ORIGINAL PAPER

# Phenylisoserine in the gas-phase and water: Ab initio studies on neutral and zwitterion conformers

**Joanna E. Rode · Jan Cz. Dobrowolski · Joanna Sadlej**

**Abstract** The conformational landscape of phenylisoserine (PhIS) was studied. Trial structures were generated by allowing for all combinations of single-bond rotamers. Based on the B3LYP/aug-cc-pVDZ calculations 54 conformers were found to be stable in the gas phase. The six most stable conformers were further optimized at the B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVDZ levels for which characteristic intramolecular hydrogen bond types were classified. To estimate the influence of water on PhIS conformation, the IEF-PCM/B3LYP/aug-cc-pVDZ calculations were carried out and showed 51 neutral and six zwitterionic conformers to be stable in water solution. According to DFT calculations, the conformer equilibrium in the gas phase is dominated by one conformer, whereas the MP2 calculations suggest three PhIS structures to be significantly populated. Comparison of DFT and MP2 energies of all 57 structures stable in water indicates that, in practice, one zwitterionic and one neutral conformer

determine the equilibrium in water. Based on the AIM calculations, we found that for the neutral conformers in vacuum and in water, d(H...B) is linearly correlated with Laplacian at the H-bond critical point.

## Introduction

The molecules important in biology are remarkable, considering that they can display a wide variety of structures and functions. Most of the interesting functions carried out by these biomolecules are dependent on stable well-defined structures. This factor can be considered in the development of unnatural backbone modified peptides. β-amino acids [1] belong to such a group. The additional carbon atom incorporated into the backbone structure in natural α-amino acids results in unnatural oligomers containing side chains, which are significantly different from that found in natural compounds, but that can be recognized by side chain of receptors. The title molecule, phenylisoserine (Scheme 1), is one of such examples.

Among the new families of anticancer agents the natural product paclitaxel (Taxol®, Scheme 2) and its derivative docetaxel (Taxotere®) are used in the treatment of breast, ovarian cancers and lung carcinomas [2–4]. Paclitaxel is a natural product isolated from the bark of the western pacific yew tree. It contains 10-deacetylbaccatin-III (DBAC, Scheme 3) - a natural substance isolated from the leaves of the *Taxus baccata* species and phenylisoserine side chain [5, 6]. It has been shown that the side chain at C-13 and the diterpene moiety of taxol (DBAC) are both essential for the anticancer activity of the taxoid family. The baccatin III

J. E. Rode · J. Cz. Dobrowolski
Industrial Chemistry Research Institute,
8 Rydygiera Street,
01-793 Warsaw, Poland

J. Cz. Dobrowolski (✉) · J. Sadlej
National Medicines Institute,
30/34 Chełmska Street,
00-725 Warsaw, Poland
e-mail: janek@il.waw.pl

J. Sadlej (✉)
Faculty of Chemistry, Warsaw University,
1 Pasteura Street,
02-093 Warsaw, Poland
e-mail: sadlej@chem.uw.edu.pl
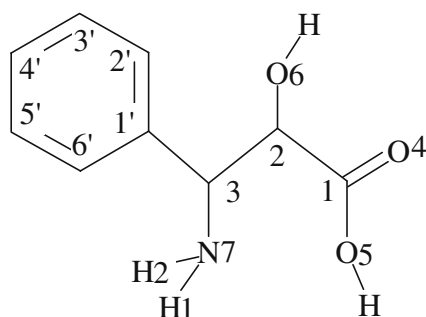
**Scheme 1** Phenylisoserine
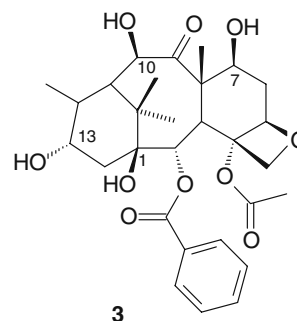


**Scheme 3** DBAC

tetracyclic ring system has been shown to be a rigid segment. Moreover, the structure-activity studies demonstrated that the presence of the phenylisoserine side chain (as N-benzoyl-(2R,3S)-3-phenylisoserine, Scheme 2) with correct absolute stereochemistry is significant for strong cytotoxicity. As all amino acids, the side chain of paclitaxel, is flexible and can adopt different conformations. The total synthesis of paclitaxel was done in the early 70s [7], however, because of complicated multistep synthesis and small yield it is not applied in the pharmaceutical industry. A semisynthetic way of producing paclitaxel seems to be the most promising: it is based on semiproduct (10-deacetylbaccatin-III) extracted from leaves of *Taxus baccata* to condensate it with N-benzoyl-(2R,3S)-3-phenylisoserine [8–10]. Therefore, study on conformational landscape of phenylisoserine may provide insights and may support practical efforts to obtain the very valuable medicine.

To study the conformational landscape of phenylisoserine as one of the β-amino acids is important not only because of paclitaxel and related compounds. It was found that β-amino acids can form stable α-helix, β-sheet, and hairpin motifs, similar to those in α-peptides. Indeed, while ca.15-20 residues are required to obtain a stable α-peptide α-helix, only six β-amino acids are necessary to form a stable β-peptide-based helix [11–14]. Because of its special importance in so many aspects, it is highly desirable to explore conformations of the β-amino acids thoroughly.

Likewise other amino acids, the phenylisoserine molecule has no symmetry and exhibits a significant internal
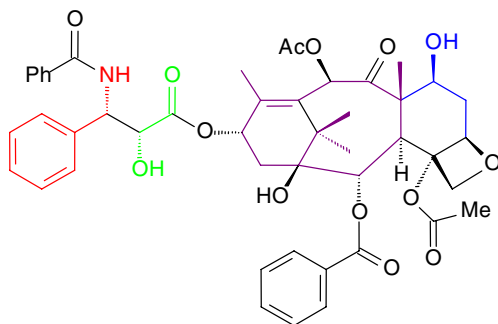


**Scheme 2** Taxol®

rotational degree of freedom due to rotations around the three single bonds in the central part of the carbon skeleton. Therefore, it is likely to occur as numerous low energy conformers. Amino acids are known to exist as zwitterions in the solid state and in solution and in the nonionized, neutral form in the gas phase. Study of the conformational behavior of the neutral amino acids is important for understanding the dependence of the peptides and protein backbone.

The gas-phase data are, in general qualitatively consistent with the matrix-isolation results. Several amino acids have been studied experimentally by using the IR inert-matrix-isolation method [15–24]. In the low-temperature matrices three conformers were found for glycine [15–18], three for valine [19], two for proline [20], four for serine [21], six for phenylalanine [24], and more than three for tryptophane [25]. Tyrosine has been studied only by ab initio method [26]. Recently, for cysteine and β-alanine we identified in the low-temperature matrices, 3 and 6 conformers, respectively [22, 23]. Also, in few laboratories, investigations on gaseous amino acids have been carried out by microwave spectroscopy during the past two decades [26–41]. To the best of our knowledge, so far, phenylisoserine has not been studied either by matrix-isolation IR spectroscopy nor by microwave spectroscopy.

The conformational characteristic of phenylisoserine has been the subject of only a few theoretical studies. In the most detailed theoretical study, the conformation of the side chain (with the glycolate fragment) have been investigated at the HF/6-31G(d,p) and PM3 levels [42]. Nine conformers have been found and it has been suggested that only in one of them, the most stable, the C3 phenyl is accommodated in a favorable position with its π orbital far away from the carbonyl oxygen lone pairs. Onsager continuum approach (SCFR) and supermolecular models with two water molecules have been used to study the conformation in polar solvent. The conformation similar to the molecule of the crystal structure of paclitaxel is now the most favorable [42] (and refs. therein). The second part of paclitaxol – DBAC have also been studied by the calculations. The benzoate and acetate are necessary to

allow binding to tubulin in the humane body. The importance of the oxetane ring was found by ZINDO method, which is consistent with experimental data [43].

The structure-activity relationship in the phenylisoserine moiety of the paclitaxel are dominated by the discussion of its different conformations named as polar and non-polar [42, 44, 45]. If the solvent is non-polar (CDCl$_3$), then one assumes that paclitaxel binding site on microtubules is hydrophobic and the molecule adopts a staggered conformation around the C2-C3 bond. On the other hand, if the solvent is polar, the C3 phenyl group is positioned close to C2 benzoyl and C4 acetyl groups [42]. To summarize, the conformation of the phenylisoserine moiety in different solvents depends on several factors.

Continuing our previous study on β-amino acids [22, 23] we undertook a detailed investigation on the conformation behavior phenylisoserine isolated in the gas-phase and in the water solvent. Since there has been no experimental attempt to determine either microwave and matrix isolation IR spectra of gaseous phenylisoserine nor x-ray structure, the calculated data may be useful in searching of both gas-phase and condensed phases conformers with the spectroscopic methods. Therefore, the goal of this study is twofold: (i) to estimate the conformer population of phenylisoserine in the gas-phase; (ii) to predict the conformer population of phenylisoserine in the water solution considering zwitterionic structures. In this way, a platform for a better distinction of various conformers in future liquid-phase experiments is settled. To the best of our knowledge, no systematic and accurate study on the conformational behavior of phenylisoserine in gas-phase and in water solution has been reported.

## Computational details

### Modeling the gas phase

All the calculations were performed by using the hybrid Becke three-parameter Lee–Yang–Parr DFT B3LYP functional [46, 47] which reliability in calculations of the ground state geometries has been widely assessed [48]. The aug-cc-pVDZ Dunning's [49, 50] (and for the most stable gas-phase conformers the aug-cc-pVTZ) basis sets were employed. These basis sets are known to be adequate to describe both organic molecules and their hydrogen-bonded systems. The conformational space of phenylisoserine was explored through a systematic variation of dihedral angles: d(O4C1O5H), d(O5C1C2C3), d(C1C2O6H), d(C1C2C3N7), d(C2C3N7H) and d(C2'C1'C3C2) (for atom labeling see Scheme 1). The stationary structures are found by ascertaining that all the harmonic frequencies are real. The relative abundances of the most stable conforma-

tions were then estimated by using the Gibbs free energy at 298.15 K, ∆G, relative to the most stable conformer. At the B3LYP/aug-cc-pVDZ level, 54 conformers were found to be stable in the gas phase.

All the calculations were performed by using the *Gaussian 03* [51] *and Gaussian 09* [52] programs.

### Modeling the aqueous phase

A systematic search for the stable zwitterions within the supermolecule approach goes far beyond this study; therefore, we decided to study the solvent effect by the IEF-PCM model [53, 54], which includes long-range interactions only. Continuum models are widely used nowadays to simulate solvent effects on conformation [55]. Among many methods used, the IEF-PCM method introduced in the late 1990s, was shown to simulate the aqueous phase fairly well [55]. Combined supermolecular-continuum approaches would probably be more successful, if they could be applied for phenylisoserine surrounded by hundreds of water molecules. To this aim, 54 gas-phase nonionized structures were reoptimized in water solvent at the IEFPCM/B3LYP/aug-cc-pVDZ level to yield 51 non-ionized and six stable zwitterion structures.

## Results and discussion

### Geometry and energy of the gas-phase phenylisoserine monomers

The phenylisoserine molecule has six single bonds around which free rotations may lead to as much as ca. 40 000 local minima. To generate the structures of it the relevant torsion angles were varied. Based on semiempirical AM1 search, 158 stable phenylisoserine conformers were appointed and reoptimized at the B3LYP/aug-cc-pVDZ level yielding 54 stable structures. Each of the phenyl-isoserine functional groups, carboxyl and alcoholic OH, C=O, and NH$_2$, may operate as both a hydrogen bond donor or acceptor. Thus, because of internal hydrogen bonds some conformers are stabilized more than the other. Six low-energy structures of phenylisoserine (Fig. 1, Table 1) were further studied by using the B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVDZ methods. They were found to be in the 3.5 kcal mol$^{-1}$ energy range relative to the most stable form. Thermochemical properties at 298.15 K were computed for all the conformers and the relative abundances were estimated for six of them. Figures, energetic and selected dihedral angles of all 54 phenylisoserine conformers optimized at the B3LYP/aug-cc-pVDZ level are available in the electronic supplementary materials (Tables 1SI and 2SI, Fig. 1SI). In Table 3SI Cartesian

**1** (0.0) [0.0]    **2** (2.47) [2.35]    **3** (2.99) [2.32]    **4** (3.08) [2.76]    **5** (3.22) [2.65]    **6** (3.42) [2.83]
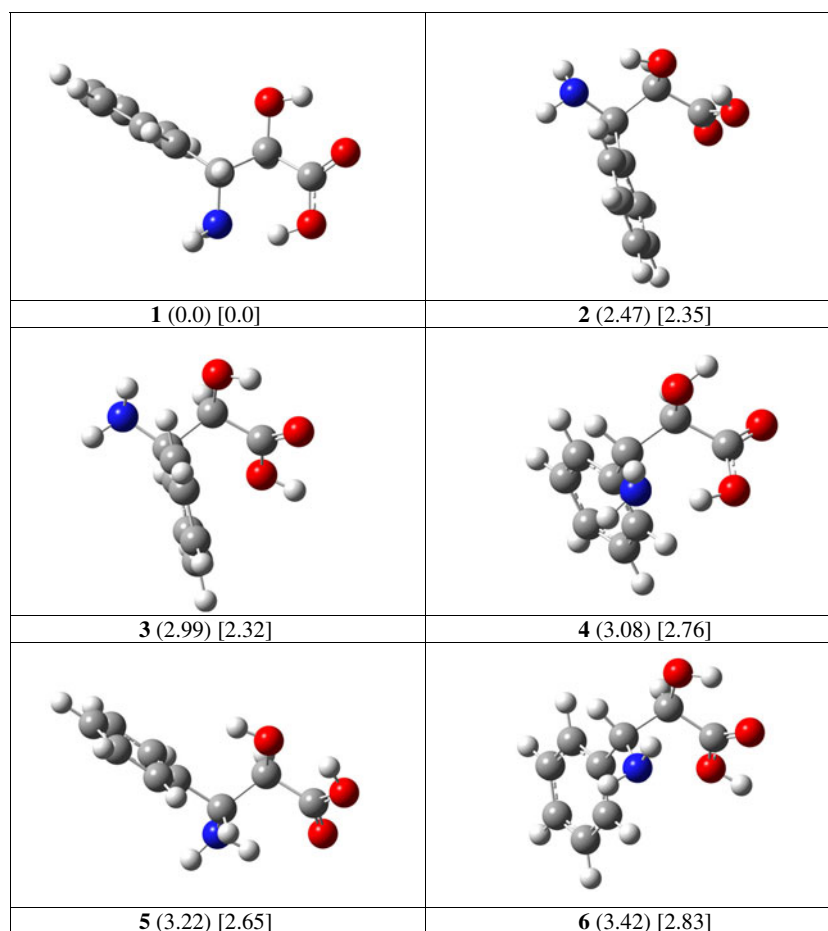
coordinates are gathered for the six most stable structures calculated at the B3LYP/aug-cc-pVDZ level.

Now let us compare the energetic results generated by using different methods and basis sets. The conformers are ordered according to the increasing total energy obtained at the B3LYP/aug-cc-pVDZ level (Table 1, Table 1SI). First and foremost, at the B3LYP/aug-cc-pVDZ level the most stable conformer **1** is more stable than the next one by as much as 2.5 kcal mol$^{-1}$. Upon the thermochemical corrections, the energetic order changes but only slightly: conformer **3** is located at the 2nd position and conformer **5** at the 4th position (Table 1). However, the order of conformers **4**, **5**, and **6** is slightly perturbed (Table 1). Conformer **1** is found to be still the global minimum when the basis set is increased to aug-cc-pVTZ and is still separated from the next one by more than 2.3 kcal mol$^{-1}$ (Table 1). The energetical picture undergoes significant change at the MP2/aug-cc-pVDZ level. First, conformer **2** becomes global minimum, separated from the next one, **1**, by only 0.13 kcal mol$^{-1}$. Moreover, the third stable conformer **3** is only 0.86 kcal mol$^{-1}$ less stable than **2**, and should be seriously taken into account if equilibrium conformer mixture in the gas phase is considered (Table 1).

The conformers may be characterized by the intramolecular hydrogen bonds (IHB). The presence of three hydrogen bond donors (NH$_2$, OH$_c$, OH$_a$) and four hydrogen bond acceptors (O=C, OH$_a$, NH$_2$, OH$_c$) allows for a wide range of hydrogen bond combinations and large number of stable forms. Let us now discuss the six lowest energy conformers. All the most stable six structures, exhibit the presence of the intramolecular hydrogen bonds (Fig. 1, Table 2). Two of them, **1** and **4**, exhibit the OH$_c$...NH$_2$ hydrogen bond, in which the OH$_c$ carboxylic group, as a proton donor, is in trans position to the C=O group. This IHB forms a six-membered ring (Fig. 1). In conformers **1** and **4** this intramolecular H-bond is quite short, i.e., ca. 1.74 Å (Table 2). This type of hydrogen bond OH$_c$...NH$_2$ is the same as that of phenylalanine [24], tyrosine [26] and tryptophan [56] global minimum. In structures **2** and **3**, the OH$_a$ alcoholic hydroxyl group is the proton donor and proton acceptor at the same time. In conformer **2** it is the proton donor to the NH$_2$ group and proton acceptor from the OH$_c$ group forming the OH$_a$...NH$_2$ and OH$_c$...OH$_a$ IHB, respectively. In conformer **3** the situation is different: the OH$_a$ group is the proton acceptor from the NH$_2$ group and the proton donor to the C=O group, forming NH$_2$...OH$_a$ and OH$_a$...O=C IHB, respectively. The other very important IHB

**Table 1** The B3LYP/aug-cc-pVDZ, B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVDZ optimized for the six most stable structures of phenylisoserine (PhIS) in gas phase, in water as neutral and zwitterion forms. The relative total energies ($\Delta E$, kcal mol$^{-1}$) and free Gibbs energies ($\Delta G$, kcal mol$^{-1}$) referred to the most stable conformer. The populations X (G) were calculated based on $\Delta G$ values (DFT) and $\Delta E$ values (MP2) separately for the 6 structures of each type

| gas phase | B3LYP/aug-cc-pVDZ | | | | B3LYP/aug-cc-pVTZ | | | | MP2/aug-cc-pVDZ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | X(G) |
| **1** | 0.00 | 0.00 | 0.00 | 93.61 | 0.00 | 0.00 | 0.00 | 94.83 | 0.13 | 37.29 |
| **2** | 2.47 | 2.36 | 2.35 | 1.77 | 2.33 | 2.36 | 2.52 | 1.34 | 0.00 | 46.07 |
| **3** | 2.99 | 2.52 | 2.32 | 1.86 | 2.65 | 2.35 | 2.34 | 1.84 | 0.86 | 10.77 |
| **4** | 3.08 | 3.10 | 2.76 | 0.89 | 3.21 | 3.34 | 3.33 | 0.34 | 1.61 | 3.05 |
| **5** | 3.22 | 2.85 | 2.65 | 1.07 | 2.97 | 2.77 | 2.78 | 0.87 | 2.04 | 1.48 |
| **6** | 3.42 | 2.96 | 2.83 | 0.79 | 3.13 | 2.83 | 2.85 | 0.77 | 2.09 | 1.35 |
| neutral in water | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | X(G) |
| **1w** | 0.00 | 0.00 | 0.00 | 97.86 | 0.00 | 0.00 | 0.00 | 98.20 | 0.00 | 86.60 |
| **29w** | 2.51 | 2.52 | 2.39 | 1.74 | 2.59 | 2.54 | 2.48 | 1.50 | 1.96 | 3.18 |
| **4w** | 3.52 | 3.64 | 3.43 | 0.30 | 3.69 | 3.80 | 3.68 | 0.20 | 1.65 | 5.35 |
| **2w** | 4.46 | 4.62 | 4.23 | 0.08 | 4.28 | 4.46 | 4.20 | 0.08 | 1.74 | 4.61 |
| **10w** | 5.77 | 6.10 | 5.64 | 0.01 | 5.42 | 5.75 | 5.27 | 0.01 | 4.38 | 0.05 |
| **3w** | 6.26 | 5.95 | 5.23 | 0.01 | 5.88 | 5.65 | 5.18 | 0.02 | 3.57 | 0.21 |
| zwitterion in water | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | $\Delta E_0$ | $\Delta G$ | X(G) | $\Delta E$ | X(G) |
| **1z** | 0.00 | 0.00 | 0.00 | 99.76 | 0.00 | 0.00 | 0.00 | 99.78 | 0.00 | 98.59 |
| **2z** | 4.62 | 3.80 | 3.75 | 0.18 | 4.77 | 3.93 | 3.87 | 0.14 | 4.38 | 0.06 |
| **3z** | 4.75 | 5.13 | 5.09 | 0.02 | 4.73 | 5.18 | 5.04 | 0.02 | 4.23 | 0.08 |
| **4z** | 4.83 | 4.65 | 4.62 | 0.04 | 4.99 | 4.84 | 4.52 | 0.05 | 3.13 | 0.50 |
| **5z** | 5.79 | 5.98 | 5.89 | 0.00 | 5.82 | 6.04 | 5.81 | 0.01 | 3.13 | 0.50 |
| **6z** | 6.11 | 6.81 | 6.75 | 0.00 | 5.89 | 6.70 | 6.59 | 0.00 | 3.48 | 0.28 |

is the OH$_a$...O=C bond (ca. 2.0 Å) stabilizing the **1**, **3**, and **6** conformers by forming a five-membered ring (Fig. 1). The next IHB, OH$_c$...OH$_a$ is quite short (ca. 1.9 Å) and it is present in two conformers **2** and **5**. In the case of conformers **3**, **4**, **5**, and **6**, there is also an IHB in which NH$_2$ group is a proton donor to either OH$_a$ or O=C groups (Table 2).

The population of conformers at room temperature are shown in Table 1. In the gas phase the distribution has a peak: the most stable form **1** is at the DFT level populated in at least 93% (X(G) in terms of $\Delta G$/B3LYP) (Table 1). On the other hand, the populations based on total energies calculated at the MP2/aug-cc-pVDZ level indicate three conformers: **1**, **2**, and **3** to be stable in significant amounts of ca. 37%, 46%, and 11%, respectively. Thus the population analysis is definitely dependent on energy correlation factor, however, to go far beyond the level applied here seems now to be quite a difficult task for so many conformers.

Energy of the aqueous-phase phenylisoserine conformers

*Neutral forms*

To estimate the influence of water on phenylisoserine molecule conformations, all 54 conformers stable in the gas phase were reoptimized by using the IEF-PCM/B3LYP/aug-cc-pVDZ method. The aqueous surrounding affects both the shape of the conformer distribution curve and the energetic differences between PhIS conformers. It is noticeable that in the first six stable conformations there are four conformers stable in the gas phase, yet, in a different order. Conformer **1w** (1 stands for number of the conformer in the gas phase, w denotes water) remains the most stable one and is still separated from the next one by 2.4 kcal mol$^{-1}$ (in terms of $\Delta G$, Table 1). Furthermore, the conformers are ordered differently (Table 1); in particular, the energy differences between them are larger. Additionally, among the most stable conformers there are forms denoted as **29w** and **10w**, that have been much less populated in the gas phase. This confirms a significant effect of solvation on geometrical parameters and solvatation energies of solutes.

All the most stable six neutral structures exhibit the presence of the intramolecular hydrogen bonds (IHB, Fig. 2, Table 2). Out of these six neutral conformers, three exhibit the OH$_c$...NH$_2$ and two OH$_a$...NH$_2$ internal hydrogen bonds (Fig. 2, Table 2). The shortest OH$_c$...NH$_2$ hydrogen bond is found for **1w**, which is the most stable neutral form in water. It is noticeable, that all structures with the OH$_c$...NH$_2$ IHB (**1w**, **29w**, **4w**) are the most stable in water. The longest OH$_c$...NH$_2$ hydrogen bond is detected for the **2w** conformer, which is the fourth stable neutral form in water. Thus, the presence of the internal OH$_c$...NH$_2$ hydrogen bond is probably the main factor controlling conformer stability in water. Additionally, the distance of

**Table 2** The intramolecular hydrogen bonds present in the most stable six conformers of phenylisoserine calculated at B3LYP/aug-cc-pVDZ level in the gas phase and at IEF-PCM/B3LYP/aug-cc-pVDZ level in "water"

| H-Bond type | Conformers in the gas phase | | | | | | Neutral conformers in "water" | | | | | | H-Bond type | Zwitterions in "water" | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 1w | 29w | 4w | 2w | 10w | 3w | | 1z | 2z | 3z | 4z | 5z | 6z |
| $OH_c{\cdots}NH_2$ | 1.741 | | | 1.743 | | | 1.621 | 1.632 | 1.652 | | | | $H_2NH{\cdots}O_1{=}C$ | 1.565 | 1.454 | 1.661 | 1.522 | 1.620 | |
| $OH_a{\cdots}NH_2$ | | 2.162 | | | | | | | | 2.060 | 1.988 | | $H_2NH{\cdots}OH_a$ | | | | | | 2.263 |
| $OH_a{\cdots}O{=}C$ | 2.014 | | 2.088 | 2.924 | | 2.036 | 2.041 | | | | | 2.129 | $OH_a{\cdots}O_2{=}C$ | 1.919 | | 2.014 | 2.932 | | 1.868 |
| $OH_c{\cdots}OH_a$ | | 1.940 | | | 1.886 | | | | | 1.921 | 1.903 | | $OH_a{\cdots}O_1{=}C$ | | | | | 2.423 | |
| $HNH{\cdots}OH_a$ | | | 2.440 | | | 2.550 | | | | | | 2.501 | $H_2NH{\cdots}C1'(Phe)$ | 2.646 | 2.662 | 2.565 | 2.604 | 2.577 | 2.613 |
| $HNH{\cdots}O{=}C$ | | | | | 2.184 | | | | 2.600 | | | | $OH_a{\cdots}C1'(Phe)$ | | 2.634 | | | | |
| $OH_a{\cdots}C1'(Phe)$ | | | | | 2.374 | | | 2.536 | | | | | | | | | | | |

$OH_a$...C1'(Phe) is shorter in **29w** than in the other conformers. This suggests that the interaction with the ring π-electrons may also be an important factor. Figures, energies and selected dihedral angles of all 51 neutral PhIS conformers in water optimized at the IEF-PCM/B3LYP/aug-cc-pVDZ level are available in the electronic supplementary materials (Tables 4SI and 5SI, Fig. 2SI). Additionally, Table 3SI lists Cartesian coordinates of the six most stable neutral forms in water.
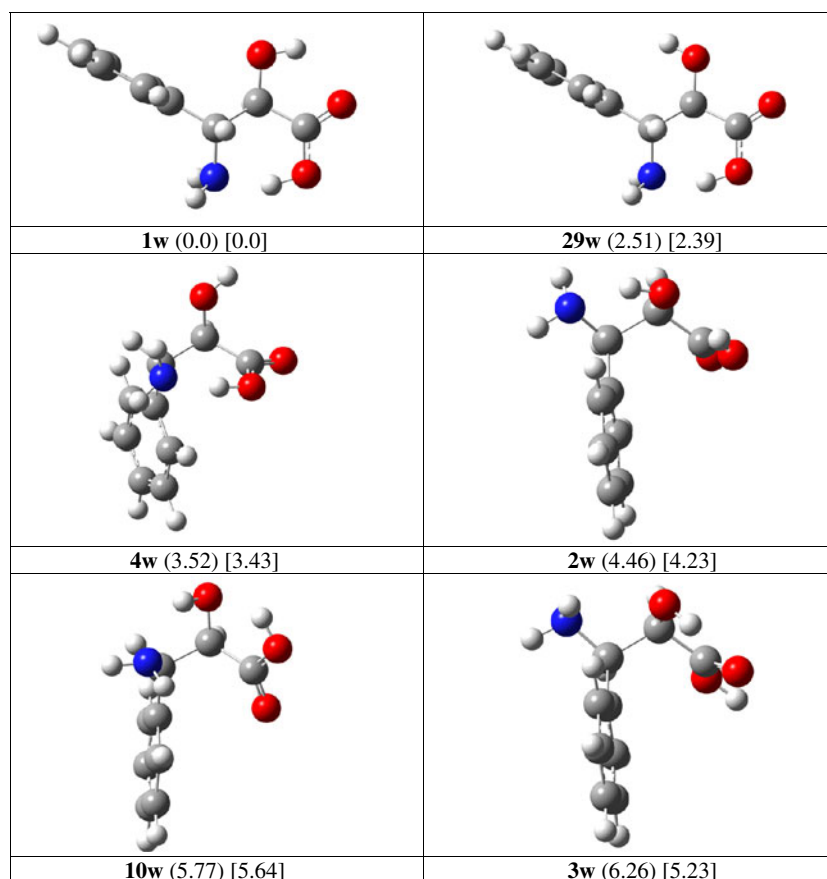
### Zwitterionic forms

The zwitterionic phenylisoserine conformers are unstable in the gas phase and thus, there are no such minima on the potential energy surface of an isolated monomer. Yet, in the water medium they become quite populated. Therefore, we carried out calculations for the zwitterion form of PhIS in the water phase. Out of 51 neutral structures stable in water only six appeared to exist as zwitterions (z) (Fig. 3, Table 1). Moreover, the zwitterionic **1z** conformer came out to be the most stable and definitely dominating in water. Energies and selected dihedral angles of the six zwitterion PhIS conformers in water, optimized at the IEF-PCM/B3LYP/aug-cc-pVDZ level, are available in the electronic supplementary materials (Tables 6SI and 7SI). Additionally, in Table 3SI Cartesian coordinates of these zwitterion structures are listed.

According to energetic ΔG results (Table 1), the most stable five zwitterion forms are those where the internal hydrogen bond is formed between the protonated amino group, $NH_3^+$, and the carboxylic $COO^-$ moiety. For example, in **1z** this distance is equal to 1.565 Å and is shorter by 0.056 Å than $OH_c$...$NH_2$ in neutral **1w** form and shorter by 0.178 Å than $OH_c$...$NH_2$ in **1** in the gas phase. Additionally, the interaction of the $NH_3^+$ group with the $OH_a$ group stabilizes conformer **6z**. Moreover, the interaction of $OH_a$ with the ring, $OH_a$...C1'(Phe), seems to stabilize all zwitterionic conformers.

Let us now discuss conformer population in water solution. For water phase it is necessary to consider both the neutral and zwitterionic forms because for all aminoacids there are equilibria of these forms in solution. However, first let us focus on each type separately. The population of neutral conformers in water (Table 1) shows the maximum peak for conformer **1w**, which is populated in at least 98% (X(G) in terms of ΔG/B3LYP). In contrast to the gas phase population, the MP2 level does not change the quantitative population picture. Indeed, according to MP2/aug-cc-pVDZ total energies conformer **1w** is populated in ca. 87% (Table 1). Similarly, the population of zwitterionic conformers in water (Table 1) shows the maximum peak for conformer **1z**, which is populated in ca. 99% (X(G) in terms of ΔG/B3LYP and

Fig. 2 The B3LYP/aug-cc-pVDZ optimized the most stable six neutral structures of phenyl-isoserine in water modeled by IEF-PCM method. The relative total energies ($\Delta E$, kcal mol$^{-1}$) and free Gibbs energies ($\Delta G$, kcal mol$^{-1}$) referred to the most stable conformer are given in parentheses and square brackets, respectively



**1w** (0.0) [0.0]     **29w** (2.51) [2.39]

**4w** (3.52) [3.43]     **2w** (4.46) [4.23]

**10w** (5.77) [5.64]     **3w** (6.26) [5.23]

98% at the MP2 level (Table 1). Now, when the most stable six neutral and six zwitterionic conformers are considered all together (Table 8SI) it appears that, at the DFT level, only the most stable neutral conformer **1w** and the most stable zwitterionic conformer **1z** are present in amounts of ca. 86% and 12%, respectively. Qualitatively, the MP2 calculations yield a similar result: in practice only the most stable neutral conformer **1w** and the most stable zwitterionic conformer **1z** are present. However, the MP2 level indicates that the zwitterions can be populated much more (ca. 36%) than predicted by the DFT calculations. Thus we expect both forms to be detected in equilibrium in water.
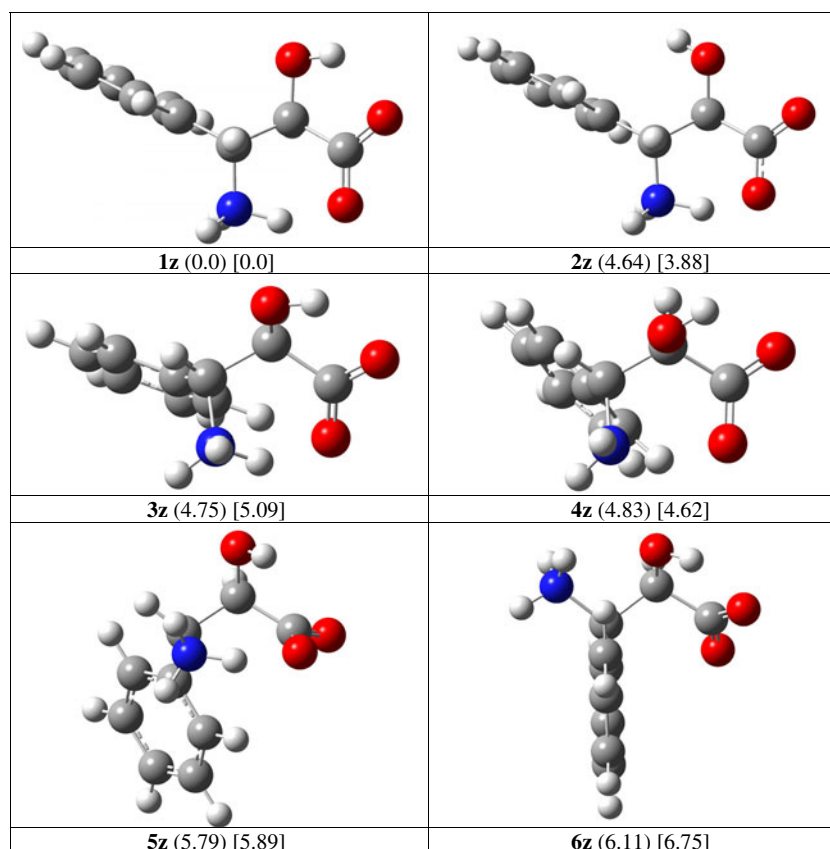
At the end let us comment on the question of how much stabilization is provided by each of the hydrogen bonds identified in each of the important conformers, both in vacuum and in water. Notice, that the studied conformers are stabilized by a variety of different hydrogen bonds and close contacts (see Table 2). Moreover, the repulsive interactions are also important for conformer diversification.

A more detailed inspection into the influence of intramolecular hydrogen bonds can be obtained by AIM analysis [57]. It was shown that Laplacian value of the electron density as well as electron density in bond and ring

critical points of the hydrogen bonds correlates with hydrogen bond strength [58–61]. Also, the hydrogen bond length correlates with its strength. Therefore, an interdependence between intramolecular hydrogen bond length and the Laplacian in its critical point is presented for all the most stable conformers in Fig. 4 (the values are listed in Table 9SI in Supplementary Information file). It is clear that the stronger the hydrogen bond the shorter it is and the more negative the Laplacian value. We found that for the neutral conformers in vacuum and in water, d(H...B) is linearly correlated with $\nabla^2\rho$ (R=0.930, Fig. 4), whereas the points corresponding to the zwitterions are spread out. Thus, the weak CH...OH$_a$ bond is placed in the top right of Fig. 4, the medium strong OH$_a$...NH$_2$ and OH$_a$...O=C bonds are located in the center, whereas the strongest OH$_c$...NH$_2$ bonds are concentrated in the bottom left of Fig. 4. We found that for zwitterions the hydrogen bond length is not correlated with $\nabla^2\rho$ value (Fig. 4), which is not surprising because of strong localization of the charges at the ends of the H-bonds.

As we stress above, the hydrogen bond strength is not a unique factor ordering the conformers stability. Thus, despite the fact that one is able to scale the intramolecular H-bond strength in the studied conformers, the other factors are hidden yet not negligible.

**Fig. 3** The B3LYP/aug-cc-pVDZ optimized the most stable six structures of phenylisoserine zwitterions in water modeled by IEF-PCM method. The relative total energies ($\Delta E$, kcal mol$^{-1}$) and free Gibbs energies ($\Delta G$, kcal mol$^{-1}$) referred to the most stable conformer are given in parentheses and square brackets, respectively



**1z** (0.0) [0.0]                **2z** (4.64) [3.88]

**3z** (4.75) [5.09]             **4z** (4.83) [4.62]

**5z** (5.79) [5.89]             **6z** (6.11) [6.75]

## Summary

Continuing our earlier studies on β-amino-acid conformations we found 54 conformers of penylisoserine to be stable in the gas phase, and 51 neutral and six zwitterionic conformers to be stable in water. The calculations were performed at the B3LYP/aug-cc-pVDZ, B3LYP/aug-cc-pVTZ and MP2/aug-cc-pVDZ levels. In the gas phase, in the most stable six conformers the following intramolecular hydrogen bonds are present: OH$_a$...O=C, OH$_c$...NH$_2$, OH$_a$...NH$_2$, OH$_c$...OH$_a$, HNH...O=C, and HNH...OH$_a$. The most stable conformer in the gas phase and in water phase is

**Fig. 4** The hydrogen bond lengths d(H...B) vs Laplacians $\nabla^2\rho$ in the hydrogen bond BCPs calculated for the most stable neutral conformers in vacuum (**v**) and in water (**w**) and for zwitterionic forms in water (**z**) based on B3LYP/aug-cc-pVDZ optimized structures
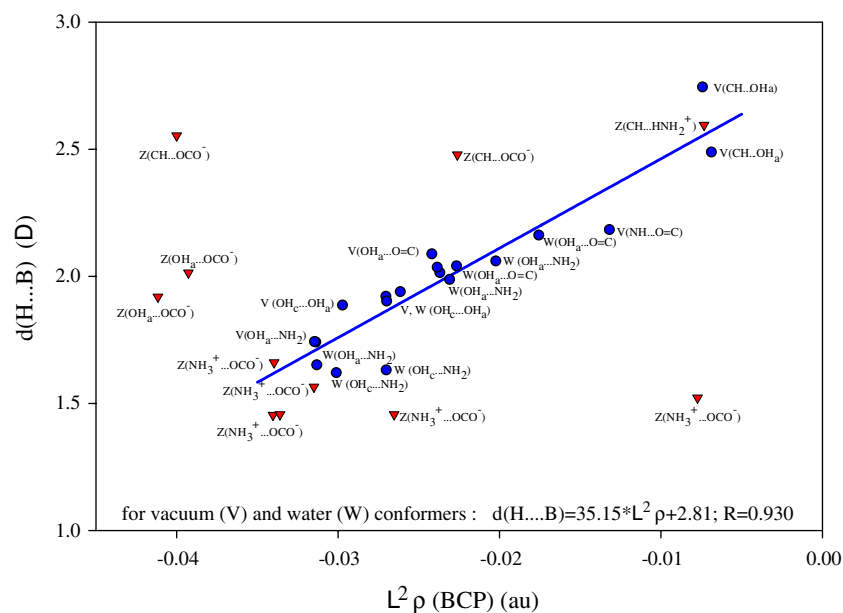
stabilized due to strong $OH_c...NH_2$ and $OH_a...O=C$ intramolecular hydrogen bonds. On the other hand, the zwitterionic forms in water are stabilized not only by interaction of $NH_3^+$ and $COO^-$ groups but also by interaction between $H_2NH$ and the phenyl ring. In the gas phase, at the DFT level, the most stable conformer is populated in at least 93%, whereas at the MP2/aug-cc-pVDZ level three conformers seems to be present in significant amounts of ca. 37%, 46%, and 11%. This shows that the population analysis is strongly dependent on electron correlation. Population analysis of the conformers in water (six neutral and six zwitterionic conformers) performed at the DFT and MP2 levels show concordantly that only the most stable neutral conformer and the most stable zwitterionic conformer are present. However, conformer ratio is again strongly dependent on electron correlation. Thus we expect both forms to be detected in equilibrium in water.

## References

1. Lelais G, Seebach D (2004) $\beta^2$-amino acids - syntheses, occurrence in natural products, and components of β-peptides. Biopolymers 76:206–243
2. Ojima I, Lin SN, Wang T (1999) Recent advances in the medicinal chemistry of taxoids with novel β-amino acid side chains. Curr Med Chem 6:927–954
3. Guenard D, Gueritte-Voegelein F, Potier P (1993) Taxol and taxotere: discovery, chemistry, and structure-activity relationships. Acc Chem Res 26:160–167
4. Nicolaou KC, Yang Z, Liu JJ, Ueno H, Nantermet PG, Guy RK, Claiborne CF, Renaud J, Couladouros EA, Paulvannan K, Sorensen FJ (1994) Total synthesis of taxol. Nature 367:630–634
5. DeGrado WF, Schneider JP, Hamuro Y (1999) The twists and turns of β-peptides. J Pept Res 54:206–217
6. Quinkert G, Egert E, Griesinger C (1995) In: Gutte B (ed) Peptides: synthesis, structure, and applications. Academic, San Diego
7. Nicolaou KC (2009) Inspirations, discoveries, and future perspectives in total synthesis. J Org Chem 74:951–972
8. Denis JN, Greene AE, Guenard D, Gueritte Voeglein F, Mangatal L, Potier P (1988) Highly efficient, practical approach to natural taxol. J Am Chem Soc 110:5917–5919
9. Denis JN, Kanazawa AM, Greene AE (1994) Taxotere® by esterification with stereochemically wrong (2s, 3s)-phenylisoserine derivatives. Tetrahedron Lett 35:105–108
10. Jezierska-Zięba M (2010) Badania nad asymetryczną syntezą estru t-butylowego kwasu 3-fenyloglicydowego–prekursora paklitakselu, Ph. D. Thesis, Industrial Chemistry Research Institute, Warsaw
11. Cheng RP, Gellman SH, DeGrado WF (2001) Beta-peptides; from structure to function. Chem Rev 101:3219–3232
12. Gellman SH (1998) Foldamers: a manifesto. Acc Chem Res 31:173–180
13. Hill DJ, Mio MJ, Prince RB, Hughes TS, Moore JS (2001) A field guide to foldamers. Chem Rev 101:3893–4011
14. Seebach D, Keiderling TA, Sebesta R, Campo MA, Beck AK (2004) How we drifted into peptide chemistry and where we have arrived at. Tetrahedron 60:7455–7506
15. Reva ID, Plokhotnichenko AM, Stepanian SG, Ivanov AYu, Radchenko ED, Sheina GG, Blagoi YP (1995) The rotamerization of conformers of glycine isolated in inert gas matrices. An infrared spectroscopic study. Chem Phys Lett 232:141–148, Erratum (1995) Ibid 235:617
16. Stepanian SG, Reva ID, Radchenko ED, Rosado MTS, Duarte MLTS, Fausto R, Adamowicz L (1998) Matrix-isolation infrared and theoretical studies of the glycine conformers. J Phys Chem A 102:1041–1054
17. Ivanov AY, Sheina GG, Blogoi YP (1999) FTIR spectroscopic study of the UV-induced rotamerization of glycine in the low temperature matrices (Kr, Ar, Ne). Spectrochim Acta A 55:219–228
18. Huisken F, Werhahn O, Ivanov AY, Krasnokutski SA (1999) The O–H stretching vibrations of glycine trapped in rare gas matrices and helium clusters. J Chem Phys 111:2978–2984
19. Stepanian SG, Reva ID, Radchenko ED, Adamowicz L (1999) Combined matrix-isolation infrared and theoretical dft and ab initio study of the nonionized valine conformers. J Phys Chem A 103:4404–4412
20. Stepanian SG, Reva ID, Radchenko ED, Adamowicz L (2001) Conformers of nonionized proline matrix-isolation infrared and post-hartree-fock ab initio study. J Phys Chem A 105:10664–10672
21. Lambie B, Ramaeckers R, Maes G (2004) Conformational behavior of serine: an experimental matrix-isolation FT-IR and theoretical DFT(B3LYP)/6-31++G** study. J Phys Chem A 108:10426–10433
22. Dobrowolski JCz, Jamróz MH, Kołos R, Rode JE, Sadlej J (2007) Theoretical prediction and the first IR-matrix observation of several L-cysteine molecule conformers. Chemphyschem 8:1085–1094
23. Dobrowolski JCz, Jamróz MH, Kołos R, Rode JE, Sadlej J (2008) IR low-temperature matrix and ab initio study on β-alanine conformers. Chemphyschem 9:2042–2051
24. Kaczor A, Reva ID, Proniewicz LM, Fausto R (2006) Importance of entropy in the conformational equilibrium of phenylalanine: a matrix-isolation infrared spectroscopy and density functional theory study. J Phys Chem A 110:2360–2370
25. Kaczor A, Reva ID, Proniewicz LM, Fausto R (2007) Matrix-isolated monomeric tryptophan: electrostatic interactions as nontrivial factors stabilizing conformers. J Phys Chem A 111:2957–2965
26. Zhang M, Huang Z, Lin Z (2005) Systematic ab initio studies of the conformers and conformational distribution of gas-phase tyrosine. J Chem Phys 122:134313(1-7)
27. Brown RD, Godfrey PD, Storey JWV, Bassez MP (1978) Microwave spectrum and conformation of glycine. J Chem Soc Chem Commun 547-548
28. Suenram RD, Lovas FJ (1978) Millimeter wave spectrum of glycine. J Mol Spectrosc 72:372–382
29. Suenram RD, Lovas FJ (1980) Millimeter spectrum of glycine. A new conformer. J Am Chem Soc 102:7180–7184
30. Godfrey PD, Firth R, Hatherley LD, Brown RD, Pierlot AP (1993) Millimeter- wave spectroscopy of biomolecules: alanine. J Am Chem Soc 115:9687–9691

31. Lovas FJ, Kawashima Y, Grabow J-U, Suenram RD, Fraser GT, Hirota E (1995) Microwave spectra, hyperfine structure, and electronic dipole moments for conformers I and II of Glycine. Astrophys J 455:L201–L204

32. McGlone SJ, Godfrey PD (1995) Rotational spectrum of a neurohormone: β-Alanine. J Am Chem Soc 117:1043–1048

33. Chapo CJ, Paul JB, Provencal RA, Roth K, Saykally RJ (1998) Is arginine zwitterionic or neutral in the gas phase? results from IR cavity ringdown spectroscopy. J Am Chem Soc 120:12956–12957

34. Lesarri A, Mata S, Cocinero EJ, Blanco S, Lopez JC, Alonso JL (2002) The structure of neutral proline. Angew Chem Int Ed 41:44673–44676

35. Blanco S, Lesarri A, Lopez JC, Alonso JL (2004) The gas-phase structure of alanine. J Am Chem Soc 126:11675–11683

36. Lesarri A, Cocinero EJ, Lopez JC, Alonso JL (2004) The shape of neutral valine. Angew Chem Int Ed 43:605–610

37. Lesarri A, Cocinero EJ, Lopez JC, Alonso JL (2005) Shape of 4 (S)- and 4(R)-hydroxyproline in gas phase. J Am Chem Soc 127:2272–2279

38. Lesarri A, Sanchez R, Cocinero EJ, Lopez JC, Alonso JL (2005) Coded amino acids in gas phase: The shape of isoleucine. J Am Chem Soc 127:12952–12956

39. Ilyushin VV, Alekseev EA, Dyubko SF, Motiyenko RA, Lovas FJ (2005) Millimeter wave spectrum of glycine. J Mol Struct 231:15–22

40. Alonso JL, Cocinero EJ, Lesarri A, Sanz ME, Lopez JC (2006) The glycine-water complex. Angew Chem Int Ed 45:3471–3474

41. Sanz ME, Lesarri A, Pena MI, Vanquero V, Cortijo V, Lopez JC, Alonso JL (2006) The Shape of β-Alanine. J Am Chem Soc 128:3812–3817

42. Milanesio M, Ugliengo P, Viterbo D (1999) Ab initio conformational study of the phenylisoserine side chain of paclitaxel. J Med Chem 42:291–299

43. Braga SF, Galvao DS (2002) A semiempirical study on the electronic structure of 10-deacetylbaccatin-III. J Mol Graph Model 21:57–70

44. Motorina IA, Huel Ch, Quiniou E, Mispelter J, Adjadj E, Grierson DS (2001) Phenylisoserine: a veraitile amino acid for the construction of novel β-peptide structures. J Am Chem Soc 123:8–17

45. Heider EM, Harper JK, Grant DM (2007) Structural characterization of an anhydrous polymorph of paclitaxel by solid-state NMR. PhysChemChemPhys 9:6083–6097

46. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652

47. Burke K, Perdew JP, Wang Y (1998) In: Dobson JF, Vignale G, Das MP (eds) Electronic density functional theory: recent progress and new directions. Plenum

48. Janoschek R (2001) Quantum chemical B3LYP/cc-pvqz computation of ground-state structures and properties of small molecules with atoms of z≤18 (hydrogen to argon). Pure Appl Chem 73:521–1553

49. Dunning TH Jr (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J Chem Phys 90:1007–1023

50. Kendall RA, Dunning TH, Harrison RJ (1992) Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J Chem Phys 96:6796–6806

51. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision C.02. Gaussian Inc, Wallingford, CT

52. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.1. Gaussian Inc, Wallingford, CT

53. Cances MT, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics. J Chem Phys 107:3032–3041

54. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999–3094

55. Cammi R, Mennucci B, Tomassi J (2003) In: Leszczyński J (ed) Computational chemistry, review of current trends, vol 8. World Scientific, Singapore

56. Compagnon I, Hagemeister FC, Antoine R, Rayane D, Broyer M, Dugourd P, Hudgins RR, Jarrold MF (2001) Permanent electric dipole and conformation of unsolvated tryptophan. J Am Chem Soc 123:8440–8441

57. Popelier PLA (2000) Atoms in molecules. An introduction. Prentice-Hall, Harlow, England

58. Espinosa E, Molins E, Lecomte C (1998) Chem Phys Lett 285:170–173

59. Alkorta I, Rozas I, Elguero J (1998) Chem Soc Rev 27:163–170

60. Grabowski SJ (2002) Monatsh Chem 133:1373–1380

61. Grabowski SJ (2001) Chem Phys Lett 338:361–366