DOI: 10.1002/ange.201405347

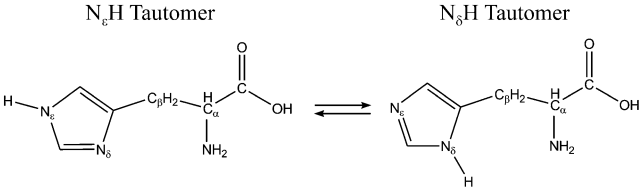
Microwave Spectroscopy

Tautomerism in Neutral Histidine**

*Celina Bermúdez, Santiago Mata, Carlos Cabezas, and José L. Alonso**

Abstract: *Histidine is an important natural amino acid, involved in many relevant biological processes, which, because of its physical properties, proved difficult to characterize experimentally in its neutral form. In this work, neutral histidine has been generated in the gas phase by laser ablation of solid samples and its $N_\varepsilon H$ tautomeric form unraveled through its rotational spectrum. The quadrupole hyperfine structure, arising from the existing three ^{14}N nuclei, constituted a site-specifically probe for revealing the tautomeric form as well as the side chain configuration of this proteogenic amino acid.*

Histidine (His, H_2N-CH-(CH_2-$C_3N_2H_3$)-COOH) is one of the twenty proteogenic amino acids present in many relevant proteins.[1] The imidazol group of its side chain, makes His to be one of the two natural amino acids that can exhibit tautomeric equilibrium (Scheme 1). Either N_δ or N_ε of the imidazol ring might be protonated, and, thus, both can participate in inter- and intramolecular hydrogen bonds, acting as donor or as acceptor. As such, His residue is an excellent binding site either for other molecules, or to induce stabilization by intramolecular interactions within the protein structure.[2] Several spectroscopic techniques have been devoted to analyze His tautomeric balance in condensed phases, where His is stabilized as a zwitterion (^+H_3N-CH(CH_2-$C_3N_2H_3$)-COO^-),[3] and, thus, it does not represent the neutral canonical form present in peptide side chains. The tautomeric fraction of the imidazole ring of His for proteins in solution varies significantly among different positions of His in the same protein, reflecting the importance of the environment in determining the tautomeric behavior.[4] For these reasons, the determination of tautomeric behavior of His in gas phase, where it is stabilized in its neutral canonical form, is of utmost importance to gain knowledge on their intrinsic tautomeric/conformational properties.



Scheme 1. Tautomeric equilibrium in histidine.

His, a solid with high melting point (m.p. about 290 °C) and low vapor pressure, is well-known for its thermal instability, preventing easy measurements of its gas-phase spectra in static thermally heated gas cells. Hence, studies on the tautomeric/conformational preferences of neutral form have been, until now, restricted solely to the theoretical field.[5] Nowadays, advances in laser ablation of solid biomolecules have allowed to overcome vaporization problems. Thus, high-resolution spectroscopy techniques have enabled scientific community to observe individual conformers of biomolecules with unprecedented clarity.[6] Particularly, the combination of Fourier transform microwave spectroscopy with laser ablation techniques conducted in a supersonic expansion (LA-MB-FTMW) has provided a new approach to the structural studies of amino acids,[7] nitrogen bases,[8] and other relevant building blocks.[9]

Nevertheless, His still remains as a challenging problem for a rotational spectroscopic study. His possesses three ^{14}N nuclei with nuclear spin $I = 1$, one located in the amino group ($^{14}N_a$) and the other two in the imidazole ring (N_ε and N_δ). These three interact with the electric field gradient created by the rest of the molecule at the nucleus, splitting into a very complex hyperfine pattern each rotational transition and, thus, increasing the difficulty of spectral interpretation. In this project, we have faced these challenges successfully and here we report the first rotational study of neutral His using a newly constructed Fourier transform microwave spectrometer (see Supporting Information).

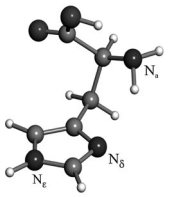
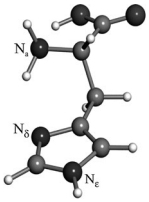
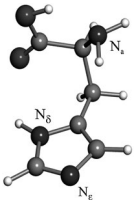
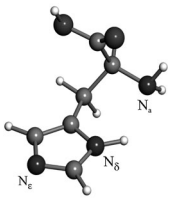
With this aim, solid His has been transferred into the gas phase by laser ablation. Vaporized molecules are then seeded in a stream of Ne carrier gas, and supersonically expanded, becoming ideally frozen in their most stable forms. These can be interrogated by a microwave radiation pulse in the prepared solvent-free environment of the supersonic expansion. Spectroscopic searches in wide frequency regions revealed the rotational spectrum of only one rotamer with the characteristic pattern of a near-prolate asymmetric top with sets of a-type R-branch transitions. As anticipated, all observed transitions were split into many close hyperfine components, arising from the coupling of three non-equivalent $^{14}N_\delta$, $^{14}N_\varepsilon$, and $^{14}N_a$ nuclei. At first glance, the assignment of hyperfine structure seemed impossible. Hence, the rotational frequencies of nine a-type transitions were roughly

[*] C. Bermúdez, S. Mata, Dr. C. Cabezas, Prof. J. L. Alonso
Grupo de Espectroscopia Molecular (GEM)
Edificio Quifima Laboratorios de Espectroscopia y Bioespectroscopia, Parque Científico UVa
Universidad de Valladolid, 47011 Valladolid (Spain)
E-mail: jlalonso@qf.uva.es
Homepage: http://www.gem.uva.es

[**] This research was supported by the Ministerio de Ciencia e Innovación (grant numbers CTQ 2010-19008 and Consolider Ingenio 2010 CSD 2009-00038), Junta de Castilla y León (grant number VA070A08). C.B. thanks Ministerio de Ciencia e Innovación for a FPI grant (grant number BES-2011-047695). C.C. thanks the Junta de Castilla y León for the postdoctoral contract (grant number CIP13/01).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201405347.

Table 1: Ab initio predicted rotational and quadrupole coupling constants for the four lowest-energy conformers[a] of histidine.

				
	εII_a	εII_b	δII_a	δI_a
A/B/C[b]	1818/862/772	2993/568/501	1917/826/729	1805/847/785
$\chi_{aa}/\chi_{bb}/\chi_{cc}$ (N_δ)	1.66/−3.51/1.85	0.67/−2.18/1.51	0.14/1.22/−1.36	0.20/−0.17/−0.03
$\chi_{aa}/\chi_{bb}/\chi_{cc}$ (N_ε)	−0.19/1.09/−0.90	1.15/1.02/−2.16	−2.27/1.63/0.64	−1.77/0.05/1.72
$\chi_{aa}/\chi_{bb}/\chi_{cc}$ (N_a)	0.58/1.93/−2.52	−1.53/2.08/−0.55	−0.67/1.41/−0.74	−4.04/2.69/1.35
$\lvert\mu_a\rvert/\lvert\mu_b\rvert/\lvert\mu_c\rvert$	3.5/1.5/0.2	7.8/0.4/1.9	1.8/3.1/1.3	4.4/2.3/1.6
ΔE	0	742	422	838
ΔG	0	678	480	731

[a] Conformers are labelled following the nomenclature used in previous studies of amino acids. The first label distinguishes between tautomers; $\boldsymbol{\varepsilon}$ (for tautomer $N_\varepsilon H$) and δ (for tautomer $N_\delta H$). The second index indicates the type of hydrogen bond between the amino and the carboxylic group (Ref. [7b]). Finally, lower labels a and b designate the increasing energy order within each type of hydrogen bonding. [b] A, B, C are the rotational constants (in MHz); χ_{aa}, χ_{bb}, χ_{cc} are the ^{14}N nuclear quadrupole coupling constants (in MHz); $|\mu_a|$, $|\mu_b|$, $|\mu_c|$ are the absolute values of the electric dipole moment components (in D); ΔE and ΔG are the MP2/6-311++G(d,p) electronic energies and Gibbs free energies (298 K), respectively (in cm^{-1}), with respect to the global minimum.

measured as the intensity-weighted mean of the line clusters and fitted[10] to a rigid rotor Hamiltonian leading to a preliminary set of rotational constants A about 1848, B about 832, and C about 746 (all in MHz). To ascertain which His structure is responsible for the observed spectrum, theoretical values of the rotational constants of the most stable forms of His are required. Thus, the lowest energy conformations reported in previous theoretical studies[5] were re-optimized using ab initio calculations[11] at the MP2/6-311++G(d,p) level of theory. The spectroscopic parameters for the conformers lying in an energy window of 1000 cm^{-1} are collected in Table 1. The values of the experimental rotational constants were found to be consistent with those predicted for the conformers labeled as εII_a, δII_a, and δI_a. Unfortunately, the difference in the rotational constants values for these three species is not large enough to allow discrimination and a conclusive identification cannot be reached on this basis.

A different and independent way of identifying structures is based on the presence of ^{14}N nuclei in the molecule. While rotational constants are strongly related to mass distribution, nuclear quadrupole coupling interactions depend critically on the electronic environment, position and orientation of the ^{14}N nuclei. The quadrupole coupling constants[12] have been used as fingerprints in conformational analysis of amino acids,[7] as well as for tautomeric identification of nucleobases.[8] For His, the predicted values of the quadrupole coupling constants (χ_{aa}, χ_{bb} and χ_{cc}) for the ^{14}N nuclei (see Table 1) are distinct because of the dissimilar nature of bonding in the vicinity of the nucleus. Thus, they could provide an independent approach to discriminate tautomeric species. In order to unveil, in a conclusive fashion, the observed tautomeric species, it becomes necessary to resolve and interpret the quadrupole hyperfine structure of His, an asymmetric top with three ^{14}N nuclei of different electronic environment.

The next stage of the investigation covered the analysis of the nuclear quadrupole hyperfine structure. b-Type R-branch transition, whose hyperfine components are predicted to be most spread in frequency, have to be first analyzed. The construction of the new LA-MB-FTMW spectrometer, covering lower frequency regions, made possible to record the b-type R-branch transition $1_{1,1}\leftarrow 0_{0,0}$ at about 2.5 GHz (Figure 1). Interpretation of the quadrupole coupling pattern led to the assignment of ten hyperfine components of this transition, essential as starting point of the analysis. New predictions allowed the assignment of a total of 75 hyperfine components (Table S1 of Supporting Information) belonging to nine a- and three b-type R-branch transitions. They were analyzed[10] using a Watson's A-reduced semirigid rotor

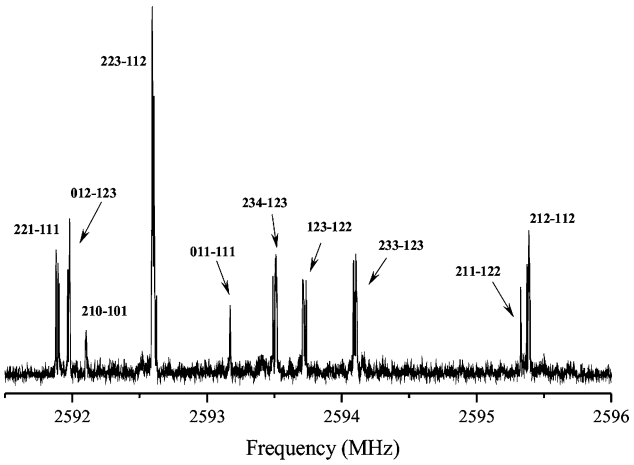


Figure 1. 1_{11}–0_{00} rotational transition in the 2 GHz frequency region for the observed rotamer of His. The ^{14}N quadrupole components are labeled with the quantum numbers F_1', F_2', F'-F_1'', F_2'', F'' (nuclear quadrupole scheme: $I_1+J=F_1$, $I_2+F_1=F_2$, $I_3+F_2=F$ where I stands for the nuclear spin for each ^{14}N).

Table 2: Experimental and predicted rotational and quadrupole coupling constants for the observed rotamer of histidine.

	Experimental	Theoretical MP2/cc-pVTZ
A[a]	1847.53472 (52)[b]	1839[c]
B	831.71551 (16)	859
C	745.94445 (18)	770
Δ_J	0.2651 (53)	–
χ_{aa}/N_δ	1.6113 (17)	1.62
χ_{bb}/N_δ	−3.4973 (16)	−3.49
χ_{cc}/N_δ	1.8860 (16)	1.87
χ_{aa}/N_ε	−0.17933 (26)	−0.18
χ_{bb}/N_ε	1.12207 (87)	0.97
χ_{cc}/N_ε	−0.94273 (87)	−0.79
χ_{aa}/N_a	0.0052 (22)	0.04
χ_{bb}/N_a	2.0982 (43)	2.10
χ_{cc}/N_a	−2.1034 (43)	−2.14
N	75	–
σ	1.9	–

[a] A, B, C are the rotational constants (in MHz); Δ_J is a quartic centrifugal distortion constant (in kHz); χ_{aa}, χ_{bb}, χ_{cc} are the ^{14}N nuclear quadrupole coupling constants (in MHz); *N* is the number of fitted transitions; σ is the rms deviation of the fit (in kHz). [b] Standard error in parentheses in units of the last digit. [c] Values calculated at MP2/cc-pVTZ level of theory.

Hamiltonian in the *I*′-representation[13] supplemented with a term to account for the nuclear quadrupole coupling contribution.[12] Table 2 illustrates how such analysis rendered accurate rotational and nuclear quadrupole constants. Comparison of experimental quadrupole coupling constants with those predicted for conformers εII_a, δII_a, and δI_a allowed the unequivocal identification of the observed rotamer as conformer εII_a. This assignment is further confirmed attending to the predicted values of the dipole moments and the type of spectra observed. A close look into the values of nuclear quadrupole coupling constants for $^{14}N_a$ nucleus indicates small discrepancies between experimental and theoretical values. Although the difference is not large enough to raise doubts about the identity of the conformer detected, it suggests the existence of some discrepancy between the calculated geometries and the actual ones. Such discrepancies have lead to the correction of the orientation of the amino group in the εII_a conformer, by performing ab initio calculations for different values of the dihedral angle ($>HN_aC_\alpha C_\beta$). An improved matching was found when this angle was rotated from the initial ab initio value of −16° to −23° (see Figure S1 of Supporting Information). This fact prompted us to make other MP2 optimizations using a larger basis set as cc-pVTZ in an attempt to find a more suitable calculation basis set to reproduce our experimental nuclear quadrupole coupling constants. In this manner, we have re-optimized the ab initio structure at the MP2/cc-pVTZ level of theory obtaining a value for this dihedral angle of −21° and, consequently, a closer concordance between the experimental and predicted quadrupole coupling constants (see Table 2). This results in a more favorable arrangement to the establishment of N_a–H···N_δ interactions between the amino group and imidazol ring analogs to the H bond that stabilized one of the species found for its homologue histamine.[14] The 3D structure of εII_a conformer, shown in the Figure 2, has been taken from that predicted by ab initio calculations (cartesian coordinates are given in Table S2 of the Supporting Information), based in the good agreement between experimental and theoretical values of rotational constants (relative errors less than 3%).

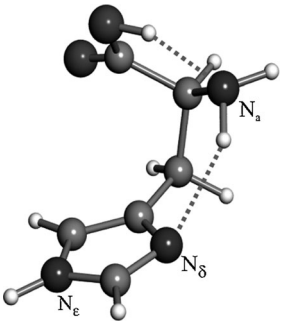


Figure 2. Observed conformer for the histidine molecule, showing the intramolecular hydrogen bonds that stabilize the structure.

Summarizing, the present study provides the first experimental information on the conformational and tautomeric properties of neutral His. The capability of LA-MB-FTMW spectroscopy to undertake very accurate spectroscopic constants and their direct comparison with ab initio computations provides an unmatched means for the unequivocal identification of the observed species. The three $^{14}N_\delta$, $^{14}N_\varepsilon$, and $^{14}N_a$ nuclei of His present at defined sites introduce hyperfine rotational probes that further expand the utility of this spectroscopic technique. Our results indicate that neutral His exits in the gas phase in the $N_\varepsilon H$ tautomeric form in a single εII_a conformation which is the one predicted as the global minimum. This is stabilized by an O–H···N_a hydrogen bond in an α-COOH trans configuration also found in the rest of essential aromatic amino acids.[15] One of the hydrogen atoms of the amino group is pointing towards the imidazol ring indicating the existence of an N_a–H···N_δ interaction, since none H is attached to the N_δ atom, this can act as a proton acceptor in the intramolecular H-bond. Both hydrogen bonds form an intramolecular network O–H···N_a–H···N_δ, which could be the stabilization motif of the observed species. The present state-of-the-art of microwave spectroscopy, as illustrated in the present study, is paving the way towards the study of larger, more complex, biological systems which have been previously considered as being out of reach of high-resolution spectroscopic studies.

Received: May 16, 2014
Published online: August 21, 2014

Keywords: conformational analysis · histidine · laser ablation · microwave spectroscopy · tautomerism

[1] A. L. Lehninger, *Principles of Biochemistry*, 4edth edW. H. Freeman, **2004**.

[2] a) J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, *Nature* **1960**, *185*, 422 – 427; b) M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T. North, *Nature* **1960**, *185*, 416 – 422; c) A. Volbeda, M. H. Charon, C. Piras, E. C. Hatchikian, M. Frey, J. C. Fontecillacamps, *Nature* **1995**, *373*, 580 – 587; d) S. Dementin, V. Belle, P. Bertrand, B. Guigliarelli, G. Adryanczyk-Perrier, A. L. De Lacey, V. M. Fernandez, M. Rousset, C. Leger, *J. Am. Chem. Soc.* **2006**, *128*, 5209 – 5218.

[3] a) F. Blomberg, W. Maurer, H. Ruterjans, *J. Am. Chem. Soc.* **1977**, *99*, 8149 – 8159; b) M. S. Lehmann, T. F. Koetzle, W. C. Hamilton, *Int. J. Pept. Protein Res.* **1972**, *4*, 229 – 239; c) J. J. Madden, E. L. McGandy, N. C. Seeman, *Acta Crystallogr. Sect. B* **1972**, *28*, 2377 – 2382; d) P. Edington, M. M. Harding, *Acta Crystallogr. Sect. B* **1974**, *30*, 204 – 206; e) F. Pflüger, B. Hernandez, M. Ghomi, *J. Phys. Chem. B* **2010**, *114*, 9072 – 9083; f) S. H. Li, M. Hong, *J. Am. Chem. Soc.* **2011**, *133*, 1534 – 1544.

[4] J. A. Vila, Y. A. Arnautova, Y. Vorobjev, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5602 – 5607.

[5] a) B. Kovačević, M. Rožman, L. Klasinc, D. Srzić, Z. B. Maksić, M. Yáñez, *J. Phys. Chem. A* **2005**, *109*, 8329 – 8335; b) Z. J. Huang, W. B. Yu, Z. J. Lin, *J. Mol. Struct. (Theochem)* **2006**, *801*, 7 – 20; c) Z. J. Huang, Z. J. Lin, C. Song, *J. Phys. Chem. A* **2007**, *111*, 4340 – 4352; d) Z. A. Tehrani, E. Tavasoli, A. Fattahi, *J. Mol. Struct. (Theochem)* **2010**, *960*, 73 – 85.

[6] J. P. Shermann, *Spectroscopy and Modelling of Biomolecular Builiding Blocks*, Elsevier, Amsterdam, **2008**.

[7] a) C. Cabezas, M. Varela, I. Peña, S. Mata, J. C. López, J. L. Alonso, *Chem. Commun.* **2012**, *48*, 5934 – 5936; b) I. Peña, M. E. Sanz, J. C. López, J. L. Alonso, *J. Am. Chem. Soc.* **2012**, *134*, 2305 – 2312.

[8] J. L. Alonso, V. Vaquero, I. Peña, J. C. López, S. Mata, W. Caminati, *Angew. Chem.* **2013**, *125*, 5934 – 5936; *Angew. Chem. Int. Ed.* **2013**, *52*, 2331 – 2334.

[9] a) C. Cabezas, M. Varela, J. L. Alonso, *ChemPhysChem* **2013**, *14*, 2539 – 2543; b) J. L. Alonso, M. A. Lozoya, I. Peña, J. C. López, C. Cabezas, S. Mata, S. Blanco, *Chem. Sci.* **2014**, *5*, 515 – 522.

[10] H. M. Pickett, *J. Mol. Spectrosc.* **1991**, *148*, 371 – 377.

[11] Gaussian 09, Revision A.1, M. J. Frisch et al., Gaussian Inc., Wallingford CT, **2009**.

[12] W. Gordy, R. L. Cook, *Microwave Molecular Spectra*, *Vol. IX and XIV*, Wiley-Interscience, New York, **1984**.

[13] J. K. G. Watson, *Vibrational Spectra and Structure*, *Vol. 6*, Elsevier, Amsterdam, **1977**.

[14] P. D. Godfrey, R. D. Brown, *J. Am. Chem. Soc.* **1998**, *120*, 10724 – 10732.

[15] a) C. Pérez, S. Mata, S. Blanco, J. C. López, J. L. Alonso, *J. Phys. Chem. A* **2011**, *115*, 9653; b) M. E. Sanz, C. Cabezas, S. Mata, J. L. Alonso, *J. Chem. Phys.* **2014**, *140*, 204308.

