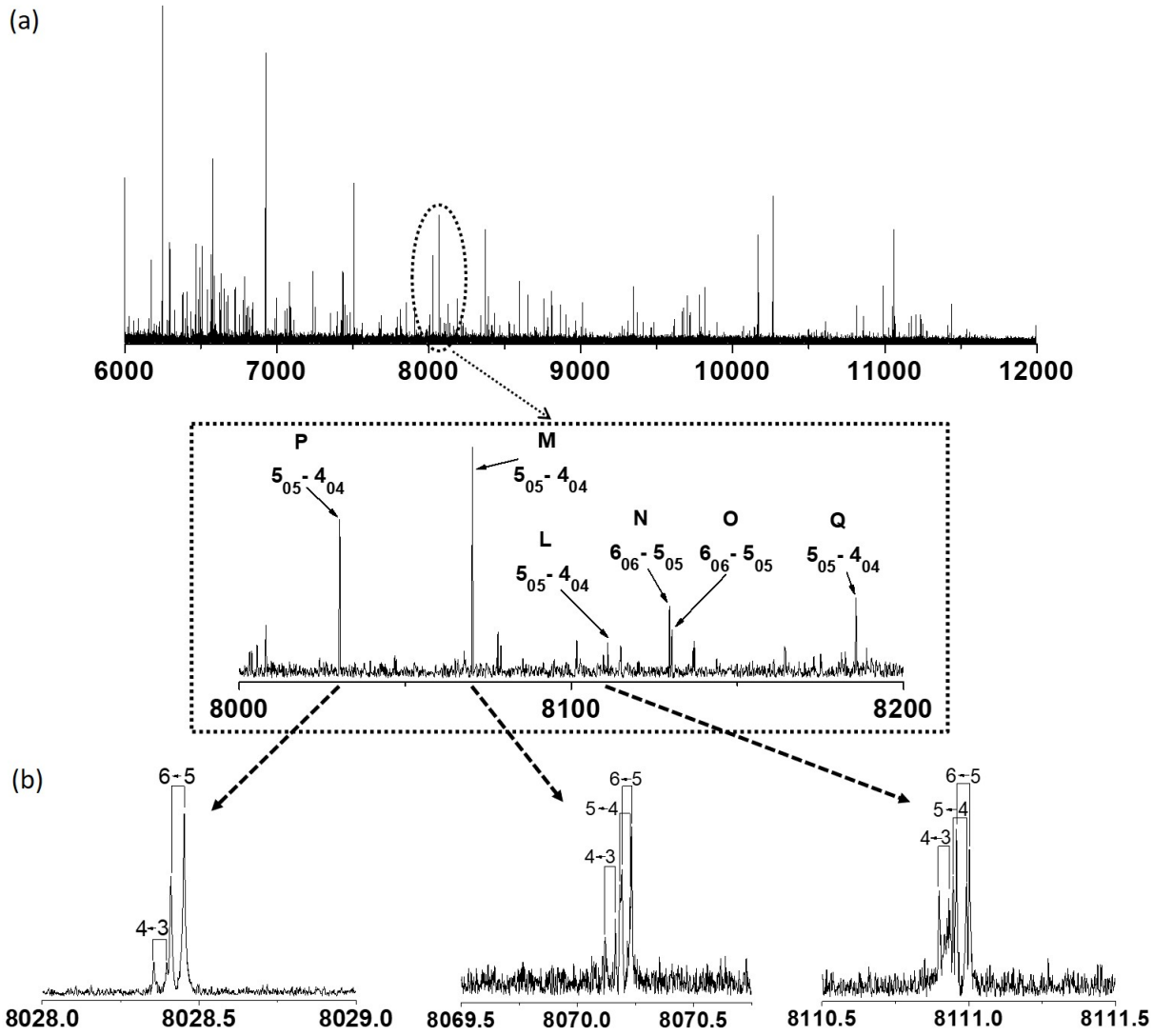


FIG. 5. (a) Broadband rotational spectrum of homocysteine in the 6.0-12.0 GHz frequency region recorded with the LA-CP-FTMW spectrometer. The inset shows some transitions belonging to different rotamers (see text for explanation). (b) The $5_{0,5} \leftarrow 4_{0,4}$ rotational transition for three of the observed rotamers of homocysteine (*P*, *M* and *L*, from left to right) obtained with the LA-MB-FTMW spectrometer. The hyperfine components (*F'* ← *F"*) are fully resolved and labeled using the **F=J+I** coupling scheme. Each component appears as a doublet because of the Doppler effect.

larger than the average error for the other conformers (9 MHz), we tried to understand whether such a behavior was related to any soft degrees of freedom. Inspection of Table S3 of the SI points out that the computed ϕ' and σ_1 dihedral angles are quite different in the *L,M* pair of conformers, whereas they are very similar in the *N,O* pair. Figure S3 in the SI shows that the effect of ϕ' is opposite on A_0 and χ, whereas σ_1 modifications of 10-15 degrees improve significantly the computed values of both spectroscopic parameters. Although small variations of the σ_1 dihedral angle have a limited effect on the electronic energy, the close agreement between rDSD and DZ-F12+CV geometries (see Tables S3 and S5) suggests a dynamical rather than structural origin of this behaviour. Furthermore, re-computation of the quadrupole coupling constants of the *L,N,O* conformers at the CCSD(T)/jun-cc-pV(T+d)Z level[32,33,55], also including CV contributions at the MP2/aug-cc-pwCV(T+d)Z level[37,56], led to results close to the rDSD counterparts (see Table I), thus excluding the possibility of an inaccurate evaluation of this property.

Conformational abundances in the molecular beam can be estimated from accurate measurements of the relative intensities of rotational transitions, provided that cooling in the supersonic jet brings the molecules to the lowest vibrational state of each conformer.[12] Under these circumstances, the measured intensities are proportional to the number density of the considered species and the corresponding electric dipole moment component. Assumed that the supersonic expansion determines effective rotational and vibrational cooling, the

post-expansion abundances of the different conformers should be governed by the relative enthalpies at 0 K rather than Gibbs free energies at room temperature. Considering either ΔH_0° or ΔG° only affects the relative population of the IIg$^-$g$^-$g$^-$ conformer, which results to be slightly more or less stable, respectively, than Ig$^-$tg$^-$ (see Table I). The experimental estimates are in general agreement with the theoretical trends, but cannot disentangle such small differences. Furthermore, it has to be taken also into account that relaxation mechanisms can impair a full quantitative comparison[17]. Despite this caveat, the remarkable agreement obtained in this study confirms the added value of state-of-the-art QC computations.

III. CONCLUSION

In conclusion, an integrated experimental-computational strategy for the accurate structural and energetic characterization of the conformational landscape of flexible biomolecule building blocks has been validated using the challenging test case offered by homocysteine. This allowed us to accurately characterize six conformers and to understand why other low-energy conformers are not experimentally observable. Indeed, several additional structures correspond to true energy minima, but escape experimental detection due to fast relaxation to more stable conformers during the supersonic expansion. Overall, the present work provides the first global picture of the conformational behavior of the neutral form of homocysteine. Together with the intrinsic interest of the studied molecule, some general remarks are deserved, namely: (i) high-resolution spectroscopic studies can be guided, supported and complemented by computational simulations owing to the development of state-of-the-art AI and QC methodologies, thus leading to the set up of a powerful integrated strategy; (ii) as a consequence of point (i), a fully *a priori*, highly reliable analysis of the conformational landscape is today possible for flexible building blocks of biomolecules in the gas phase; (iii) the use of MW spectroscopy provides an unbiased and rigorous check of the developed computational approaches. Future perspectives rely on the implementation of the integrated experimental-computational strategy in a general, user friendly platform, with the aim of permitting its access also to non-specialists.

IV. SUPPLEMENTARY MATERIAL

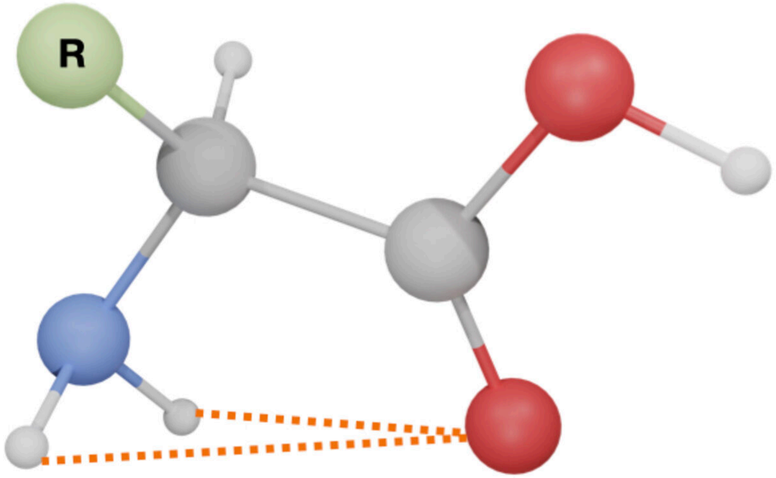
Computational (exploration and exploitation steps, tables with energetic and spectroscopic details, optimized geometries of all the stationary points) and experimental details (including observed lines and fitting errors for all the detected conformers).

ACKNOWLEDGMENTS

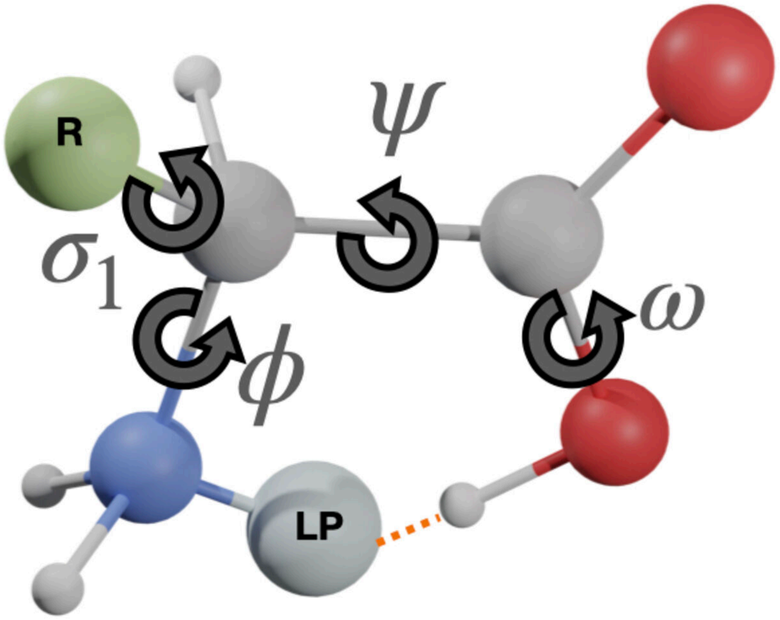
The financial fundings from Ministerio de Ciencia e Innovación (PID2019-111396GB-I00), Junta de Castilla y León (Grant VA244P20), Italian Ministry of University and Research (Grant 2017A4XRCA) and European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC-2013-SyG, Grant Agreement n. 610256 NANOCOSMOS, are gratefully acknowledged. We also thank the technical staff at SNS' SMART Lab. for managing the computational facilities.

[1] E. G. Robertson and J. P. Simons, "Getting into shape: Conformational and supramolecular landscapes in small biomolecules and their hydrated clusters," Phys. Chem. Chem. Phys. **3**, 1–18 (2001).
[2] J. L. Alonso and J. C. López, *Gas-Phase IR Spectroscopy and Structure of Biological Molecules* (Springer, 2015) pp. 335–401.
[3] V. Barone, S. Alessandrini, M. Biczysko, J. R. Cheeseman, D. C. Clary, A. B. McCoy, R. J. DiRisio, F. Neese, M. Melosso, and C. Puzzarini, "Computational molecular spectroscopy," Nat. Rev. Methods Primers **1**, 1–27 (2021).
[4] A. Karton, "A computational chemist's guide to accurate thermochemistry for organic molecules," WIRES, Comp. Mol. Sci. **6**, 292–310 (2016).
[5] J. Lupi, S. Alessandrini, V. Barone, and C. Puzzarini, "junChS and junChS-F12 models: Parameter-free efficient yet accurate composite schemes for energies and structures of noncovalent complexes," J. Chem. Theory Comput. **17**, 6974–6992 (2021).
[6] G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli, and V. Barone, "Unsupervised search of low-lying conformers with spectroscopic accuracy: A two-step algorithm rooted into the island model evolutionary algorithm," J. Chem. Phys. **153**, 124110 (2020).
[7] J. U. Grabow, S. Mata, J. L. Alonso, I. Peña, S. Blanco, J. C. López, and C. Cabezas, "Rapid probe of the nicotine spectra by high-resolution rotational spectroscopy," Phys. Chem. Chem. Phys. **13**, 21063–21069 (2011).
[8] I. León, E. R. Alonso, S. Mata, C. Cabezas, and J. L. Alonso, "Unveiling the neutral forms of glutamine," Angew. Chem. Int. Ed. Engl. **58**, 16002–16007 (2019).
[9] M. F. Wahab, S. Aslani, A. V. Mikhonin, J. L. Neill, and D. W. Armstrong, "Enhancing sensitivity for high-selectivity gas chromatography-molecular rotational resonance spectroscopy," Anal. Chem. **93**, 15525–15533 (2021).
[10] G. Schols, *Atomic and Molecular Beam Methods* (Oxford University Press, 1988).
[11] A. Lesarri, S. Mata, J. C. López, and J. L. Alonso, "A laser-ablation molecular-beam fourier-transform microwave spectrometer: The rotational spectrum of organic solids," Rev. Scient. Instr. **74**, 4799–4804 (2003).
[12] T. J. Balle and W. H. Flygare, "Fabry-Perot cavity pulsed Fourier transform microwave spectrometer with a pulsed nozzle particle source," Rev. Sci. Instrum. **52**, 33–45 (1981).
[13] G. G. Brown, B. C. Dian, K. O. Douglass, S. M. Geyer, S. T. Shipman, and B. H. Pate, "A broadband fourier transform microwave spectrometer based on chirped pulse excitation," Rev. Sci. Instrum. **79**, 053103 (2008).
[14] T. F. Miller, D. C. Clary, and A. J. H. M. Meijer, "Collision-induced conformational changes in glycine," J. Chem. Phys. **122**, 244323 (2005).
[15] R. S. Ruoff, T. D. Klots, T. Emilsson, and H. S. Gutowsky, "Relaxation of conformers and isomers in seeded supersonic jets of inert gases," J. Chem. Phys. **93**, 3142–3150 (1990).
[16] P. D. Godfrey and R. D. Brown, "Proportions of species observed in jet spectroscopy-vibrational energy effects: Histamine tautomers and conformers," J. Am. Chem. Soc. **120**, 10724–10732 (1998).
[17] G. M. Florio, R. A. Christie, K. D. Jordan, and T. S. Zwier, "Conformational preferences of jet-cooled melatonin: Probing trans- and cis-amide regions of the potential energy surface," J. Am. Chem. Soc. **124**, 10236–10247 (2002).
[18] M. E. Sanz, S. Blanco, J. C. López, and J. L. Alonso, "Rotational probes of six conformers of neutral cysteine," Angew. Chem. Int. Ed. Engl. **47**, 6216–6220 (2008).
[19] C. Cabezas, M. Varela, and J. L. Alonso, "The structure of the elusive simplest dipeptide gly-gly," Angew. Chem. Int. Ed. Engl. **56**, 6420–6424 (2017).
[20] J. L. Alonso, I. Peña, J. C. López, E. R. Alonso, and V. Vaquero, "The shape of the simplest non-proteinogenic amino acid α-aminoisobutyric acid (aib)," Chem. Eur. J. **25**, 2288–2294 (2019).

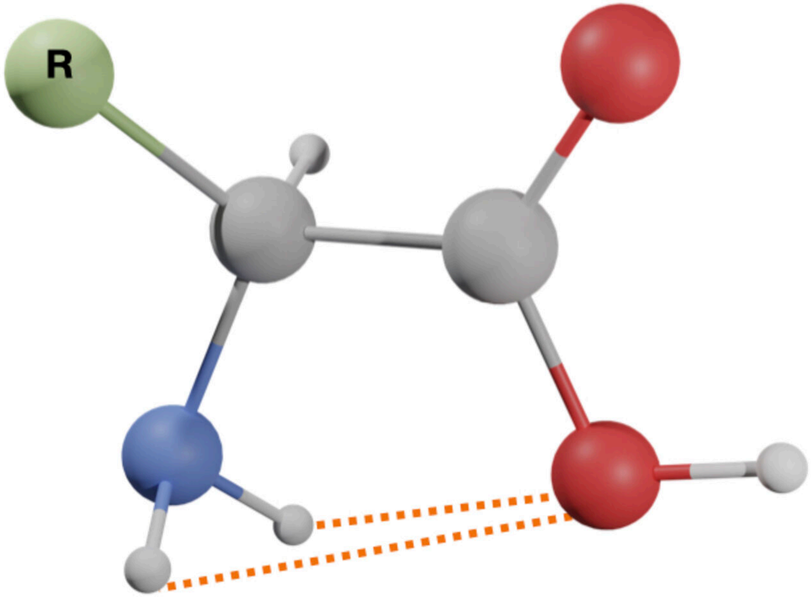
[21]E. R. Alonso, I. León, and J. L. Alonso, *Intra- and Intermolecular Interactions Between Non-Covalently Bonded Species* (Elsevier, 2020) pp. 93–141.
[22]H. V. L. Nguyen and I. Isabelle Kleiner, "Understanding (coupled) large amplitude motions: the interplay of microwave spectroscopy, spectral modeling, and quantum chemistry," Phys. Sci. Rev. **7**, 679–726 (2022).
[23]S. Bakels, M.-P. Gaigeot, and A. M. Rijs, "Gas-phase infrared spectroscopy of neutral peptides: Insights from the Far-IR and THz domain," Chem. Rev. **120**, 3233–3260 (2020).
[24]M. S. de Vries and P. Hobza, "Gas-phase spectroscopy of biomolecular building blocks," Ann. Rev. Phys. Chem. **58**, 585–612 (2007).
[25]J. J. Wilke, M. C. Lind, H. F. I. Schaefer, A. G. Csaszar, and W. D. Allen, "Conformers of gaseous cysteine," J. Chem. Theory Comput. **5**, 1511–1523 (2009).
[26]C. Cabezas, M. Varela, I. Peña, S. Mata, J. C. Lopez, and J. L. Alonso, "The conformational locking of asparagine," Chem. Comm. **48**, 5934–5936 (2012).
[27]D. L. Nelson, A. L. Lehninger, and M. M. Cox, *Lehninger principles of biochemistry* (Macmillan, 2008).
[28]F. Lazzari, A. Salvadori, G. Mancini, and V. Barone, "Molecular Perception for Visualization and Computation: The Proxima Library," J. Chem. Inf. Model. **60**, 2668–2672 (2020).
[29]S. Potenti, L. Spada, M. Fusè, G. Mancini, A. Gualandi, C. Leonardi, P. G. Cozzi, C. Puzzarini, and V. Barone, "4-fluoro-threonine: from diastereoselective synthesis to pH-dependent conformational equilibrium in aqueous solution," ACS Omega **6**, 13170–13181 (2021).
[30]C. Bannwarth, S. Ehlert, and S. Grimme, "GFN2-xTB–An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions," J. Chem. Theory Comput. **15**, 1652–1671 (2019).
[31]Y. Zhao and D. G. Truhlar, "Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions," J. Phys. Chem. A **109**, 5656–5667 (2005).
[32]T. H. Dunning, K. A. Peterson, and A. K. Wilson, "Gaussian basis sets for use in correlated molecular calculations. X. the atoms aluminum through argon revisited," J. Chem. Phys. **114**, 9244–9253 (2001).
[33]E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, and D. G. Truhlar, "Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions," J. Chem. Theory Comput. **7**, 3027–3034 (2011).
[34]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," J. Chem. Phys. **132**, 154104 (2010), https://doi.org/10.1063/1.3382344.
[35]G. Santra, N. Sylvetsky, and J. M. L. Martin, "Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4," J. Phys. Chem. A **123**, 5129–5143 (2019).
[36]A. D. Becke, "A new mixing of Hartree–Fock and local density-functional theories," J. Chem. Phys. **98**, 1372–1377 (1993), https://doi.org/10.1063/1.464304.
[37]C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," Phys. Rev. **46**, 618–622 (1934).
[38]J. Lupi, C. Puzzarini, C. Cavallotti, and V. Barone, "State-of-the-art quantum chemistry meets variable reaction coordinate transition state theory to solve the puzzling case of the H_2S + Cl system," J. Chem. Theory Comput. **16**, 5090–5104 (2020).
[39]T. B. Adler, G. Knizia, and H.-J. Werner, "A simple and efficient CCSD(T)-F12 approximation," J. Chem. Phys. **127**, 221106 (2007).
[40]G. Knizia, T. B. Adler, and H.-J. Werner, "Simplified CCSD(T)-F12 methods: Theory and benchmarks," J. Chem. Phys. **130**, 054104 (2009).
[41]H.-J. Werner, G. Knizia, and F. R. Manby, "Explicitly correlated coupled cluster methods with pair-specific geminals," Mol. Phys. **109**, 407–417 (2011).
[42]K. A. Peterson, T. B. Adler, and H.-J. Werner, "Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar," J. Chem. Phys. **128**, 084102 (2008).
[43]T. Helgaker, W. Klopper, H. Koch, and J. Noga, "Basis-set convergence of correlated calculations on water," J. Chem. Phys. **106**, 9639–9646 (1997).
[44]J. G. Hill, S. Mazumder, and K. A. Peterson, "Correlation consistent basis sets for molecular core-valence effects with explicitly correlated wave functions: The atoms B–Ne and Al–Ar," The J. Chem. Phys. **132**, 054108 (2010).
[45]M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko, and V. Barone, "Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: Validation and application to semi-rigid organic molecules," J. Phys. Chem. A **119**, 2058–2082 (2015).
[46]G. Ceselin, V. Barone, and N. Tasinato, "Accurate biomolecular structures by the nano-LEGO approach: Pick the bricks and build your geometry," J. Chem. Theory Comput. **17**, 7290–7311 (2021).
[47]R. M. Balabin, "The identification of the two missing conformers of gas-phase alanine: a jet-cooled raman spectroscopy study," Phys. Chem. Chem. Phys. **12**, 5980–5982 (2010).
[48]A. I. Jimenéz, V. Vaquero, C. Cabezas, J. C. López, C. Cativela, and J. L. Alonso, "The singular gas-phase structure of 1-aminocyclopropanecarboxylic acid (ac3c)," J. Am. Chem. Soc. **133**, 10621–10628 (2011).
[49]S. Blanco, M. E. Sanz, J. C. López, and J. L. Alonso, "Revealing the multiple structures of serine," Proc. Natl. Acad. Sci. U.S.A. **104**, 20183–20188 (2007).
[50]V. Barone, M. Biczysko, J. Bloino, and C. Puzzarini, "Accurate structure, thermodynamic and spectroscopic parameters from CC and CC/DFT schemes: the challenge of the conformational equilibrium of glycine," Phys. Chem. Chem. Phys. **15**, 10094–10111 (2013).
[51]S. Mata, I. Peña, C. Cabezas, J. C. López, and J. L. Alonso, "A broadband fourier-transform microwave spectrometer with laser ablation source: The rotational spectrum of nicotinic acid," J. Mol. Spectrosc. **280**, 91–96 (2012).
[52]I. León, E. R. Alonso, S. Mata, and J. L. Alonso, "A rotational study of the AlaAla dipeptide," Phys. Chem. Chem. Phys. **22**, 13867–13871 (2020).
[53]W. Gordy, R. L. Cook, and A. Weissberger, *Microwave molecular spectra*, Vol. 18 (Wiley New York, 1984).
[54]I. León, E. R. Alonso, S. Mata, C. Cabezas, M. A. Rodriguez, J. U. Grabow, and J. L. Alonso, "The role of amino acid side chains in stabilizing dipeptides: the laser ablation fourier transform microwave spectrum of Ac-Val-NH_2," Phys. Chem. Chem. Phys. **19**, 24985–24990 (2017).
[55]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories," Chem. Phys. Lett. **157**, 479 – 483 (1989).
[56]K. A. Peterson and T. H. Dunning, "Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al-Ar, and the first row atoms B-Ne revisited," J. Chem. Phys. **117**, 10548–10560 (2002).



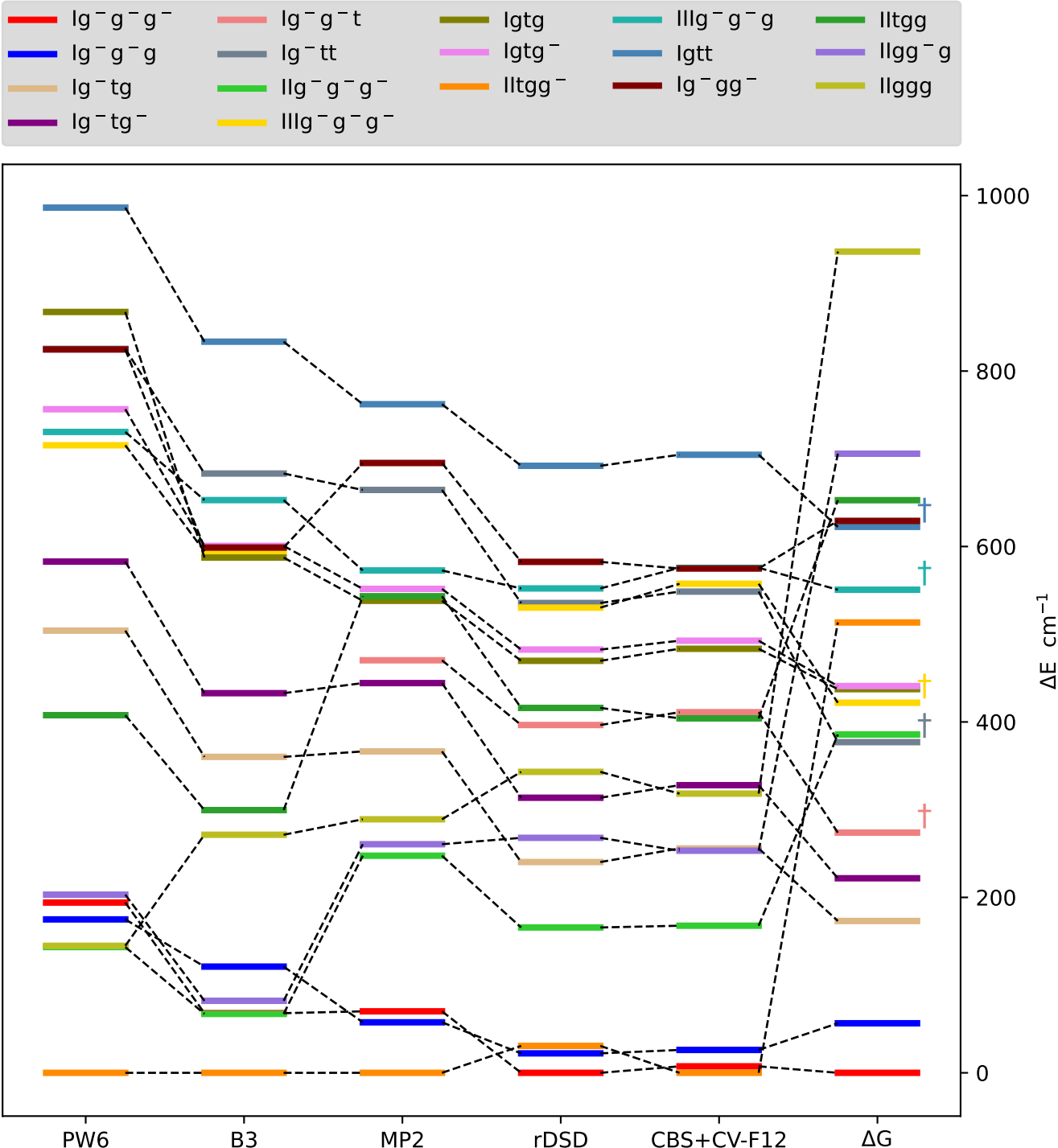
Type I

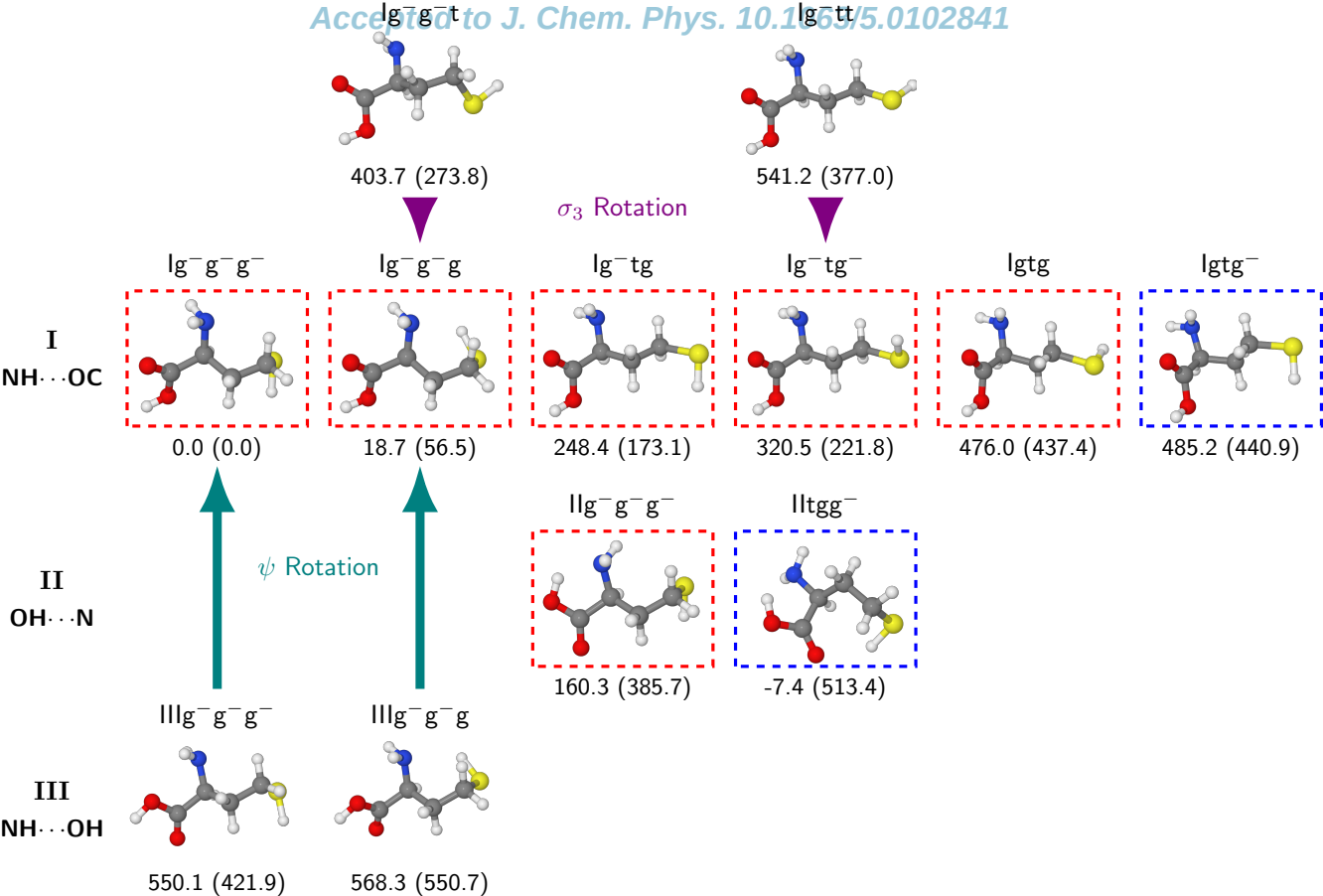


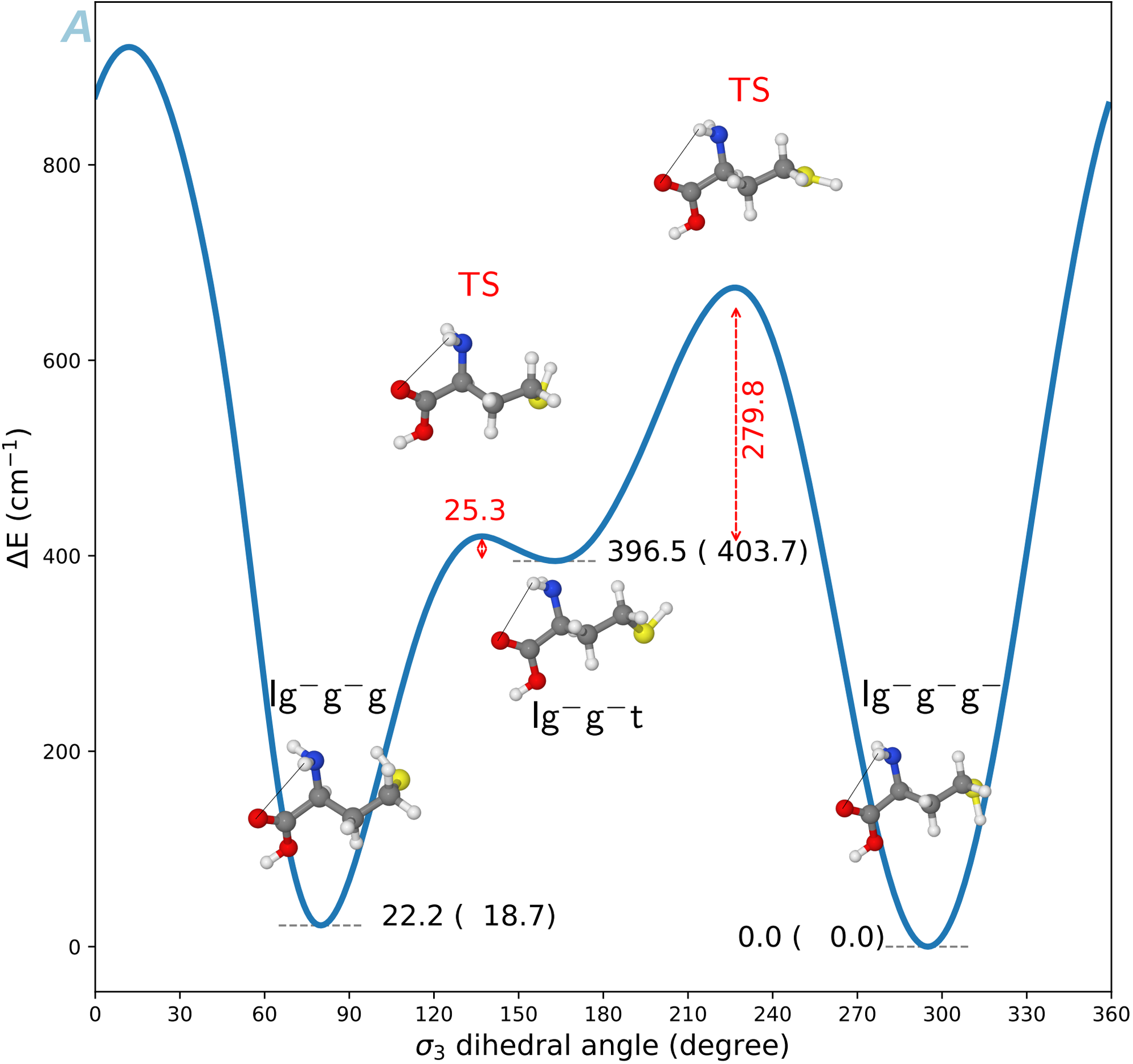
Type II

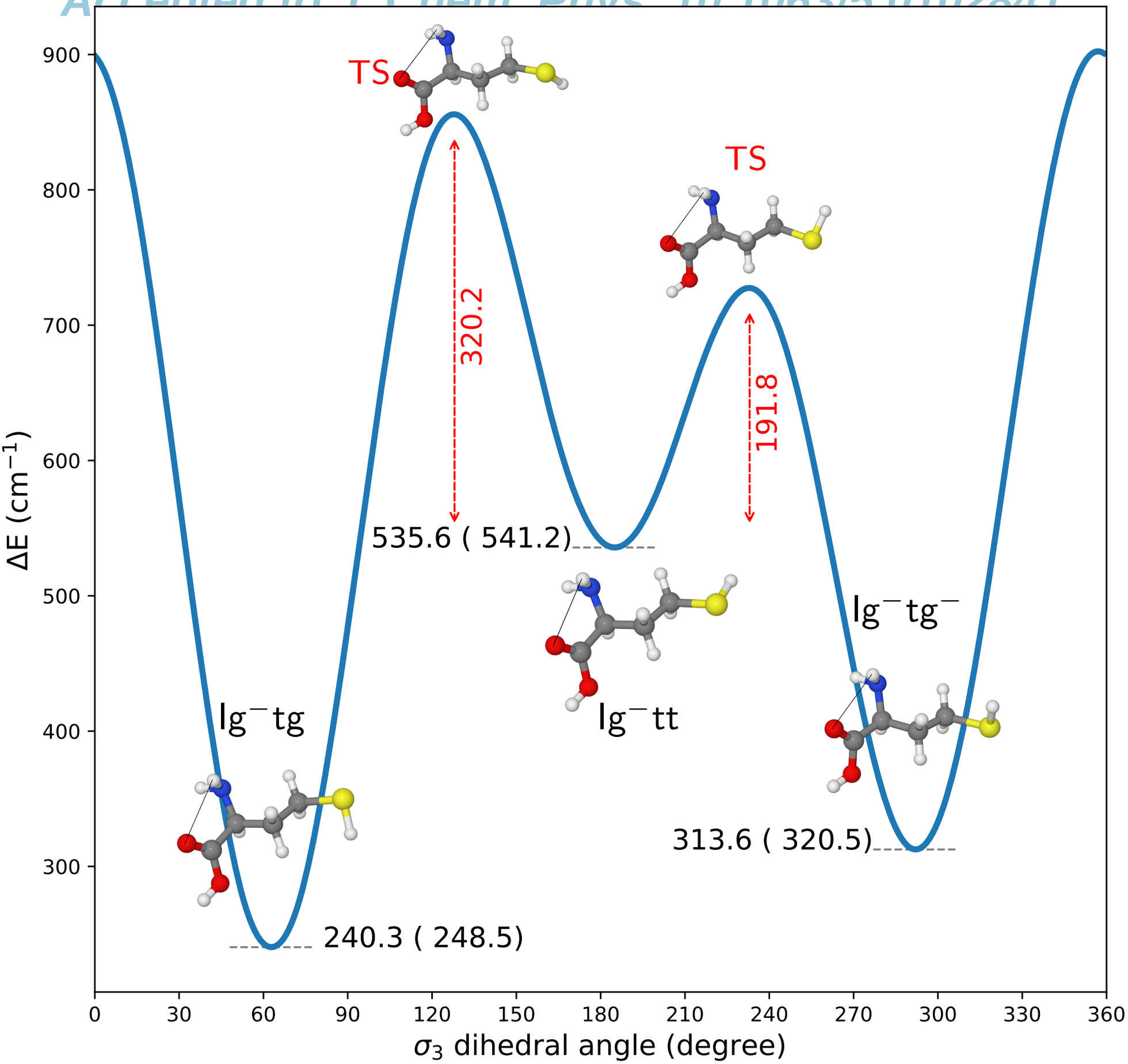


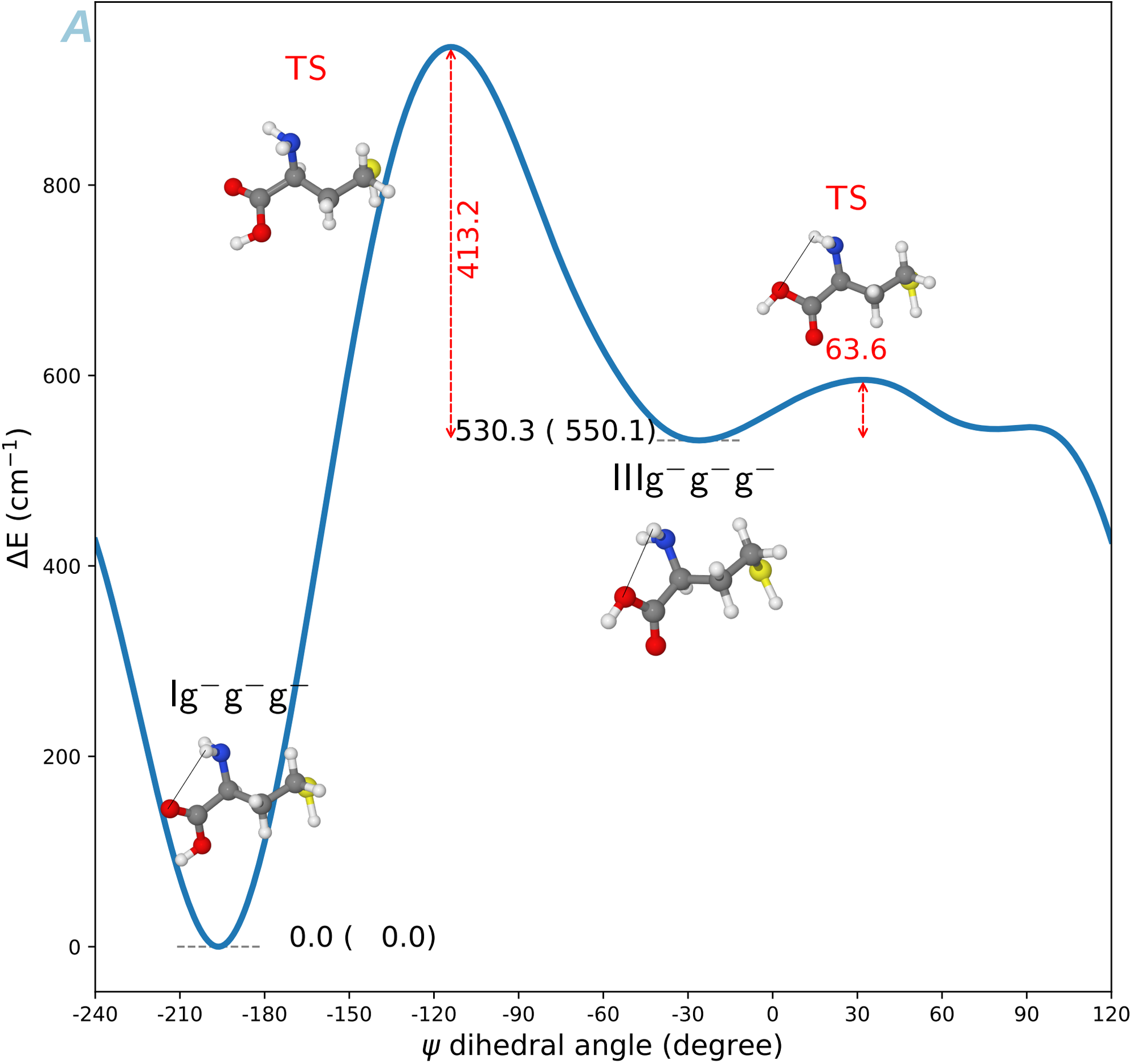
Type III

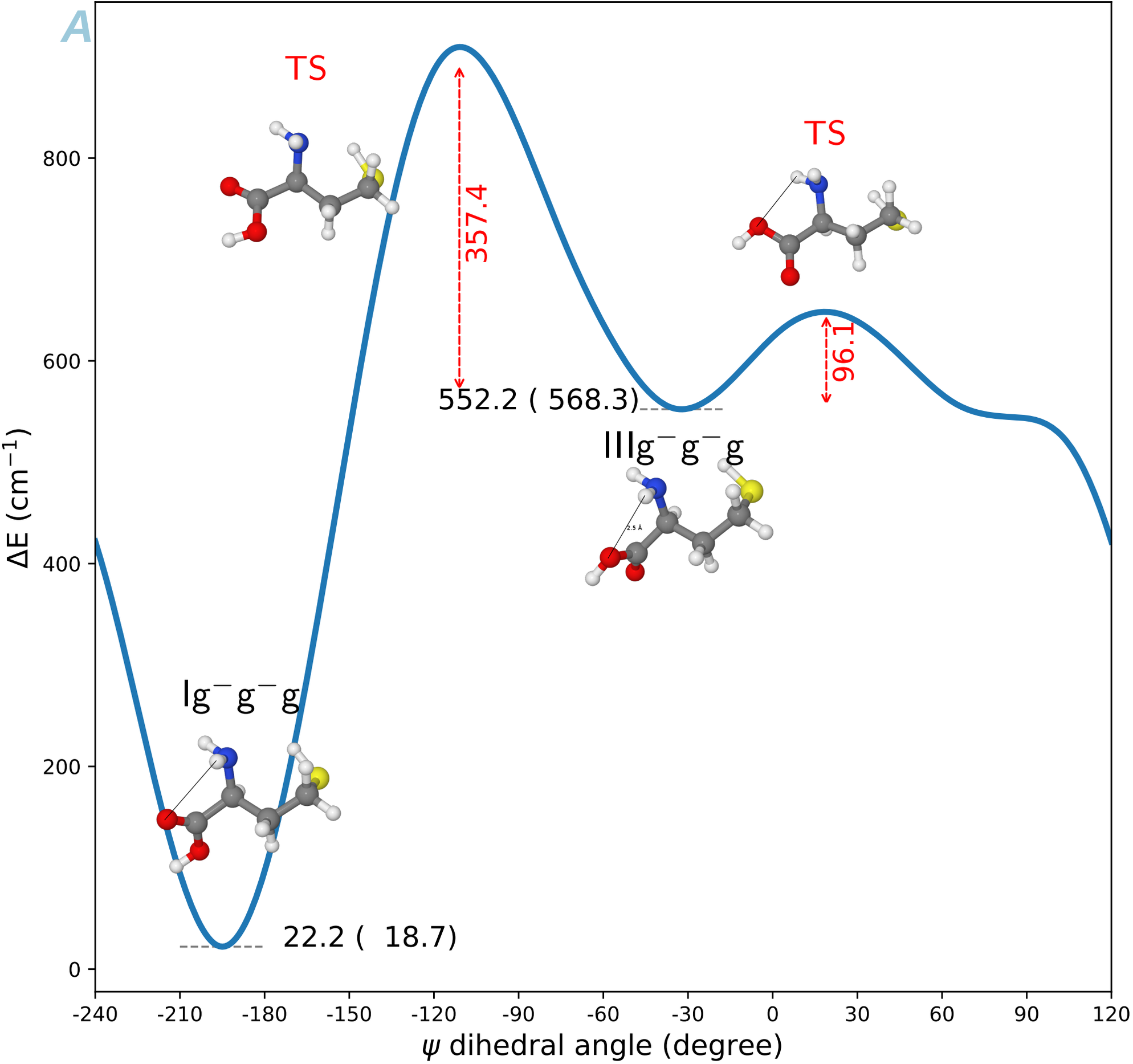

lg⁻g⁻g⁻
lg⁻g⁻g
lg⁻tg
lg⁻tg⁻
lg⁻g⁻t
lg⁻tt
llg⁻g⁻g⁻
lllg⁻g⁻g⁻
lgtg
lgtg⁻
lltgg⁻
llllg⁻g⁻g
lgtt
lg⁻gg⁻
lltgg
llgg⁻g
llggg
1000
800
600
400
200
0
ΔE cm⁻¹
PW6
B3
MP2
rDSD
CBS+CV-F12
ΔG

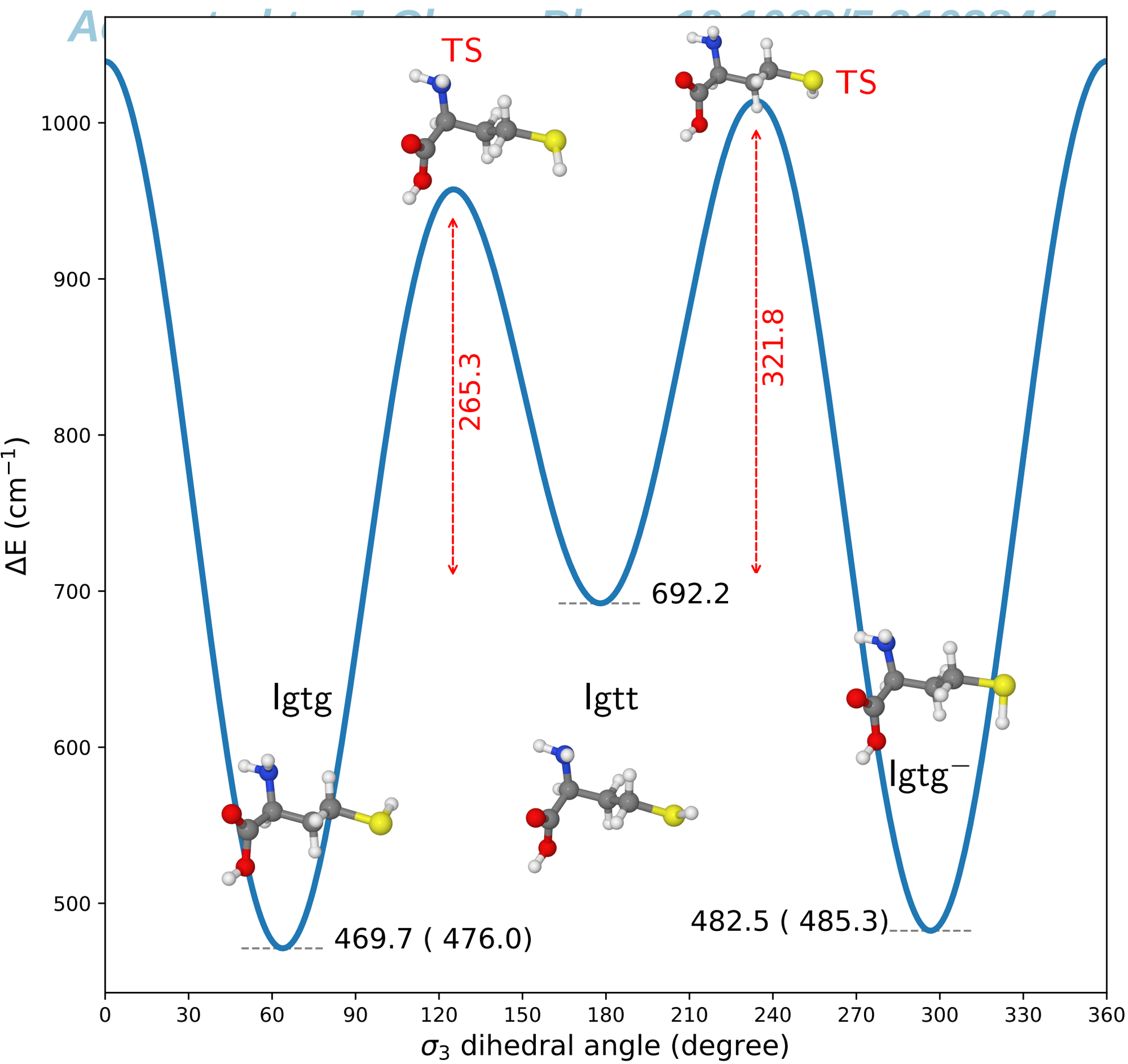


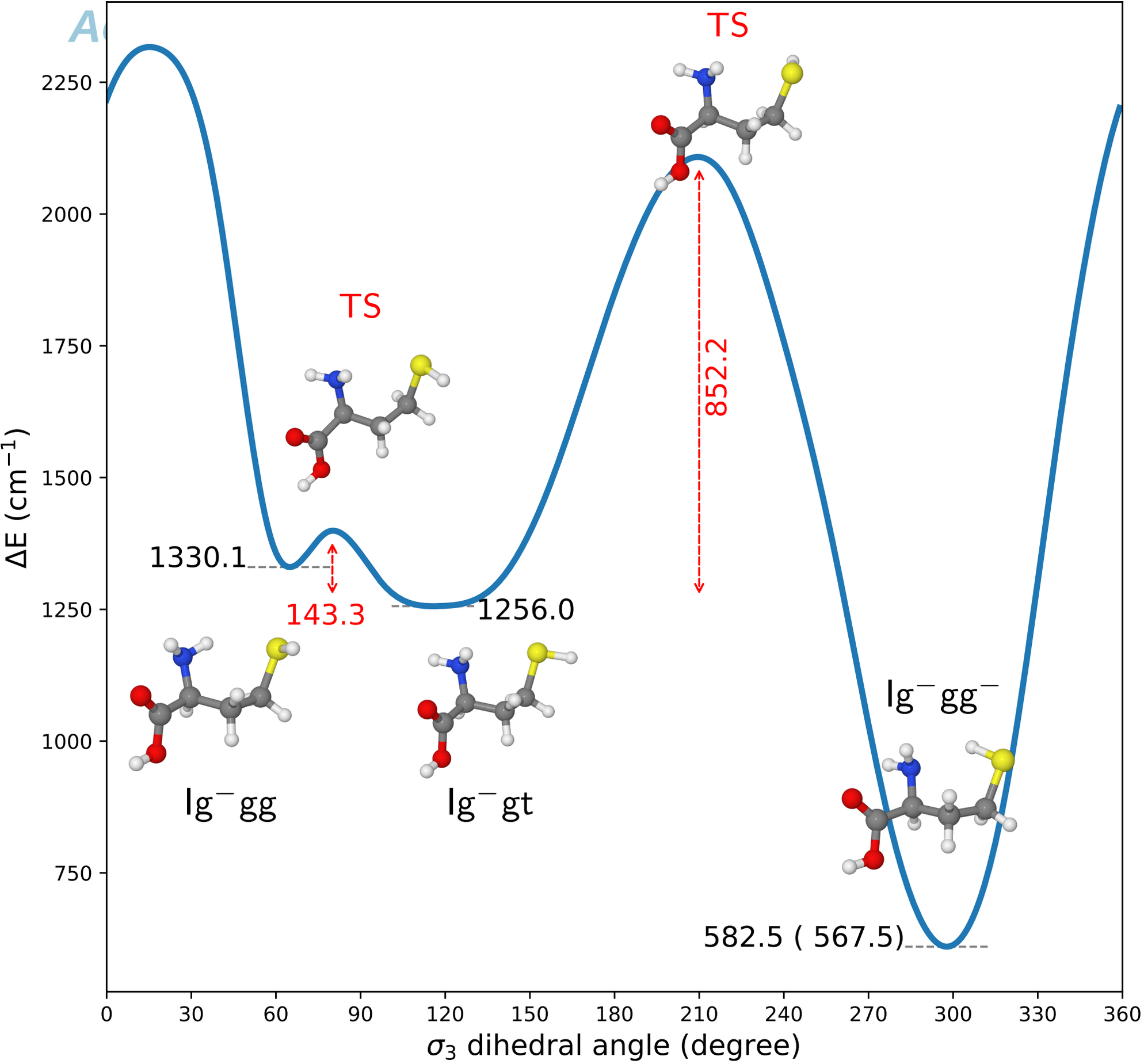

A
TS
TS
279.8
25.3
396.5 (403.7)
lg⁻g⁻g
lg⁻g⁻t
lg⁻g⁻g⁻
22.2 (18.7)
0.0 (0.0)
ΔE (cm⁻¹)
σ3 dihedral angle (degree)
0
200
400
600
800
0
30
60
90
120
150
180
210
240
270
300
330
360

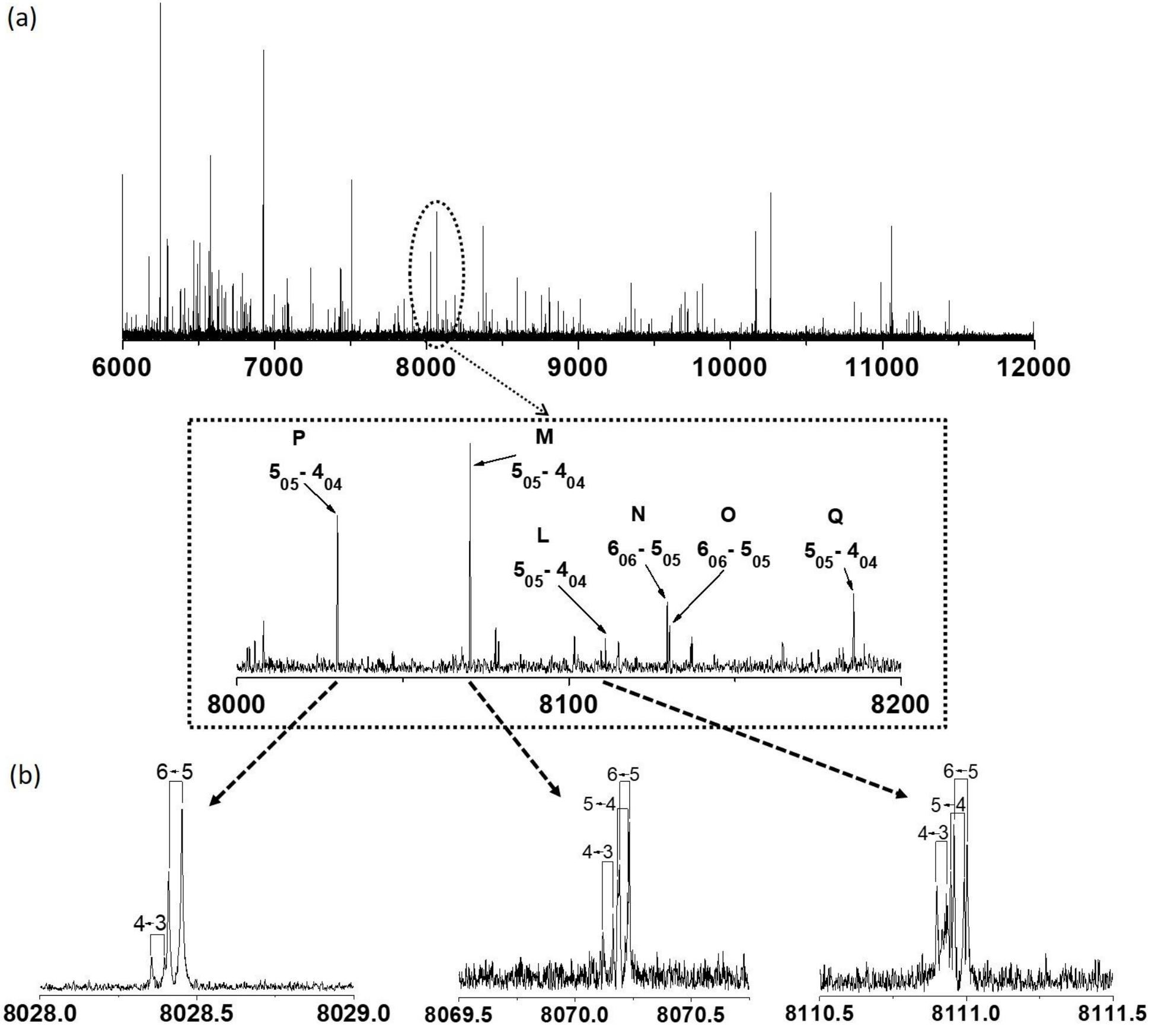



A
TS
413.2
TS
63.6
530.3 (550.1)
III g⁻ g⁻ g⁻
I g⁻ g⁻ g⁻
0.0 (0.0)
ΔE (cm⁻¹)
0
200
400
600
800
-240
-210
-180
-150
-120
-90
-60
-30
0
30
60
90
120
ψ dihedral angle (degree)


A
TS
TS
357.4
96.1
552.2 (568.3)
IIIg⁻g⁻g
Ig⁻g⁻g
22.2 (18.7)
2.5 Å
ΔE (cm⁻¹)
0
200
400
600
800
-240
-210
-180
-150
-120
-90
-60
-30
0
30
60
90
120
ψ dihedral angle (degree)


TS
TS
265.3
321.8
692.2
lgtg
lgtt
lgtg⁻
469.7 (476.0)
482.5 (485.3)
ΔE (cm⁻¹)
σ3 dihedral angle (degree)
1000
900
800
700
600
500
0
30
60
90
120
150
180
210
240
270
300
330
360


TS
TS
852.2
1330.1
143.3
1256.0
582.5 (567.5)
lg⁻gg
lg⁻gt
lg⁻gg⁻
ΔE (cm⁻¹)
σ3 dihedral angle (degree)
2250
2000
1750
1500
1250
1000
750
0
30
60
90
120
150
180
210
240
270
300
330
360


(a)
6000
7000
8000
9000
10000
11000
12000
P
5_{05} - 4_{04}
M
5_{05} - 4_{04}
L
5_{05} - 4_{04}
N
6_{06} - 5_{05}
O
6_{06} - 5_{05}
Q
5_{05} - 4_{04}
8000
8100
8200
(b)
6←5
4←3
8028.0
8028.5
8029.0
6←5
5←4
4←3
8069.5
8070.0
8070.5
6←5
5←4
4←3
8110.5
8111.0
8111.5