HHS Public Access

Author manuscript

J Mol Struct. Author manuscript; available in PMC 2018 April 15.

Published in final edited form as:

J Mol Struct. 2017 April 15; 1134: 576–581. doi:10.1016/j.molstruc.2017.01.022.

Limiting Values of the one-bond C–H Spin-Spin Coupling Constants of the Imidazole Ring of Histidine at High-pH

Jorge A. Vila[1,*] and **Harold A. Scheraga**[2]

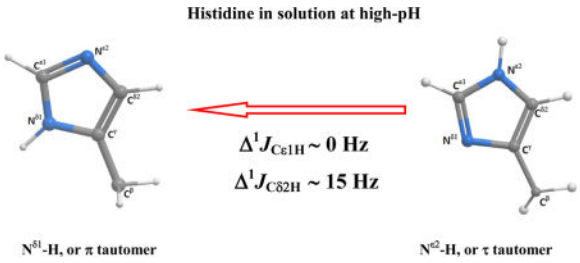
[1]IMASL-CONICET, Universidad Nacional de San Luis, Ejército de Los Andes 950, 5700-San Luis, Argentina

[2]Baker Laboratory of Chemistry, Cornell University, Ithaca, NY, USA

Summary

Assessment of the relative amounts of the forms of the imidazole ring of Histidine (His), namely the protonated (H^+) and the tautomeric $N^{\varepsilon 2}$-H and $N^{\delta 1}$-H forms, respectively, is a challenging task in NMR spectroscopy. Indeed, their determination by direct observation of the ^{15}N and ^{13}C chemical shifts or the one-bond C–H, $^1J_{CH}$, Spin-Spin Coupling Constants (SSCC) requires knowledge of the "canonical" limiting values of these forms in which each one is present to the extent of 100%. In particular, at high-pH, an accurate determination of these "canonical" limiting values, at which the tautomeric forms of His coexist, is an elusive problem in NMR spectroscopy. Among different NMR-based approaches to treat this problem, we focus here on the computation, at the DFT level of theory, of the high-pH limiting value for the $^1J_{CH}$ SSCC of the imidazole ring of His. Solvent effects were considered by using the polarizable continuum model approach. The results of this computation suggest, first, that the value of $^1J_{C\varepsilon 1H}$ = 205 ± 1.0 Hz should be adopted as the canonical high-pH limiting value for this SSCC; second, the variation of $^1J_{C\varepsilon 1H}$ SSCC during tautomeric changes is minor, *i.e.*, within ±1Hz; and, finally, the value of $^1J_{C\delta 2H}$ SSCC upon tautomeric changes is large (15 Hz) indicating that, at high-pH or for non-protonated His at any pH, the tautomeric fractions of the imidazole ring of His can be predicted accurately as a function of the observed value of $^1J_{C\delta 2H}$ SSCC.

Graphical abstract



*Corresponding author: vila@unsl.edu.ar.

Keywords

SSCC; one-bond spin-spin coupling constant; high-pH limiting values; Histidine tautomers; protonated form of His; solvent effects

Introduction

The role of Histidine (His) in many biological functions and activities is very well documented [1–4], and the reason for such versatility can be found in three distinctive properties of the His amino acid residue: (*i*) existence of two neutral, chemically-distinct forms ($N^{\delta 1}$–H and $N^{\varepsilon 2}$–H tautomers, also known as **π** and **τ** tautomers, respectively [5], and a charged H^+ form, shown in Figure 1), with one form favored over the other by the protein environment and pH [6]; (*ii*) the only ionizable residue (the charged form) that titrates around neutral pH has a pK° of 6.6 [7] and (*iii*) appearance of a population of ~50% in all enzyme active sites [8].

Despite these well-known facts, the physical properties of neutral His are difficult to characterize experimentally [9], making a proper determination of the fractions of the tautomeric forms of the imidazole ring of His a challenging problem in NMR spectroscopy. Among the experimental methods in current use, are those based on the observed: (*a*) ^{15}N chemical shifts [10–12]; (*b*) $^{13}C^{\gamma}$ and $^{13}C^{\delta 2}$ chemical shifts [6,13]; and (*c*) $^{1}J_{CH}$ Spin-Spin Coupling Constant (SSCC) of the imidazole ring of His [13–15]. As with any experimental method, *all* these mentioned approaches possess shortcomings: (*i*) the tautomeric fractions obtained by the ^{15}N-based method may differ significantly depending on the adopted canonical limiting values of the ^{15}N chemical shift [16]; (*ii*) the $^{13}C^{\gamma}$ and $^{13}C^{\delta 2}$ chemical shifts cannot always be observed. In fact, only 106 $^{13}C^{\gamma}$, versus 4,703 $^{13}C^{\delta 2}$, chemical shifts of the imidazole ring of histidine have been deposited in the Biological Magnetic Resonance data Bank (BMRB) [17]. Hence, problems in the determination of the chemical shifts for these nuclei, such as that for the ground state of His 40 in the protein Im7 [14], often prevent the use of this methodology; and (*iii*) the observed one bond C–H SSCC value at the high-pH limit is ambiguous, as will be discussed below.

It should be noted, from Figure 1, that there are *only* two one-bond C–H's, $^{1}J_{CH}$, SSCC's of the imidazole ring of Histidine, namely the $^{13}C^{\varepsilon 1}$–H, $^{1}J_{C\varepsilon 1H}$, and the $^{13}C^{\delta 2}$–H, $^{1}J_{C\delta 2H,}$ SSCC, respectively. Absences of an accurate value for each of these SSCC's, at the high-pH limit, gives rise to two different kinds of problems as explained below.

The first problem pertains to the use of $^{1}J_{C\varepsilon 1H}$ SSCC to determine the protonation fraction of His, e.g., to detect sparsely populated, short-lived, protein states [14]. In detail, the low-pH limiting value for $^{1}J_{C\varepsilon 1H}$ SSCC appears to be quite well defined (221 ± 1.0 Hz [14]), for the $^{1}J_{C\varepsilon 1H}$ SSCC pH-dependence of four titrating His residues (His 6, His 13, His 26, His 87) of the PLCC**γ** SH2 protein domain [14,18]. However, the observed high-pH limit for $^{1}J_{C\varepsilon 1H}$ SSCC differs among *five* His residues, of the PLCC**γ** SH2 protein domain, by up to ~6Hz [14], i.e., four titrating His residues converge to a high-pH limiting value of 207 ± 1.0 Hz while the remaining one (His 57), which is the *only* non-titrating His residue, shows an almost flat, pH-independent, value of 203 ± 1.0 Hz [14]. The existence of two

possible high-pH limiting values for $^1J_{C\varepsilon 1H}$ SSCC, namely 207 ± 1.0 Hz or 203 ± 1.0 Hz [14], is a source of ambiguity. A similar ambiguity is found for four non-titrating His residues of subtilisin BPN′ having $^1J_{C\varepsilon 1H}$ SSCC in the range of ~205Hz to ~209Hz [19].

The second problem pertains to a potential contradiction between an assumption [15] and existing evidence about the variation of $^1J_{C\delta 2H}$ SSCC upon tautomeric change [13]. Platzer *et al.* [15] had proposed that the variations of $^1J_{C\delta 2H}$ SSCC should be independent of the forms of the His tautomer, as for $^1J_{C\varepsilon 1H}$ SSCC. On the other hand, there is experimental evidence for L-histidine at pH 12 in 80% d6-ethanol/20% H_2O at −55 °C [13], showing the existence of a large, rather than a small change, of $^1J_{C\delta 2H}$ SSCC upon changes of the tautomeric forms.

As can be inferred from the above, a common problem, in both $^1J_{CH}$ SSCC's and the ^{15}N-based methods, is the need for accurate knowledge of the "canonical" limiting values of the imidazole ring of His in which each form of His, namely the protonated (H^+) and the tautomeric $N^{\varepsilon 2}$-H and $N^{\delta 1}$-H forms, respectively, is present to the extent of 100%. In this regard, the canonical limiting values of the ^{15}N chemical shift have already been analyzed [16] and, hence, here we will determine the high-pH limiting values for both $^1J_{CH}$ SSCC's of the imidazole ring of His. By doing this, we will be able to: (*i*) eliminate any possible ambiguity about the actual value of $^1J_{C\varepsilon 1H}$ SSCC; and, (*ii*) resolve a possible contradiction associated with the variations of $^1J_{C\delta 2H}$ SSCC upon changes in the relative amounts of the tautomeric forms.

Materials and Methods

Calculations details

All DFT-calculations of the two $^1J_{CH}$ SSCC's, of the imidazole ring of **His** in the Ac-**His**-NMe molecule, were carried out by using the Gaussian 09 suite of programs [20]; the Keywords used in Gaussian 09 (listed here for assessing the reproducibility of the calculations) were: "NMR=Mixed", with the options "CPHF=Conv=10" and "Int=ultrafine" [21]. Additional Keywords, such as "Readatoms", were also tested (see Results and Discussion section).

There are four contributions to the NMR coupling constants [22], namely, the Fermi Contact (**FC**), the Spin Dipolar (**SD**), the Paramagnetic Spin-Orbit (**PSO**), and the Diamagnetic Spin-Orbit (**DSO**) contribution, respectively. All four are known as the Ramsey contributions. For this reason, in each of the Tables, we have listed: (*i*) the sum (***Σ***) of *all* four Ramsey contributions to each DFT-computed $^1J_{CH}$ SSCC, as computed with the Gaussian 09 suite of programs; and (*ii*) the predicted values for the $^1J_{CH}$ SSCC's, listed in the last column of each Table, are obtained after adding an *ad-hoc* contribution of 5Hz, to the ***Σ*** term, due to the Zero Point Vibrational Contribution [23].

All the results in Table 1 correspond to gas-phase DFT calculations, while the ones in Tables 2–4 include DFT-calculations in both gas-phase and in the presence of solvent (see *Solvent Effects* section below). For the latter, all the results obtained with solvent are highlighted in italics and bold face.

Structural geometry of His

All the $^{1}J_{CH}$ SSCC calculations, at the DFT level of theory, were carried out by using the histidine geometry as defined in the Empirical Conformational Energy Program for Peptides and Proteins (ECEPP) in which their bond-lengths and bond-angles were parameterized by Momany *et al.* [24] and updated by Némethy *et al.* [25] by using a high-resolution (***R*** = 0.037) X-ray crystal structure of histidine. In this regard, (*i*) less than 10% of more than 500,000 crystal structures deposited in the Cambridge Structural Database have an ***R***-factor < 0.04 [26]; and (*ii*) a comparison between DFT-computed and observed $^{13}C^{\alpha}$ chemical shifts of two different structures of Ubiquitin [27], one that possesses non-regular geometry which has been obtained by X-ray diffraction at 1.8 Å resolution (PDB id 1UBQ [29]) and the other structure with regularized ECEPP geometry [27], $1UBQ_{reg}$, with both, in terms of rmsd, leading to 3.28 ppm and 2.38 ppm, respectively. Supplementary analysis of the agreement between these structures with the deposited electron density data of 1UBQ, in terms of the ***R***-factor, leads to 19.2% and 23.1% for 1UBQ and $1UBQ_{reg}$, respectively, while the all-heavy-atom rmsd between these two structures is only 0.142 Å.

Overall, the better agreement, in terms of $^{13}C^{\alpha}$ chemical-shifts, obtained with $1UBQ_{reg}$, rather than 1UBQ, is consistent with the well-known recognition that the bond lengths and bond angles of both X-ray and NMR-derived structures of proteins are not defined as highly accurately as in studies of small molecules [29] with which the ECEPP geometry has been parameterized [24,25].

Solvent effects

It is well known that sizable solvent effects are important for the computation of $^{n}J_{CH}$ SSCC *only* for $n = 1$ [30], with n being the number of intervening bonds between the Carbon and the Hydrogen. Consequently, dielectric solvent effects were taken into account during the DFT-calculations by using a Polarizable Continuum Model (PCM) [31,32] as implemented in the Gaussian package [20]. Because we are interested primarily in proteins in water, we decided to model the dielectric medium of such a system by using theoretical evidence [23] indicating that the average dielectric constants (D_i) inside and at-the-surface of a protein are around 6–7 and 20–30, respectively. Hence, the dielectric solvent effects on the DFT-computation of the $^{1}J_{CH}$ SSCC for the imidazole ring of His were computed by using the PCM approach with an *effective* D_i of 6.5 and 25.5, respectively.

It is worth noting that there is a *direct* and an *indirect* contribution of the solvent effects on the $^{1}J_{CH}$ SSCC [30], and both effects were taken into account here. The *direct* contribution, due to the polarization of the molecular electronic structure by the solvent, was considered here by using the PCM approach which, during the calculation of $^{1}J_{CH}$ SSCC, takes into account the induced surface charges derived from the boundary conditions at the cavity surface. The *indirect* effect, caused by the change of the molecular geometry due to the solvent, was also considered here because we used a high-resolution X-ray crystal structure of histidine and, as is well known, the crystals contain, on average, ~50% of solvent [34].

Determination of the factors affecting the computation of $^1J_{CH}$ SSCC

Among the factors affecting the computation of the four Ramsey contributions to $^1J_{CH}$ SSCC, we first analyzed their dependence on the functionals *and* on the basis sets chosen for the DFT calculations. This dependence is discussed below.

Selecting the functional

Maximoff *et al.* [35] ranked 20 functionals for their ability to reproduce the observed $^1J_{CH}$ SSCC of 31 chemical compounds containing 11 aromatic molecules. Among all the 20 functionals, we selected *only* those showing a mean absolute error < 3Hz (see Table 1). In Table 1, we highlight in bold face the results from those functionals giving closer predictions to the observed $^1J_{C\varepsilon 1H}$ SSCC for the non-titrating His 57 of the PLCCγ SH2 protein domain, namely ~203Hz [14]. These selected *best* functionals, namely, OPW91 and OPBE, are in agreement with the conclusion reached by Maximoff *et al.* [35] after testing 20 functionals. To decide which one of these two functionals should be adopted for our DFT calculations, the following complementary analysis was carried out. There is only one aromatic group ($C_4H_4N_2$), among 11 tested by Maximoff *et al.*, [35] possessing a C–H between 2 nitrogen atoms, in a similar chemical arrangement as that of $^{13}C^{\varepsilon 1}$ of the imidazole ring of His (see Figure 1). For this particular aromatic group ($C_4H_4N_2$), the results obtained by Maximoff *et al.* [35] with the OPW91 (0.2Hz) and the OPBE (2.4 Hz) functionals, in terms of the difference, Δ, between the observed and DFT-computed $^1J_{CH}$ SSCC, show that OPW91 matches the experimental data better than OPBE (see Supporting Information in Table 1 from Maximoff *et al.* [35]). This result indicates that OPW91 is the best functional with which to predict the observed values of $^1J_{CH}$ SSCC for the $C_4H_4N_2$ aromatic group and, hence, OPW91 rather than OPBE was chosen for the DFT calculations in this work.

Selecting a basis set

It is well known that $^1J_{CH}$ SSCC calculations are dominated by the Fermi-Contact (**FC**) contribution which depends strongly on the electron density close to the nuclei and, hence, demands high quality of the basis set chosen [36]. Consequently, *all* the DFT-calculations in this work were carried out by using the “aug-cc-pVTZ-J” basis set [37] which is specially designed to study NMR properties. In addition, there is an important consensus that this basis set gives satisfactory close agreement between observed and computed $^1J_{CH}$ SSCC [21,35,36,38,39,40].

A locally-dense basis set approximation [41] was also used to sense the response of the DFT-computed $_1J_{CH}$ SSCC to the basis set distribution chosen, i.e., different combinations of “aug-cc-pVTZ-J” basis sets on the nuclei of interest of the imidazole ring of **His**, and a “6-31G” or “3-21G” basis set on the remaining nuclei of the Ace-**His**-NMe molecule, were tested here; the results are shown in Tables 2 and 3.

Results and Discussion

In general, the gas-phase results from Tables 2 and 3 enable us to conclude that the computed $^1J_{C\varepsilon 1H}$ SSCC with a locally-dense basis set approximation is indistinguishable

from the one obtained by using a *uniform* basis set distribution, i.e., a "aug-cc-pVTZ-J" basis set on *all* the nuclei of the Ac-**His**-NMe molecule, but with a considerable reduction in computational time. In addition, a comparison of the DFT-computed $^1J_{C\varepsilon 1H}$ SSCC values, from rows 4–7 of Table 2, shows that the results do not depend on the ϕ, ψ, $\chi 1$ and $\chi 2$ torsional angles of His. It is worth noting that, for a given basis set distribution, the use of the keyword "*Readatoms*", that focuses the computation of $^1J_{CH}$ SSCC on the chosen pair of nuclei, can speed up the calculation by up to ~7 times.

Consideration of solvent effects, by using the PCM approach with two D_i, namely 6.5 and 25.5, increases the gas-phase results for $^1J_{C\varepsilon 1H}$ SSCC on the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomers by up to ~3Hz (see the highlight in italics in rows 8–9 and 5–6 of Tables 2 and 3, respectively). In view of all the assumptions and round off made, the most accurate DFT-computed value for $^1J_{C\varepsilon 1H}$ SSCC is 205Hz ± 1.0Hz. Without doubt, within ±1 Hz, the values of $^1J_{C\varepsilon 1H}$ SSCC for the tautomers of the imidazole ring of His, are indistinguishable. This likely high-pH limiting value (205Hz ± 1.0Hz) for $^1J_{C\varepsilon 1H}$ SSCC is within the range of the *lowest* high-pH limiting value observed from non-titrating His on both the subtilisin BPN′ and the PLCγ SH2 protein domain, explicitly ~205Hz and 203 ± 1.0 Hz, respectively. Moreover, it is also in very good agreement with the observed values (~205Hz for each of the His tautomers) obtained from NMR experiments for L-histidine at pH 12 in 80% d6-ethanol/20% H_2O at −55 °C [13].

The computed $^1J_{C\delta 2H}$ SSCC in the gas-phase and in the presence of solvent, by using the PCM approach, with an "aug-cc-pVTZ-J" basis set on *all* nuclei of the imidazole ring of His and a "6-31G" basis set on the remaining nuclei of the Ac-His-NMe molecule, are listed in Table 4 for both the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomer, respectively. In the presence of solvent effects (highlighted in italics in Table 4) the most remarkable of the results obtained for $^1J_{C\delta 2H}$ SSCC, is the existence of a large difference between His tautomers, namely 15Hz. This result is counter to the assumption made by Platzer *et al.* [15] that $^1J_{C\delta 2H}$ SSCC should behave like $^1J_{C\varepsilon 1H}$ SSCC upon tautomeric changes, i.e., showing minimal or no difference. Even more important, our result for the difference between tautomers, in terms of $^1J_{C\delta 2H}$ SSCC, is in excellent agreement with existing NMR-based evidence on L-histidine at pH 12 in 80% d6-ethanol/20% H_2O at −55 °C [13], showing that the difference in $^1J_{C\delta 2H}$ SSCC between His tautomers is indeed 15Hz.

Despite the excellent agreement mentioned above, there is a conflict between the observed and the DFT-computed value for $^1J_{C\delta 2H}$ SSCC for each of the tautomeric forms of the neutral His, i.e., the computed values in the presence of solvent for the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomer are, as shown in Table 4, 165Hz and 180Hz, respectively, while the observed values are in the reverse order, namely, 180Hz and 165Hz, respectively (see Table 2S of Sudmeier *et al.* [13]). An irrefutable test to resolve this discrepancy would be to repeat the NMR experiments on L-histidine at pH 12 in 80% d6-ethanol/20% H_2O at −55 °C. However, this is beyond the scope of this work. For this reason, and in order to validate that the predictions in Table 4 actually belong to the referred His tautomer, the following analyses were carried out, first a validation in terms of the observed chemical shifts, second, a computation of the difference in the $^1J_{C\delta 2C\gamma}$ SSCC values ($\Delta_{C\delta 2C\gamma}$) upon tautomeric

changes and, third, a computation of the $^{1}J_{C\delta N\varepsilon 2}$ the and $^{2}J_{C\delta N\delta 1}$ SSCC in the $N^{\varepsilon 2}$-H and in the $N^{\delta 1}$-H tautomers, respectively.

The $^{13}C^{\gamma}$ and $^{13}C^{\delta 2}$ shieldings for each tautomer of the imidazole ring of His, mentioned in Table 5, were computed at the DFT-level of theory because these nuclei are very sensitive probes with which to confirm the tautomeric forms of His accurately [6,11]. Chemical shifts are related to shieldings as differences with respect to a reference shielding. Hence, from Table 5, we can straightforwardly infer the following chemical-shifts inequalities:

$$\left(^{13}C^{\gamma}\right)_{N\varepsilon 2-H} \gg \left(^{13}C^{\gamma}\right)_{N\delta 1-H} \text{ and } \left(^{13}C^{\delta 2}\right)_{N\delta 1-H} \gg \left(^{13}C^{\delta 2}\right)_{N\varepsilon 2-H} \quad (1)$$

The chemical-shift inequalities between tautomers given by Eq. (1) are in full agreement with both the observed chemical shifts (see Table 2S of Sudmeier *et al.* [13]) and the DFT-computed shielding values in model peptides for the tautomeric forms of the imidazole ring of histidine (see Figure 1 of Vila *et al.* [6]).

The computation, at the DFT-level of theory, of the difference ($\Delta_{C\delta 2C\gamma}$) for the $^{1}J_{C\delta 2C\gamma}$ SSCC upon tautomeric changes, viz., with $\Delta_{C\delta 2C\gamma} = (^{1}J_{C\delta 2C\gamma}|^{\varepsilon 2} - {}^{1}J_{C\delta 2C\gamma}|^{\delta 1})$, where $^{1}J_{C\delta 2C\gamma}|^{\lambda}$ with $\lambda = \varepsilon 2$ or $\delta 1$ represents the computed $^{1}J_{C\delta 2C\gamma}$ SSCC value for the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomers, respectively, was carried out by using the His structures, the functional (OPW91) and the basis-set described in Table 4. The computed difference, $\Delta_{C\delta 2C\gamma} = +2.6$ Hz, is in good agreement, but opposed in sign, with the observed difference, $\Delta_{C\delta 2C\gamma} = -3.0$ Hz (see Table 2S of Sudmeier *et al.* [13]). Although no effort was made to optimize the functional and basis set for the computed one-bond C–C, $^{1}J_{CC}$, SSCC, an analyses with *all* the functionals shown in Table 1, with a locally-dense basis set distribution mentioned in Table 4, shows consistency with the result obtained with OPW91, namely $\Delta_{C\delta 2C\gamma} = +1.8$ Hz, +2.0 Hz, +2.6 Hz, +1.8 Hz and +1.9 Hz, for the B3LYP, B3P86, OPBE, B972 and PB86 functionals, respectively.

For a long time, it has been recognized that the Carbon–Nitrogen spin-spin coupling constants (J_{CN}) can be used for an accurate determination of the tautomeric forms of His [8,42]. For this reason, and despite the fact that no effort was made to optimize the functional or the basis set with which to compute the one- and two-bond C–N, J_{CN}, an analysis of the computed J_{CN} values was carried out by using the His structures, the functional (OPW91), and the basis set described in Table 4. Among *all* possible C–N SSCC of the imidazole ring of His, we focused the calculations on both the $^{1}J_{C\delta 2N\varepsilon 2}$ and the $^{2}J_{C\delta 2N\delta 1}$ because these SSCC were suggested as the probes of choice with which to determine the ratio of the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomers, accurately [8]. Consequently, the results obtained here for $^{1}J_{C\delta 2N\varepsilon 2}$ (13.2 Hz), in the $N^{\varepsilon 2}$-H tautomer (see Figure 1a), and $^{2}J_{C\delta 2N\delta 1}$ (4.7 Hz), in the $N^{\delta 1}$-H tautomer (see Figure 1b), show the same trend as that of the observed absolute values for these tautomers, namely ~13 Hz and ~5 Hz, respectively [8,42]. Even though this analysis does show internal consistency with the previous calculations of both the chemical shifts and $^{1}J_{C\delta 2H}$ SSCC, the validation of the theoretical results remains elusive without farther experimental evidence.

Overall, under the *only* condition that His must be non-protonated and, assuming the correctness of the theoretical predictions for $^1J_{C\delta 2H}$ SSCC of the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomers, it is possible to determine the fraction of the $N^{\delta 1}$-H tautomeric form of the imidazole ring of His, $f^{\delta 1}$, as a function of the observed $^1J_{C\delta 2H}$ SSCC value. So, assuming that $J^{\mathrm{obs}} = f^{\varepsilon 2} J^{\varepsilon 2}_{\mathrm{high}-pH} + f^{\delta 1} J^{\delta 1}_{\mathrm{high}-pH}$, and $f^{\varepsilon 2} + f^{\delta 1} = 1$ the following equation can be used for this purpose:

$$f^{\delta 1} = \frac{J^{\mathrm{obs}} - J^{\varepsilon 2}_{\mathrm{high}-pH}}{J^{\delta 1}_{\mathrm{high}-pH} - J^{\varepsilon 2}_{\mathrm{high}-pH}} = \frac{J^{\mathrm{obs}} - 165}{15} \qquad (2)$$

where J refers to $^1J_{C\delta 2H}$ SSCC, "*obs*" is the value observed, $J^{\varepsilon 2}_{\mathrm{high}-pH}$ and $J^{\delta 1}_{\mathrm{high}-pH}$ refer to $^1J_{C\delta 2H}$ SSCC obtained, at the high-pH limiting value, from the $N^{\varepsilon 2}$-H and $N^{\delta 1}$-H His tautomers, respectively; as shown in Table 4, "165" is the value for the $J^{\varepsilon 2}_{\mathrm{high}-pH}$ SSCC of the $N^{\varepsilon 2}$-H tautomer, and "15" is the difference, 180–165, between the computed $^1J_{C\delta 2H}$ SSCC for the tautomeric forms of His at the high-pH limit. Naturally, the fraction of the $N^{\varepsilon 2}$-H tautomeric forms is given by: $f^{\varepsilon 2} = 1 - f^{\delta 1}$.

It is worth noting that, even if the Sudmeier *et al.* [13] observations were confirmed Equation (2) will still be valid after the substitution $\delta 1 \rightarrow \varepsilon 2$.

Conclusions

The main results obtained for $^1J_{CH}$ SSCC in the $N^{\varepsilon 2}$-H and $N^{\delta 1}$-H tautomers of the imidazole ring of His indicate that: (*i*) $^1J_{C\varepsilon 1H} = 205 \pm 1.0$ Hz should be adopted as the canonical high-pH limiting value for this SSCC; (*ii*) $^1J_{C\varepsilon 1H}$ SSCC is not sensitive enough to be used as a probe with which to determine the tautomeric states of the imidazole ring of His; and (*iii*) $^1J_{C\delta 2H}$ SSCC shows a large difference (~15Hz) between tautomers and, hence, this SSCC emerges as a very sensitive probe with which to identify the tautomeric fractions of non-protonated His.

Overall, the theoretical confirmation showing that $^1J_{C\varepsilon 1H}$ SSCC is *not* sensitive upon tautomeric change means that the use of this SSCC will enable us to predict the protonation fraction of His, as a function of the pH, accurately, because the prediction will not depend on the fraction of the tautomeric forms. In addition, the fraction of the tautomeric forms of the imidazole ring of His, at either high-pH or for non-protonated buried His, can be predicted accurately by using Equation (2) for $^1J_{C\delta 2H}$ SSCC.

Acknowledgments

This research was supported by grants from the National Institutes of Health (GM-14312), the National Science Foundation (MCB10-19767) (HAS), and PIP-112-2011-0100030 from IMASL-CONICET-Argentina, Project 3-2212 from UNSL-Argentina, and PICT-2014-0556 from AMPCyT-Argentina (JAV).

References

1. Cheng F, Sun H, Zhang Y, Mukkamala D, Oldfield E. J Am Chem Soc. 2005; 127:12544–12554. [PubMed: 16144402]
2. Jensen MR, Has MAS, Hansen DF, Led JJ. Cell Mol Life Sci. 2007; 64:1085–1104. [PubMed: 17396226]
3. Hass MAS, Hansen DF, Christensen HEM, Led JJ, Kay LE. J Am Chem Soc. 2008; 130:8460–8470. [PubMed: 18540585]
4. Kumar M, Venkatnathan A. J Phys Chem B. 2015; 119:3213–3222. [PubMed: 25634681]
5. Li S, Hong M. J Am Chem Soc. 2011; 133:1534–1544. [PubMed: 21207964]
6. Vila JA, Arnautova YA, Vorobjev YN, Scheraga HA. Proc Natl Acad Sci USA. 2011; 108:5602–5607. [PubMed: 21422292]
7. Demchuk E, Wade RC. J Phys Chem. 1996; 100:17373–17387.
8. Shimba N, Serber Z, Ledwidge R, Miller SM, Craik CS, Dotsch V. Biochem. 2003; 42:9227–9234. [PubMed: 12885258]
9. Bermúdez C, Mata S, Cabezas C, Alonso JL. Angew Chem Int Ed. 2014; 53:11015–11018.
10. Blomberg W, Rüterjans H. J Am Chem Soc. 1977; 99:8149–8159. [PubMed: 925263]
11. Pelton JG, Torchia DA, Meadow ND, Roseman S. Prot Sci. 1993; 2:543–558.
12. Kolehmainen E, Ośmiałowski B. Int Rev Phys Chem. 2012; 31:567–629.
13. Sudmeier JL, Bradshaw EM, Haddad EC, Day RM, Thalhauser CJ, Bullock PA, Bachovchin WW. J Am Chem Soc. 2003; 125:8430–8431. [PubMed: 12848537]
14. Hansen AL, Kay LE. Proc Natl Acad Sci USA. 2014; 111:E1705–E1712. [PubMed: 24733918]
15. Platzer G, Okon M, McIntosh LP. J Biomol NMR. 2014; 60:109–129. [PubMed: 25239571]
16. Vila JA. J Phys Chem B. 2012; 116:6665–6669. [PubMed: 22376024]
17. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Wenger RK, Yao H, Markley JL. Nucleic Acids Res. 2008; 36:D402–D408. [PubMed: 17984079]
18. Singer AU, Forman-Kay JD. Prot Sci. 1997; 6:1910–1919.
19. Day RM, Thalhauser CJ, Sudmeier JL, Vincent MP, Torchilin EV, Sanford DG, Bachovchin CW, Bachovchin WW. Protein Sci. 2003; 12:794–810. [PubMed: 12649438]
20. 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. Gaussian 09, revision A.1. Gaussian, Inc; Wallingford, CT: 2009.
21. Deng W, Cheeseman JR, Frisch MJ. J Chem Theory Comput. 2006; 2:1028–1037. [PubMed: 26633062]
22. Ramsey NF. Phys Rev. 1953; 91:303–307.
23. Ruden TA, Lutnæs OB, Helgaker T, Ruud KJ. Chem Phys. 2003; 118:9572–9581.
24. Momany FA, McGuire AF, Burgess AW, Scheraga HA. J Phys Chem. 1975; 79:2361–2381.
25. Némethy G, Gibson KD, Palmer KA, Yoon CN, Paterlini G, Zagari A, Rumsey S, Scheraga HA. J Phys Chem. 1992; 96:6472–6484.
26. Oda K, Koyama HA. Acta Cryst B. 1972; 28:639–642.
27. Arnautova YA, Vila JA, Martin OA, Scheraga HA. Acta Cryst. 2009; D65:697–703.
28. Vijay-Kumar S, Bugg CE, Cook WJ. J Mol Biol. 1987; 194:531–544. [PubMed: 3041007]

29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. Nucleic Acids Research. 2000; 28:235–242. [PubMed: 10592235]
30. Ruud K, Frediani L, Cammi R, Mennucci B. Int J Mol Sci. 2003; 4:119–134.
31. Tomasi J, Persico M. Chem Rev. 1994; 94:2027–2094.
32. Tomasi J. Comput Mol Sci. 2011; 1:855–867.
33. Li L, Li C, Zhang Z, Alexov E. J Chem Theory Comput. 2013; 9:2126–2136. [PubMed: 23585741]
34. Andersson KM, Hovmöller S. Acta Cryst. 2000; D56:789–790.
35. Maximoff SN, Peralta JE, Barone V, Scuseria GE. J Chem Theory Comput. 2005; 1:541–545. [PubMed: 26641673]
36. Peralta JE, Scuseria GE, Cheeseman JR, Frisch MJ. Chem Phys Lett. 2003; 375:452–458.
37. Provasi PF, Aucar GA, Sauer SPA. J Chem Phys. 2001; 115:1024–1333.
38. Contreras RH, Peralta JE, Giribet CG, de Azua MC, Facelli JC. Annu Rep NMR Spectrosc. 2000; 41:55–184.
39. Krivdin LB, Contreras RH. Ann Repts NMR Spectro sc. 2007; 61:133–245.
40. San Fabian J, Garcia de la Vega JM, Suardiaz R, Fernández-Oliva M, Pérez C, Crespo-Otero R, Contreras RH. Magn Reson Chem. 2013; 51:775–787. [PubMed: 24123317]
41. Chesnut DB, Moore KD. J Comp Chem. 1989; 10:648–659.
42. Sudmeier JL, Ash EL, Günter UL, Luo X, Bullock PA, Bachovchin WW. J Mag Res B. 1996; 113:236–247.

Highlights

- The $^{1}J_{CH}$ SSCC of the imidazole ring of His was computed at the DFT-level of theory.
- Solvent effects (polarizable continuum model approach) were considered.
- The Zero Point Vibrational Contribution was included.
- $^{1}J_{C\varepsilon 1H}$ SSCC shows a slight difference (~1 Hz) between tautomers.
- $^{1}J_{C\delta 2H}$ SSCC shows a large difference (~15Hz) between tautomers.

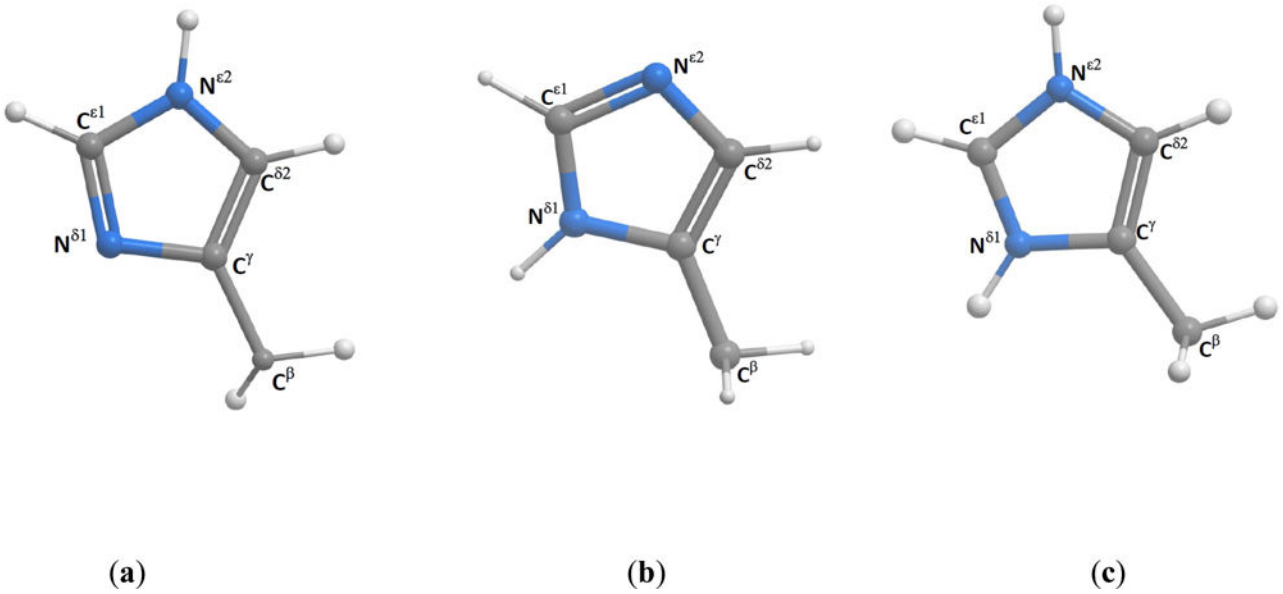


Figure 1.

Ball and stick representation of the forms of the imidazole ring of His, namely the: (**a**) $N^{\varepsilon 2}$-H, or **τ** tautomer [5], (**b**) $N^{\delta 1}$-H, or **π** tautomer [5], and (**c**) H^{+} form, respectively.

Table 1

Test of functionals for the DFT computations of $^{1}J_{C\varepsilon 1H}$ SSCC of the $N^{\varepsilon 2}$-H tautomer[a]

Functional	$\Sigma_{\{FC,SD,PSO,DSO\}}$ (Hz)	$^{1}J_{C\varepsilon 1H}$ (Hz)
OPW91	**199.02**	**~204**
B3LYP	231.65	~237
B3P86	215.54	~220
OPBE	**198.10**	**~203**
B972	212.57	~218
BP86	211.35	~216

[a]All gas-phase DFT calculations of $^{1}J_{C\varepsilon 1H}$ SSCC were carried out on Ac-His-NMe by using the Gaussian 09 suite of programs [20]; the chosen ϕ, ψ, ω, χ1 and χ2 torsional angles for His correspond to a local-minimum of the ECEPP force-field [25] in the α-helical region of the Ramachandran map: −73.563°, −35.197°, −179.856°, 66.389° and −62.607°, respectively; for each functional we used an “aug-cc-pVTZ-J” basis set [37] on *all* nuclei of the imidazole ring of His, and a “6-31G” basis set on the remaining nuclei of Ac-**His**-NMe. The total (**Σ**) is a sum over the four Ramsey contributions [22], as given by the output of the Gaussian 09 suite of programs [20], namely, the Fermi Contact (**FC**), the Spin Dipolar (**SD**), the Paramagnetic Spin-Orbit (**PSO**), and the Diamagnetic Spin-Orbit (**DSO**) contribution, respectively; the last column lists the predicted value for the one-bond $^{1}J_{C\varepsilon 1H}$ SSCC after adding 5 Hz to the **Σ** term (second column), due to the Zero Point Vibrational Contribution [23].

Table 2

High-pH limiting value for $^{1}J_{C\varepsilon 1H}$ SSCC of the $N^{\varepsilon 2}$-H tautomer[a]

$\Sigma_{\{FC,SD,PSO,DSO\}}$ (Hz)	$^{1}J_{C\varepsilon 1H}$ (Hz)
199.20	~204[b]
199.29	~204[c]
199.10	~204[d]
199.02	~204[e]
198.91	~204[f]
199.12	~204[g]
198.62	~204[h]
201.26	***~206***[i]
201.78	***~207***[j]

[a] All gas-phase DFT calculations of $^{1}J_{C\varepsilon 1H}$ SSCC, unless otherwise noted, for which see items *i*–*j* below, were carried out for Ac-His-NMe by using the Gaussian 09 suite of programs [20]; the chosen ϕ, ψ, ω, χ1 and χ2 torsional angles for His correspond to a local-minimum of the ECEPP force-field [25] in the α-helical region of the Ramachandran map: −73.563°, −35.197°, −179.856°, 66.389° and −62.607°, respectively, unless otherwise noted, for which see items *f*–*h* below; the total (**Σ**) is a sum over the four Ramsey contributions (see footnote *a* of Table 1) [22]; column 2 lists the predicted values for the one-bond $^{1}J_{C\varepsilon 1H}$ SSCC after adding 5 Hz to the **Σ** term (in column 1), due to the Zero Point Vibrational Contribution [23].

[b] Result obtained by using a *uniform* basis set "aug-cc-pVTZ-J" [37] for *all* the nuclei of Ac-His-NMe.

[c] Result obtained by using an "aug-cc-pVTZ-J" basis set for the $^{13}C^{\varepsilon 1}$, $^{1}H^{\varepsilon 1}$, $N^{\delta 1}$, $N^{\varepsilon 2}$ and $H^{\varepsilon 2}$ nuclei of the imidazole ring of His and a "3-21G" basis set for *all* the remaining nuclei of Ac-His-NMe, i.e., by using a "locally-dense" basis set approximation [41].

[d] Same as (*c*) with a "6-31G" rather than a "3-21G" basis set.

[e] Result obtained by using an "aug-cc-pVTZ-J" basis set for *all* nuclei of the imidazole ring of His, and a "6-31G" basis set for the remaining nuclei of Ac-His-NMe, i.e., by using a "locally-dense" basis set approximation.

[f] Same as (*e*) with the ϕ, ψ, ω, χ1 and χ2 torsional angles of His from a local-minimum, of the ECEPP force-field [25], in the extended region of the Ramachandran map: −142.183°, 153.652°, −179.870°, −58.886° and 113.207°, respectively.

[g] Same as (*e*) with the ϕ, ψ, ω, χ1 and χ2 torsional angles of His from a local-minimum, of the ECEPP force-field,[25] in the extended region of the Ramachandran map: −156.922°, 159.176°, 179.634°, 60.066° and 96.365°, respectively.

[h] Same as (*e*) with the ϕ, ψ, ω, χ1 and χ2 torsional angles of His from a local-minimum, of the ECEPP force-field [25], in the extended region of the Ramachandran map: −154.460°, 157.818°, −178.614°, −152.343° and −67.537°, respectively.

[i] Same as (*c*) with solvent effects computed by using the PCM [31,32] approach with $D_i = 6.5$.

[j] Same as (*i*) with $D_i = 25.5$

Table 3

High-pH limiting value for $^{1}J_{C\varepsilon 1H}$ SSCC of the $N^{\delta 1}$-H tautomer[a]

$\Sigma_{\{FC,SD,PSO,DSO\}}$ (Hz)	$^{1}J_{C\varepsilon 1H}$ (Hz)
197.98	~203[b]
198.02	~203[c]
197.84	~203[d]
19787	~203[e]
199.32	***~204***[f]
199.66	***~205***[g]

[a]All gas-phase DFT calculations of $^{1}J_{C\varepsilon 1H}$ SSCC, unless otherwise noted, for which see items *f*–*g* below, were carried out for Ac-His-NMe by using the Gaussian 09 suite of programs [20]; the total (**Σ**) is a sum over the four Ramsey contributions (see footnote *a* of Table 1); the chosen ϕ, ψ, ω, χ1 and χ2 torsional angles for His correspond to a local-minimum of the ECEPP force-field [25] in the α-helical region of the Ramachandran map: −74.737°, −39.192°, 179.928°, −67.874°, and 28.534°, respectively; column 2 lists the predicted values for the one-bond $^{1}J_{C\varepsilon 1H}$ SSCC after adding 5 Hz to the **Σ** term (in column 1), due to the Zero Point Vibrational Contribution [23].

[b]Result obtained by using a *uniform* basis set "aug-cc-pVTZ-J" [37] for *all* the nuclei of Ac-His-NMe.

[c]Result obtained by using an "aug-cc-pVTZ-J" basis set for the $^{13}C^{\varepsilon 1}$, $^{1}H^{\varepsilon 1}$, $N^{\delta 1}$, $N^{\varepsilon 2}$ and $H^{\varepsilon 2}$ nuclei of the imidazole ring of His, and a "3-21G" basis set for *all* the remaining nuclei of Ac-**His**-NMe, i.e., by using a "locally-dense" basis set approximation [41].

[d]Same as (*c*) with "6-31G" rather than a "3-21G" basis set.

[e]Result obtained by using an "aug-cc-pVTZ-J" basis set for *all* nuclei of the imidazole ring of His, and a "6-31G" basis set on the remaining nuclei of Ac-His-NMe, i.e., by using a "locally-dense" [41] basis set approximation.

[f]Same as (*e*) with solvent effects computed by using the PCM [31,32] approach with $D_i = 6.5$.

[g]Same as (*f*) with $D_i = 25.5$.

Table 4

High-pH limiting value for $^{1}J_{C\delta 2H}$ SSCC[a]

Tautomer	$\Sigma_{\{FC,SD,PSO,DSO\}}$ (Hz)	$^{1}J_{C\delta 2H}$ (Hz)
$N^{\varepsilon 2}$-H	158.54	~164[b]
	159.87	***~165***[c]
	160.08	***~165***[d]
$N^{\delta 1}$-H	176.14	~181[b]
	175.38	***~180***[c]
	175.07	***~180***[d]

[a] All gas-phase DFT calculations of $^{1}J_{C\delta 2H}$ SSCC, unless otherwise noted, for which see items *c*–*d* below, were carried out on Ac-His-NMe by using the Gaussian 09 suite of programs [20]; the chosen ϕ, ψ, ω, χ1 and χ2 torsional angles for the $N^{\varepsilon 2}$-H and $N^{\delta 1}$-H tautomers of His correspond to a local-minimum in the α-helical region of the Ramachandran map, and are given in footnote *a* of Tables 2 and 3, respectively; the total (**Σ**) is a sum over the four Ramsey [22] contributions (see item *a* of Table 1); column 3 list the predicted values for the one-bond $^{1}J_{C\delta 2H}$ SSCC after adding 5 Hz to the **Σ** term (in column 2), due to the Zero Point Vibrational Contribution [23].

[b] Result obtained by using an "aug-cc-pVTZ-J" basis set [37] for *all* nuclei of the imidazole ring of His and "6-31G" on the remaining nuclei of Ac-His-NMe.

[c] Same as (*b*) with solvent effects computed by using the PCM [31,32] approach with $D_i = 6.5$

[d] Same as (*c*) with $D_i = 25.5$

Table 5

Chemical shieldings for the $^{13}C^{\delta 2}$ and $^{13}C^{\gamma}$ nuclei[a]

Nucleus	$N^{\varepsilon 2}$-H tautomer Shielding (ppm)	$N^{\delta 1}$-H tautomer Shielding (ppm)
$^{13}C^{\gamma}$	42.1	61.9
$^{13}C^{\delta 2}$	78.1	60.8

[a]All gas-phase DFT-calculations were computed, by using the Gaussian 09 suite of programs [20], for Ace-His-NMe with the OB98 functional [6] and a “6-311+G(2d,p)” basis set [6] for *all* the nuclei of His and a “6-31G” basis set on the remaining nuclei of the molecule, i.e., by using a “locally-dense” basis set approximation.[41]