

Microwave Spectroscopy of Biomolecular Building Blocks

José L. Alonso and Juan C. López

Abstract Microwave spectroscopy, considered as the most definitive gas phase structural probe, is able to distinguish between different conformational structures of a molecule, because they have unique spectroscopic constants and give rise to distinct individual rotational spectra.

Previously, application of this technique was limited to molecular specimens possessing appreciable vapor pressures, thus discarding the possibility of studying many other molecules of biological importance, in particular those with high melting points, which had a tendency to undergo thermal reactions, and ultimately degradation, upon heating.

Nowadays, the combination of laser ablation with Fourier transform microwave spectroscopy techniques, in supersonic jets, has enabled the gas-phase study of such systems. In this chapter, these techniques, including broadband spectroscopy, as well as results of their application into the study of the conformational panorama and structure of biomolecular building blocks, such as amino acids, nucleic bases, and monosaccharides, are briefly discussed, and with them, the tools for conformational assignment – rotational constants, nuclear quadrupole coupling interaction, and dipole moment.

Keywords Amino acids · Building blocks · Carbohydrates · Conformational analysis · Laser ablation · Microwave spectroscopy · Nitrogen bases

Contents

1	Introduction	336
2	Experimental Techniques	339

J.L. Alonso (✉) and J.C. López

Grupo de Espectroscopia Molecular (GEM), Edificio Quifima, Laboratorios de Espectroscopia y Bioespectroscopia, Unidad Asociada CSIC, Universidad de Valladolid, 47005 Valladolid, Spain
e-mail: jalonso@qf.uva.es

2.1	Laser-Ablation Molecular-Beam Fourier Transform Microwave (LA-MB-FTMW) Spectroscopy	339
2.2	Chirped-Pulse Fourier Transform Microwave Spectroscopy (CP-FTMW) Coupled to Laser Ablation (LA)	341
3	Tools in Conformational Analysis	344
3.1	Model Calculations	344
3.2	Analysis of Spectra and Conformer Identification	348
4	Amino Acids	351
4.1	Introduction	351
4.2	Proteinogenic Amino Acids with Non-polar Side Chains	353
4.3	Imino Acids Proline and Hydroxyproline	357
4.4	Proteinogenic Amino Acids with Polar Side Chain	360
4.5	Amino Acids with Aromatic Side Chain	366
4.6	Non-proteinogenic Amino Acids	368
5	Nitrogen Bases	371
5.1	Uracil, Thymine, and Their Monohydrates	372
5.2	Guanine	373
5.3	Cytosine	375
6	Monosaccharides	378
6.1	C ₄ Sugars: D-Erythrose	378
6.2	2-Deoxy-D-Ribose and Ribose	381
6.3	D-Glucose and D-Xylose	385
6.4	D-Fructose	391
7	Summary and Outlook	393
	References	394

1 Introduction

The study of the structure, function, reactivity, and properties of biomolecules covers a vast range of molecular sizes, from large macromolecules such as proteins, nucleic acids, lipids, or polysaccharides, down to small molecules such as amino acids, nucleic acid bases, neurotransmitters, or monosaccharides, to name but a few. Interest in this field of research has grown exponentially in the last few decades, across different areas, such as biology, medicine, chemistry, and physics, bringing with it a multitude of continuously developing experimental and theoretical methods. A major goal of this ongoing effort is to obtain information about the molecular structure of biomolecules and their related properties at the molecular level. Consequently, over the last few years, there has been a growing interest in the study of biomolecules in isolation conditions, that is, conditions matching those found in the gas phase [1], [2 and references therein], [3–6]. Gas-phase conditions are far from the condensed phase, *in vivo*, environment in which biomolecules work and biological reactions occur. Nevertheless, gas phase does provide ideal conditions to observe the intrinsic molecular properties, and thus it is a first step into the understanding of the molecular behavior in more complex media. For instance, in physiological media, many biological molecules are electrically charged species, because of processes of protonation (or deprotonation). A well-known example is

amino acids, which exist as doubly-charged zwitterions ($^+\text{NH}_3\text{-CH(R)-COO}^-$) [7, 8] in either solution or crystal. Moreover, in the gas phase, perturbing agents, such as solvents or even intermolecular interaction, are absent and thus observation of neutral species becomes possible. Furthermore, most biochemical molecules are characterized by a great torsional flexibility, resulting in manifold conformational varieties relatively close in energy. Since studies in condensed phases are considerably affected by multiple intermolecular interactions, conformational preferences can be biased because of the surrounding environment and only gas-phase studies can reveal the inherent structural minima of bare molecules.

Despite these advantages, the number of biomolecules studied to date in the gas phase is still very small compared with the overwhelming number of condensed-phase studies. Supersonic jets [9, 10], in combination with different spectroscopic techniques, have been instrumental in providing conformational information on small biomolecules. Modern supersonic-jet laser spectroscopy [2–4, 9, 10] combined with double resonance techniques (such as hole burning UV–UV and ion-dip IR–UV) and mass detection (through resonance-enhanced multiphoton ionization – REMPI), giving electronic and vibrational information with mass and conformer selectivity, have pioneered this field. This kind of experiments must be interpreted with the aid of high level *ab initio* quantum chemical calculations, thus allowing assignment of discrete conformational structures within a reasonable degree of confidence. However, unfortunately, laser spectroscopy is hampered by the need for an absorbing chromophore possessing a sharp vibronic structure, such as an aromatic ring [2, 11].

Unlike the aforementioned techniques, microwave spectroscopy can be applied to any gaseous molecular system of moderate size, the only requirement being to have a nonzero electric dipole moment [12]. Microwave spectroscopy has long been considered to be a powerful method for precise determination of gas-phase structures, being the source of the majority of gas-phase structural data known nowadays [13, 14] (see, for example, the Landolt-Börnstein, New Series II series). The combination of Fourier transform microwave spectroscopy with supersonic jets has given rise to two different families of techniques: those using a narrow band molecular beam (MB-FTMW) [15–17] or, more recently developed, wide band chirped pulse FTMW spectroscopy (CP-FTMW [18] and IMPACT-FTMW [19]). By virtue of the extremely high resolution (sub-Doppler) and sensitivity of the FTMW techniques, all populated species in the jet, be they tautomers, conformers, or isotopomers, can be analyzed independently. Moreover, small hyperfine effects arising from electric or magnetic interactions, such as nuclear electric quadrupole coupling and nuclear spin-spin coupling, can be fully resolved, providing insight to the molecule electronic properties. Finally, the observation of tunneling doublings in the case of large amplitude motions, such as those arising from internal rotation or inversion, can give information on the intramolecular dynamics associated with these motions. These techniques have been widely used in the study of gaseous or easily vaporized compounds, weakly bounded complexes [20–22], and even short-lived species generated *in situ* by electric discharges [23–27]. In our laboratory it has been used to study axial/equatorial equilibria [28–31] in

hydrogen-bonded complexes and weak C–H···O bonds [32–35], and has been combined with electric discharges to generate and characterize new unstable compounds [36–39].

The vast majority of biomolecules have extremely low vapor pressures at room temperature and are thermally fragile. For that reason, complete series of molecules of biological interest, such as building blocks, have escaped microwave spectroscopy studies. Several laboratories have incorporated heatable nozzles [40–42] into their microwave spectrometers to overcome the problem, but, in many cases, thermal instability constituted an important drawback. The pioneering works of Brown [43] or Suenram [44] and co-workers on molecular building blocks were carried out using such heating methods. Laser ablation is an efficient method for vaporizing solid samples, which otherwise would decompose upon heating [45]. Laser-based vaporization techniques coupled with spectroscopic techniques [46–49] have been used for gas-phase studies of solid compounds. However, until recently, there have been few attempts to incorporate laser ablation into MB-FTMW spectrometers [50–55]. Suenram et al. [50] and Walker and Gerry [51] independently developed two laser ablation devices in combination with molecular-beam Fourier transform microwave spectroscopy (LA-MB-FTMW) to investigate the rotational spectra of metal oxides and halides. The first was also applied in the structural study of glycine in the gas phase, but it was reported that the classical heating method was more reliable. Thus the experiments devoted to organic solids [54, 55] were initially discontinued because of the poor experimental results. Over the last few years, several MB-FTMW spectrometers which implement laser ablation sources to vaporize molecules in the throat of the nozzle have been configured in our laboratories. This approach, laser-ablation molecular-beam Fourier transform microwave (LA-MB-FTMW) spectroscopy [56], was initially tested with several organic solids [57, 58]. With the initial design [59] and progressive improvements [60, 61 and references therein], relevant biomolecules have been studied in order to identify and characterize their most stable conformations in the gas phase. This approach has been essentially followed in other laboratories [62]. More recently, laser ablation sources have been incorporated into broad-band CP-FTMW spectrometers [63, 64], extending the application of this method to a large number of thermally fragile systems with high melting points.

In the next sections, the experimental techniques, the procedure employed to characterize the different species observed, and the results obtained from the study of molecular building blocks such as amino acids, nucleic acid bases or sugars are discussed.

2 Experimental Techniques

2.1 Laser-Ablation Molecular-Beam Fourier Transform Microwave (LA-MB-FTMW) Spectroscopy

An extensive description of the experimental setup employed in our laboratory has been given elsewhere [56–61], but can be arranged into three main blocks: a laser ablation system, a Fabry–Pérot resonator, and an FTMW spectrometer, illustrated in Fig. 1. To accomplish vaporization efficiently, while retaining the advantages of the coaxial orientation of the resonator and supersonic jet axes, it was found necessary to modify the backside of the Fabry–Pérot resonator fixed mirror so that it can hold a specially designed laser ablation pulsed-jet nozzle. The samples are prepared as solid rods by pressing a fine powder of the pure compound with minimum quantities of a commercial binder. The rod is then placed vertically in the laser ablation nozzle, where a focused laser beam falls upon the sample laterally and the vaporized molecules are seeded in the supersonic jet expansion.

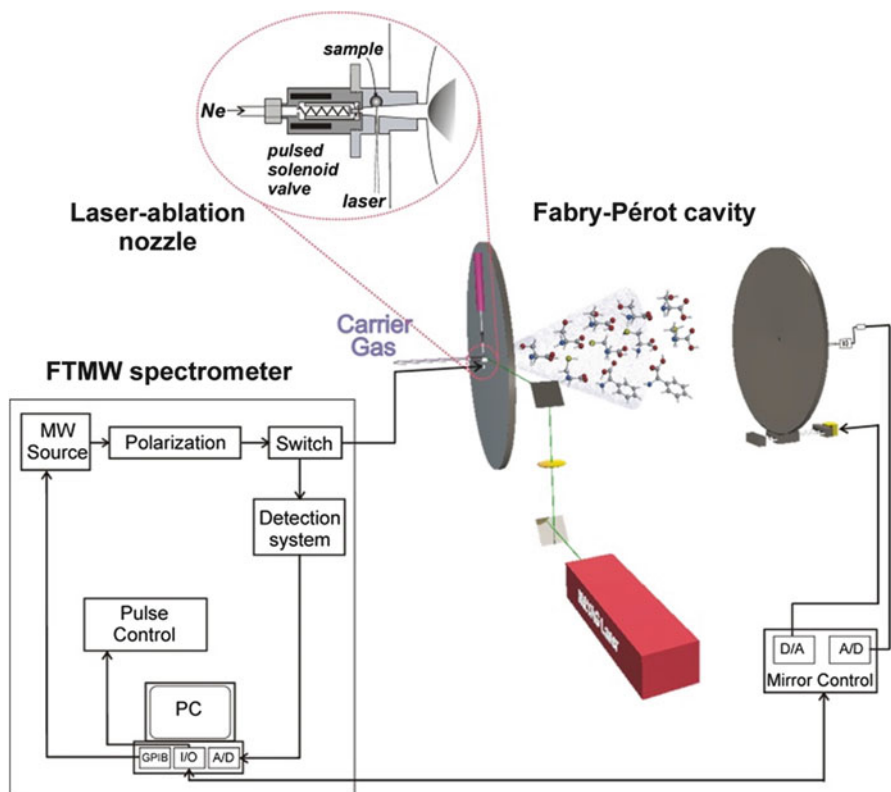


Fig. 1 Scheme of the LA-MB-FTMW spectrometer. (From [61])

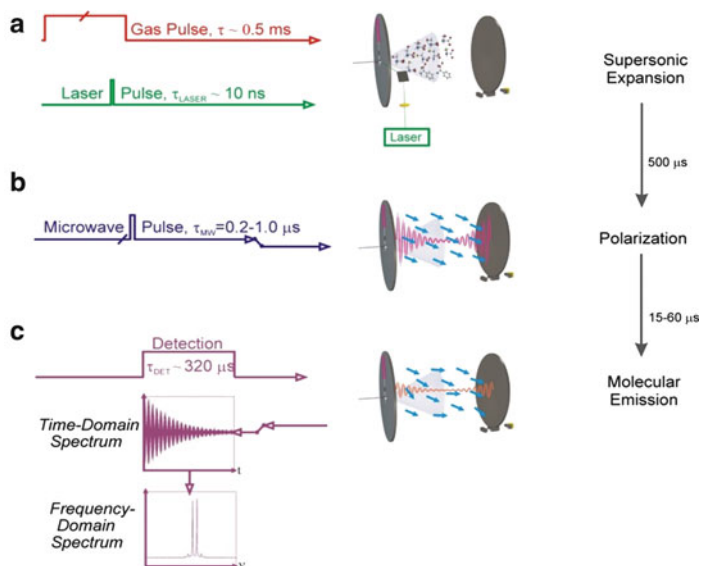


Fig. 2 Pulse sequence for a single experimental LA-MB-FTMW cycle. (From [61])

Reproducibility of the ablation process is improved by a continuous translation and rotation of the sample rod, exposing a new sample surface to each laser pulse. In our initial experimental setups, the second (532 nm), third (355 nm), or fourth (266 nm) harmonics of a nanosecond Q-switched Nd:YAG laser (ca. 50 mJ/pulse) have been used. Presently, the use of the harmonics of an Nd:YAG 20–35 ps lasers (7–15 mJ/pulse), which have been proved to be more efficient [see, for example, 65], is considered standard. The laser-ablation nozzle has been modified from its original design to provide a smoother transition between the aperture at the solenoid valve and that of the Fabry–Pérot mirror (see insert in Fig. 1), nowadays resembling a Laval nozzle [66].

A typical experimental cycle (see Fig. 2) starts with a pulse of a noble carrier gas (stagnation pressures 2–8 bar, gas pulse typically 0.5 ms). After an adequate delay, a short laser pulse hits the sample rod producing, the vaporization of the solid. The ablated molecules are then seeded in the carrier gas, which expands supersonically between the two mirrors of the Fabry–Pérot resonator. In the supersonic expansion, the seeded molecules suffer a strong cooling of the rotational and vibrational degrees of freedom, and individual conformers are frozen into the ground vibrational state of the corresponding potential energy well. Thus, the conformer distribution before the expansion may be preserved provided the interconversion barriers between conformers are high enough. Molecular collisions gradually disappear as the expansion proceeds, in such a way that the different species can be probed in a virtually isolated environment by Fourier transform microwave (FTMW) spectroscopy. A short microwave pulse (typically 0.3 μ s) is subsequently applied, which produces a macroscopic polarization of the species in the jet. Once the excitation

pulse ceases, the molecular emission signal (FID, free induction decay) is captured in the time domain. Its Fourier transformation to the frequency domain yields the molecular rotational transitions which appear as Doppler doublets because the supersonic jet travels parallel to the resonator axis. The molecular equilibrium frequencies are calculated as the arithmetic mean of the Doppler doublets, and are obtained with accuracy better than 3 kHz. A new experimental cycle can start once the vacuum cavity has been evacuated, and a repetition rate of 2 Hz is normally employed. For very weak signals, hundreds of cycles must be added coherently. To probe the jet at different microwave frequencies or conduct a frequency scan, the Fabry–Pérot resonator is tuned mechanically under computer control. The timing of the whole experiment, in particular the delay of the laser with respect to the molecular pulse, is crucial for an optimum signal. Laser pulse energy is also critical.

2.2 *Chirped-Pulse Fourier Transform Microwave Spectroscopy (CP-FTMW) Coupled to Laser Ablation (LA)*

A new approach of high-resolution rotational spectroscopy, providing a sensitive method for broadband detection (CP-FTMW), has recently been developed at the University of Virginia [18]. Such an instrument is based on the same principles as the original MB-FTMW instrument [15], but the burst of microwaves has been replaced by a microwave pulse which contains a fast linear sweep (chirp) over the entire frequency region being explored and is used to polarize the sample. This new approach increases the speed of the spectral acquisition, making the search of the different coexisting species in the jet much more efficient. In the last few years, a new alternative design of broadband Fourier transform microwave spectroscopy, called IMPACT (in-phase/quadrature-phase modulation passage acquired coherence technique) has been developed in our laboratory [19]. The aforementioned broadband techniques have been combined with a laser ablation source for the study of solid biomolecules [63, 64].

A schematic block diagram of the newly designed chirped-pulse Fourier transform microwave spectrometer CP-FTMW, combined with a *ps*-pulsed laser ablation system, is given in Fig. 3. The spectrometer, which uses the basic operation of the CP-FTMW instrument [18], is described elsewhere [63], with only the relevant details to this experiment being described here. It operates in the 6.0–18 GHz region. The solid sample, prepared as usual as a rod shape, was placed in a laser ablation nozzle, similar to that previously described [61] (1 in Fig. 3) and vaporized using the second (532 nm) or third (355 nm) harmonics of a *ps* Nd:YAG laser (i.e.: Ekspla, 20 ps, 15 mJ/pulse) (2 in Fig. 3). A motor controller (3 in Fig. 3) allows a DC motor (Oriol Motor Mike 18074) (4 in Fig. 3) to rotate and translate the rod up and down along the injection system to achieve the maximum exploitation of the

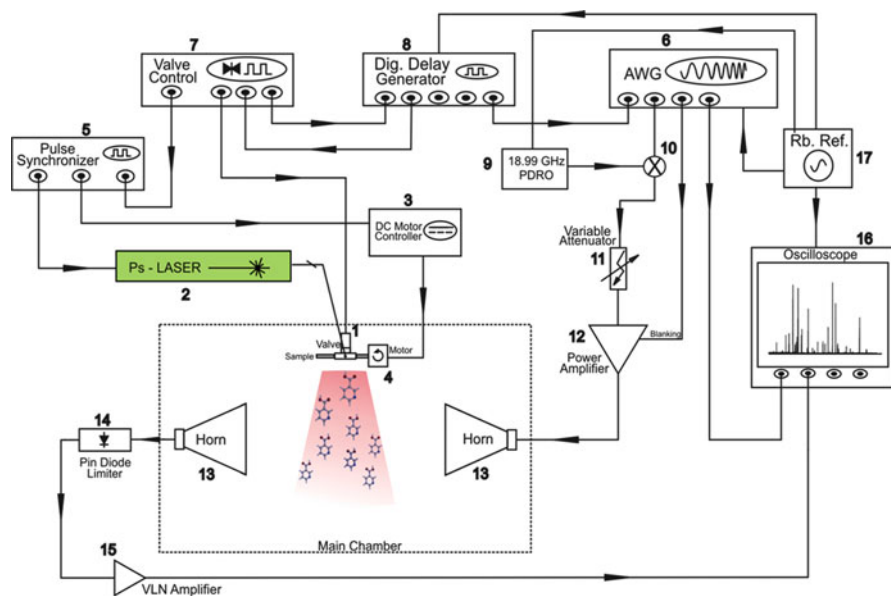


Fig. 3 Schematics of a laser ablation chirped-pulse Fourier transform microwave spectrometer. (From [63])

sample. So that good vacuum conditions are retained in the expansion chamber, and to reduce sample consumption, the repetition rate of the experiment is set to be 2 Hz, achieved via a shutter operating at said frequency, placed at the output of the laser and controlled via a pulse synchronizer (5 in Fig. 3), used to adjust the repetition rate to 2 Hz from the standard rate of the laser (10 Hz). In the ablation nozzle, the sample molecules, once vaporized, are seeded in the flow of the carrier gas (Ne at backing pressure of 15 bar), being expanded adiabatically into the vacuum chamber, where they are probed by a microwave chirped pulse.

Briefly, the broadband microwave spectrometer works as follows. A 24 GS/s arbitrary waveform generator (6 in Fig. 3) creates a fast chirp microwave pulse, covering the entire range to be explored, to polarize macroscopically the pulsed molecular beam sample created from the ablation nozzle by the valve driver (7 in Fig. 3). A digital delay generator (8 in Fig. 3) is used to trigger both the arbitrary waveform generator and the valve driver. The microwave pulse is upconverted by mixing it in a broadband mixer (10 in Fig. 3) with an 18.99 GHz signal provided by a phase-locked dielectric resonator oscillator (PDRO) (9 in Fig. 3). The upconverted signal is subsequently amplified by a 300 W traveling wave tube amplifier (12 in Fig. 3). The power level necessary for the polarization of the molecular systems can be adjusted using a variable attenuator (11 in Fig. 3). The amplified chirp pulse is broadcast across the vacuum chamber, where the jet expansion occurs, using one of the two standard horn antennas (13 in Fig. 3) placed inside the chamber. These two horn antennas are separated by approximately

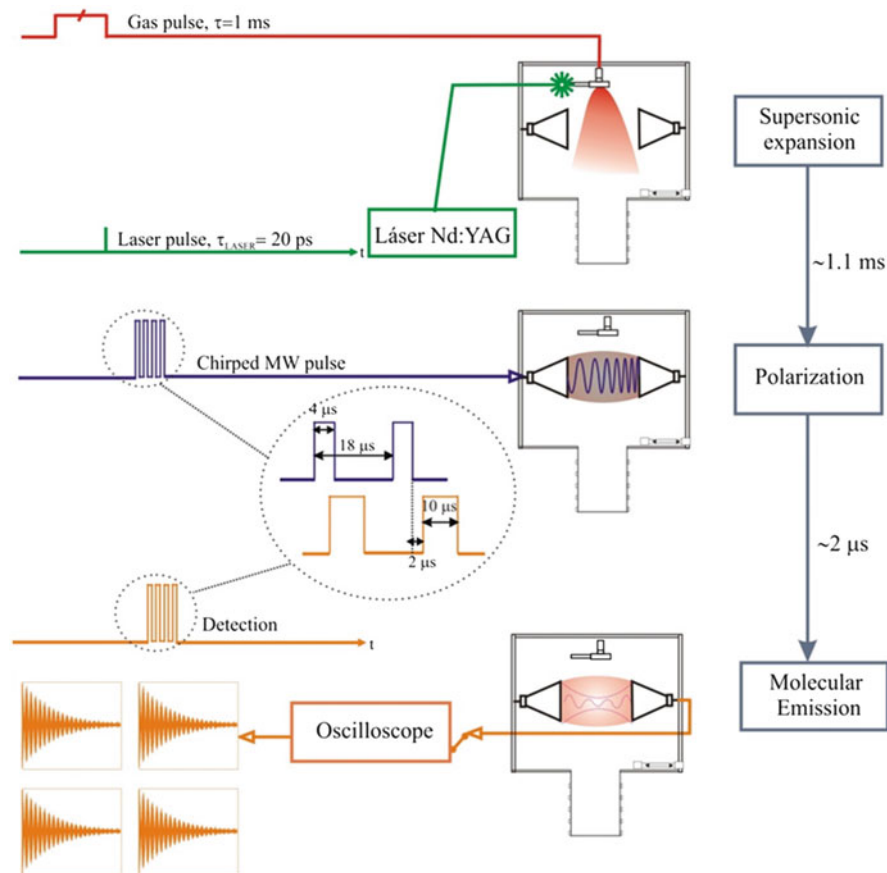


Fig. 4 Pulse sequence for a single experimental cycle including generation of a supersonic expansion and laser ablation, polarization and detection. (From [63])

60 cm. The second antenna is used to detect the free induction decay signal (FID) emitted by the sample as response to the microwave excitation. The FID is further amplified by a sensitive VLN amplifier (14 in Fig. 3) which is protected from the high-power of the TWT amplifier by a pin diode limiter (15 in Fig. 3). The amplified rotational free induction decay (FID) is recorded in the time domain by a digital oscilloscope (50 GS/s, 20 GHz hardware bandwidth) (16 in Fig. 3) and Fourier transformed to the frequency domain. The phase reproducibility of the experiment is achieved by locking all frequency sources and the digital oscilloscope to a 10-MHz rubidium frequency standard oscillator (17 in Fig. 3).

The operation sequence (see Fig. 4) starts with a molecular pulse of 1,000 μs duration which drives the carrier gas flow through the pulsed valve source. After an adequate delay (~850 μs), a laser pulse hits the solid and vaporizes the sample. To reduce sample consumption, four separate broadband rotational spectra are

acquired in each injection cycle. The four individual broadband-chirped excitation pulses, of $4 \mu\text{s}$ width, are spaced by $18 \mu\text{s}$. Then, $2 \mu\text{s}$ after each excitation pulse ceases, the rotational free induction decay is acquired for $10 \mu\text{s}$. The internal pulse generator of the valve driver is used to create the digital pulses involved in the laser generation. Since the sample injection has a perpendicular arrangement with the microwave field, the transit time of the polarized molecular jet is quite short and linewidths of about 100 kHz full-width-half-maximum (FWHM) are achieved.

3 Tools in Conformational Analysis

The high flexibility of biomolecules produces the appearance of a significant number of low-energy conformers. Whereas covalent forces determine the molecular skeleton, conformational isomerism is also controlled by weaker nonbonded interactions within the molecule, especially hydrogen bonding. To treat this conformational problem and identify the different structures in the supersonic jet, a procedure represented in Fig. 5 has been followed in all the microwave studies. It is illustrated through its application to L-threonine [61] in this section and it was also described in our study of glutamic acid [67].

3.1 Model Calculations

To have an overall picture of the conformational landscape of L-threonine, theoretical predictions are used to find the most stable conformers on the potential energy surface. Only low-lying energy conformers are sufficiently populated in the supersonic jet to be observed in the rotational spectrum. Starting geometries for ab initio calculations are initially selected by considering all possible rotations around single bonds and identifying plausible intramolecular hydrogen bonds. To predict the most

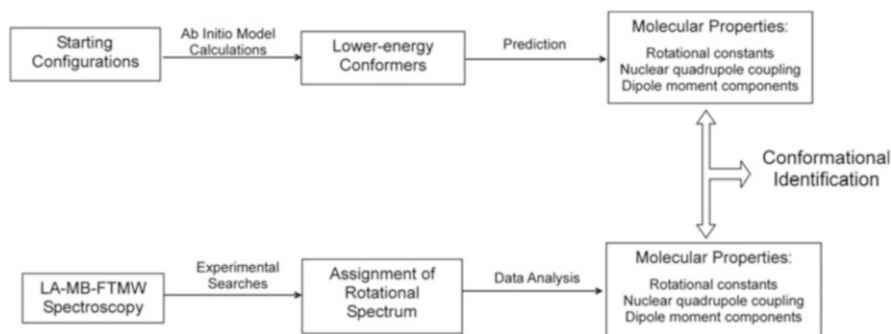


Fig. 5 Steps followed in the identification of the conformers of biomolecules. (From [61])

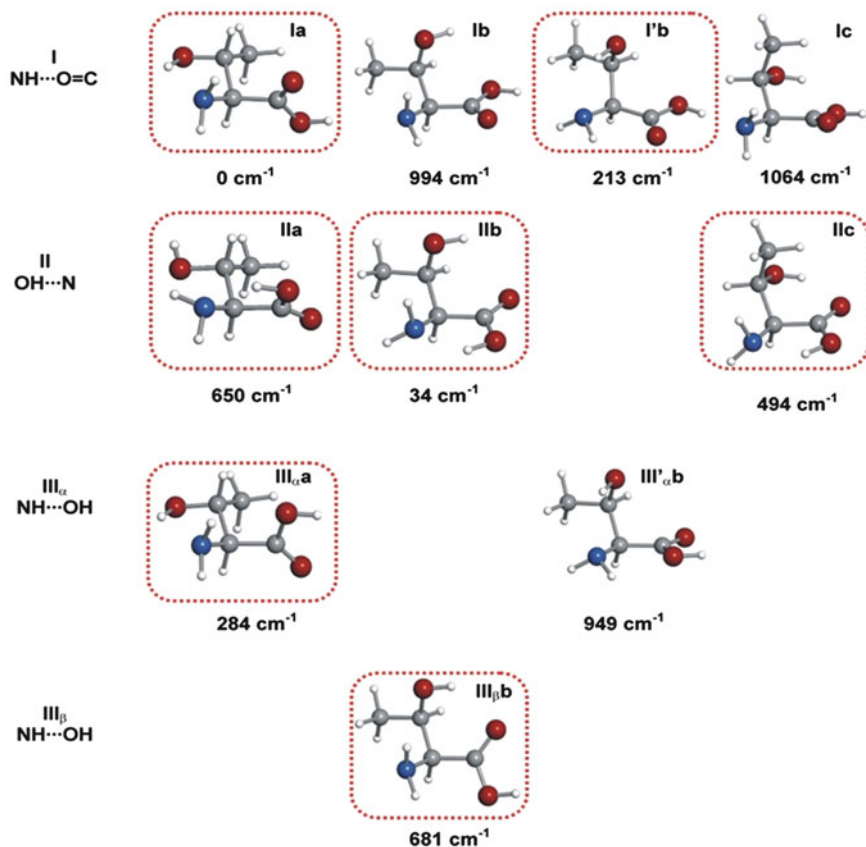


Fig. 6 Predicted low-energy conformers of threonine and energies relative (MP2/6-311++G(d,p)) to the global minimum in cm^{-1} . The detected conformers are encircled. (From [61])

stable forms, a series of structural optimizations is conducted for each of the starting configurations using the Gaussian suite of programs [68]. First, a cheap but computationally efficient calculation at the B3LYP/6-31G(d,p) level [69–73] is performed. In a second step, full geometry optimization calculations using second-order Møller–Plesset perturbation theory (MP) [74] and Pople’s 6-311++G(d,p) [75] basis set are conducted. This level of theory has been found to behave satisfactorily in all the series of amino acids studied (see below). Finally, the ten conformers of Fig. 6 were predicted to be within $1,000 \text{ cm}^{-1}$. Conformers are classified as I, II, or III depending on the hydrogen bond established between the amino and carboxylic groups and as *a*, *b*, or *c* depending on the configuration adopted by the $-\text{CH}(\text{CH}_3)\text{OH}$ side chain.

A distinguishing quality of rotational spectroscopy lies in its capability to generate very accurate spectroscopic parameters directly comparable with in vacuo ab initio predictions to provide unequivocal evidence of the conformers

observed. Three sets of spectroscopic constants are relevant for the interpretation of the rotational spectra: rotational constants, nuclear quadrupole coupling constants, and electric dipole moment components. From the ab initio optimized structures of the lower-energy conformers of threonine (Fig. 6), those spectroscopic constants are calculated and listed in Table 1.

The rotational constants A , B , and C , which provide information on the mass distribution of the molecules [12, 76, 77], are compared with those obtained experimentally to identify the conformers present in the supersonic expansion. Such comparison is normally conclusive in the identification, but sometimes the difference in the values between conformers is not large enough to allow discrimination between them. This happens with some forms belonging to the same family in Table 1 (see, for example, Ia and IIa). In these cases, a different and independent way of identifying conformers based on the ubiquitous presence of ^{14}N in amino acids can be utilized. ^{14}N nuclei possess a nonzero quadrupole moment ($I = 1$) because of a non-spherical distribution of the nuclear charge which interacts with the electric field gradient created by the rest of the molecule at the site of those nuclei. This results in a nuclear hyperfine structure in the rotational spectrum [12, 76, 77]. FTMW spectroscopy provides the high resolution needed to resolve fully the various quadrupole hyperfine components (see Fig. 7). The associated experimentally-determined molecular properties are the quadrupole coupling constants χ_{gg} ($g = a, b, c$), referred to the principal inertial axes, which are directly related to the electronic environment of the quadrupolar nucleus referred to the principal inertial axes and strongly depend on the orientation of the amino group. For example, conformers Ia and IIa with similar predicted rotational constants have different orientations of the amino group (Fig. 7), which is reflected in the nuclear quadrupole coupling constants χ_{aa} and χ_{cc} , as can be seen by the predicted values in Table 1.

A third molecular property to be considered is the electric dipole moment. The selection rules and intensity of the rotational transitions of asymmetric tops depend on the dipole moment components along the principal inertial axes, that is, on μ_a , μ_b , and μ_c , which give rise to a-, b-, and c-type spectra, respectively. All conformers possess the same connectivity between the atoms but they differ in the orientation of the functional groups, and this necessarily produces diverse charge distributions reflected in different values of the dipole moment components, as can be seen in Table 1. The microwave power necessary for optimal polarization depends on the dipole moment component involved in a rotational transition. Hence, the difference in the values of the dipole moment components of conformers can be exploited to discriminate between specific conformers just by varying the polarization power. By itself, it cannot be used as a conclusive tool, but it can always corroborate the conformer identification achieved with the previously described molecular properties.

The electric dipole moment components are necessarily used in estimating relative conformational abundances, as is explained below. The predicted values of the electric dipole moment components of the threonine conformers are collected in Table 1.

Table 1 Ab initio spectroscopic constants for the low-energy conformers of threonine of Fig. 6. (From [61])

MP2/6-311++G(d,p)	Ia	IIa	III _a a	Ib	I'b	IIb	III' _a b	III _b b	Ic	IIc
A ^a (MHz)	2,870	2,906	2,885	3,163	3,164	3,233	3,194	3,383	2,488	2,662
B (MHz)	1,608	1,668	1,578	1,542	1,521	1,543	1,498	1,488	1,838	1,793
C (MHz)	1,225	1,201	1,249	1,288	1,324	1,271	1,332	1,239	1,431	1,395
χ_{aa} (MHz)	-4.47	-0.44	-4.53	-2.82	-0.35	-3.72	0.57	-2.32	-2.32	-3.94
χ_{bb} (MHz)	2.76	2.67	2.85	1.89	2.94	1.88	2.96	-0.03	0.88	2.47
χ_{cc} (MHz)	1.71	-2.23	1.69	0.93	-2.59	1.84	-3.52	2.44	1.43	1.46
$ \mu_d $ (D)	2.3	4.3	2.1	0.3	0.1	3.4	0.2	2.3	1.0	0.1
$ \mu_{pl} $ (D)	0.1	2.2	1.5	0.5	0.6	3.3	2.8	1.1	1.0	5.6
$ \mu_d $ (D)	0.8	0.3	1.3	0.9	2.9	1.2	1.9	1.0	0.3	0.3

Source: From [61]

^aA, B, and C are the rotational constants; χ_{aa} , χ_{bb} , and χ_{cc} are the diagonal elements of the ^{14}N nuclear quadrupole coupling tensor; μ_a , μ_b , and μ_c are the electric dipole moment components

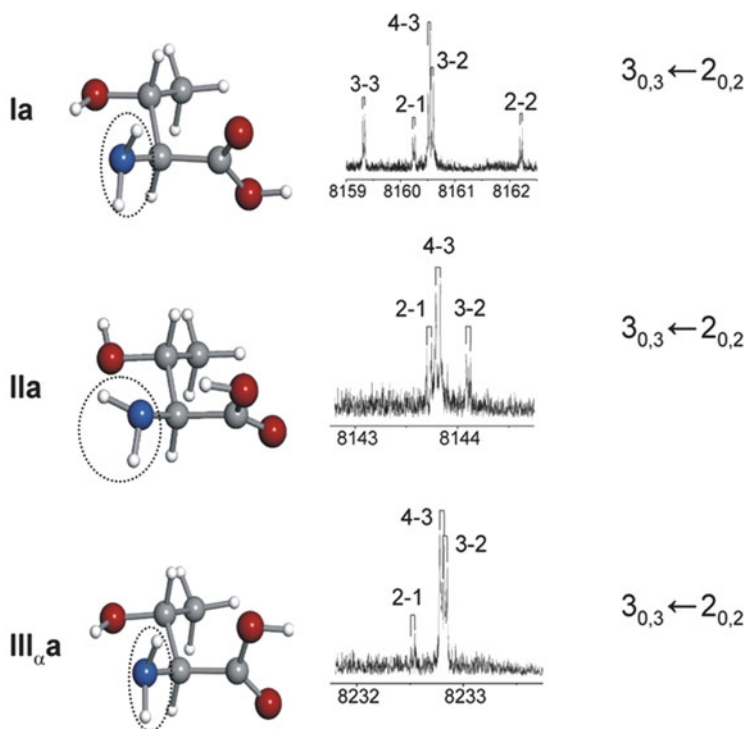


Fig. 7 Nuclear quadrupole hyperfine structure of the $3_{0,3} \leftarrow 2_{0,2}$ rotational transitions observed for rotamers M, N, and O, identified as IIa, Ia, and III_αa, respectively. (From [61])

3.2 Analysis of Spectra and Conformer Identification

According to the values of the predicted electric dipole moment components and rotational constants listed in Table 1, the rotational spectra of most of the threonine conformers are dominated by μ_a -type, R-branch transitions forming groups of lines with characteristic patterns which appear at frequency intervals equivalent to $B + C$. Wide frequency scans with low power polarization conditions were conducted to search for such rotational transitions of conformers with relatively large μ_a . Several sets of R-branch lines corresponding to five different rotamers are labeled L, M, N, O, and P. Spectral searches, conducted to detect other sets of μ_a -type R-branch transitions, using high microwave power for polarization, were unsuccessful. Therefore, wide frequency scans were carried out to identify μ_b - and μ_c -type transitions. Finally, two more rotamers labeled Q and R were identified.

Apart from the instrumental Doppler doubling, all transitions are observed to be split into several close hyperfine components (see Fig. 7) arising from the ^{14}N nuclear quadrupole interaction described above, indicating that they arise from a molecule with a nitrogen nucleus. Successive predictions and new experimental

measurements discarded or confirmed the initial assignment until a group of consistent rotational transitions, including all possible μ_a -, μ_b -, or μ_c -transitions, was collected for each conformer.

The observed transitions frequencies were analyzed with a Watson's semirigid rotor Hamiltonian ($H_R^{(A)}$) [12] supplemented with a nuclear quadrupole coupling term (H_Q): $H = H_R^{(A)} + H_Q$. The Hamiltonian is constructed in the coupled basis set $I+J=F$ and diagonalized. The A-reduced semirigid Watson Hamiltonian in the F' representation is given by

$$H_R^{(A)} = AP_a^2 + BP_b^2 + CP_c^2 - \Delta_J P^4 - \Delta_{JK} P^2 P_a^2 - \Delta_K P_a^4 - 2\delta_J P^2 (P_b^2 - P_c^2) - \delta_K [P_a^2 (P_b^2 - P_c^2) + (P_b^2 - P_c^2) P_a^2], \quad (1)$$

where the coefficients A , B , and C represent the rotational constants and Δ_J , Δ_{JK} , Δ_K , δ_J , and δ_K are the quartic centrifugal distortion constants. Only Δ_J needed to be floated to obtain an rms deviation of the fit consistent with the estimated frequency accuracy. The term H_Q accounts for the interaction energy of the ^{14}N electric quadrupole moment (eQ) with the molecular electric field gradient ($q_{\alpha\beta} = \partial^2 V / \partial\alpha\partial\beta$; $\alpha, \beta = a, b, c$) at the nitrogen nucleus. The determinable spectroscopic parameters are the elements of the nuclear quadrupole coupling tensor χ , linearly related to the electric field gradient by $\chi = -eQq$. Usually, only the diagonal elements of the tensor (χ_{aa} , χ_{bb} , χ_{cc}) are determined.

The experimental rotational constants and the χ_{aa} , χ_{bb} , and χ_{cc} nuclear quadrupole coupling constants for the rotamers L, M, N, O, P, Q, and R are collected in Table 2. A first look at the rotational constants of Table 2, and their comparison with the predicted constants of Table 1, allows us to classify easily rotamers as belonging to different families. Hence, rotamers M, N, and O belong to the "a" family, rotamers L, P, and Q belong to the "b" family, and, finally, rotamer R belongs to the "c" family. Conformers belonging to the "a" family have similar mass distributions: their rotational constants are very similar and they do not allow further discrimination. In these cases, as mentioned above, conclusive evidence comes from the values of ^{14}N quadrupole coupling constants because they are very sensitive to the orientation of the $-\text{NH}_2$ group with respect to the principal inertial axis system (see Table 1). The values of these constants, reflected in the hyperfine structure clearly identify rotamer M as IIa (see Fig. 6).

Rotamers N and O have very similar quadrupole coupling constants; comparison of their values with those predicted ab initio indicate that they are necessarily conformers Ia and III $_{\alpha}$ a (see Tables 1 and 2). Because of the similar orientation of the amino group in these forms (see Fig. 6), they cannot be discriminated on the basis of the quadrupole constants. We can distinguish them from their selection rules and intensities of the observed transitions. The rotational spectrum of rotamer N shows strong μ_a -type transitions and fairly weak μ_c -type transitions, while form O presents strong μ_a -type transitions and medium-strength μ_b - and μ_c -type transitions. No μ_b -type transitions have been detected for conformer N. Considering the predicted dipole moment components of Table 1, these data are consistent with

Table 2 Experimental spectroscopic constants for the seven observed rotamers of threonine

Experimental	L	M	N	O	P	Q	R
A ^a (MHz)	3,379.841(14) ^b	2,912.6227(20)	2,872.77049(48)	2,889.93352(45)	3,232.4827(12)	3,148.59247(32)	2,670.72096(53)
B (MHz)	1,482.04984(21)	1,660.21807(34)	1,608.95699(26)	1,572.32152(50)	1,533.71801(32)	1,506.27679(37)	1,784.66894(60)
C (MHz)	1,237.59121(22)	1,189.31443(34)	1,211.39762(38)	1,241.83423(47)	1,267.88615(34)	1,316.33575(44)	1,383.75384(51)
κ	-0.77	-0.45	-0.52	-0.60	-0.73	-0.79	-0.38
χ_{aa} (MHz)	-2.201(14)	-0.544(11)	-4.1859(25)	-4.1529(32)	-3.4971(21)	-0.7403(21)	-3.7652(73)
χ_{bb} (MHz)	-0.157(50)	2.582(16)	2.6611(42)	2.5682(46)	1.7519(27)	2.8781(28)	2.4258(75)
χ_{cc} (MHz)	2.358(64)	-2.038(50)	1.5248(17)	1.5846(46)	1.7452(60)	-2.1378(70)	1.3394(20)

Source: From [61]

^aA, B, and C are the rotational constants; χ_{aa} , χ_{bb} , and χ_{cc} are the diagonal elements of the ^{14}N nuclear quadrupole coupling tensor

^bStandard error in parentheses in units of the last digit

the identification of rotamer N as conformer Ia and rotamer O as conformer III_αa. The microwave power used to polarize optimally the rotational transitions is also in agreement with the predicted values for the dipole moment components of each conformer.

The rotamers L, P, and Q which belong to the “b” family, can be attributed to I’b, Ib, IIb, or III_βb considering the values of the rotational constants. However, detailed comparison of the values of the ¹⁴N quadrupole coupling constants conduct to the identification of rotamer L as conformer III_βb, P as IIb, and Q as I’b. In this case, the quadrupole coupling constants are essential for the discrimination of conformers.

Again, from the rotational constants alone it is only possible to match rotamer R to a conformer belonging to the “c” family, Ic or IIc. Consideration of the nuclear quadrupole coupling constants uniquely identifies form R as IIc.

The procedure illustrated here to analyze the spectrum of threonine and to identify the observed conformers has been followed throughout all the examples shown in the next sections.

4 Amino Acids

4.1 Introduction

An important subset of biologically relevant molecules is that of natural amino acids, the so-called “building blocks” of peptides and proteins. Amino acids have long been studied in solids [78–82] and in solution [83–85], where they are stabilized as doubly charged species or zwitterions (R–CH(NH₃⁺)–COO[–]). Gas-phase studies have the advantage of providing information on the neutral forms of amino acids (R–CH(NH₂)–COOH, the canonical forms present in peptide chains) and on their inherent molecular properties free from the intermolecular interactions occurring in the condensed media. Furthermore, gas-phase data can be easily contrasted with theoretical models and used to refine the latter.

Figure 8 shows the 20 proteinogenic or coded α-amino acids (NH₂–CH(R)–COOH) which can be classified according to the nature of their side chain (R). The α-amino acid backbone determines the primary sequence of a protein, but the nature of the side chains determines the protein properties [86–88]. For this reason, examination of the structural properties of amino acids with different side chains is important to the understanding of its functionality. Knowledge of the structures and conformational behavior of amino acids has been considerably expanded thanks to the extensive use of LA-MB-FTMW spectroscopy. Only the simplest of these amino acids, glycine [43, 44, 89–92] and alanine [93], were studied before by rotational spectroscopy, using heating methods. The amino acids encircled in a solid line in Fig. 9 have been studied with this technique in our laboratory and some of them are illustrated in this chapter.

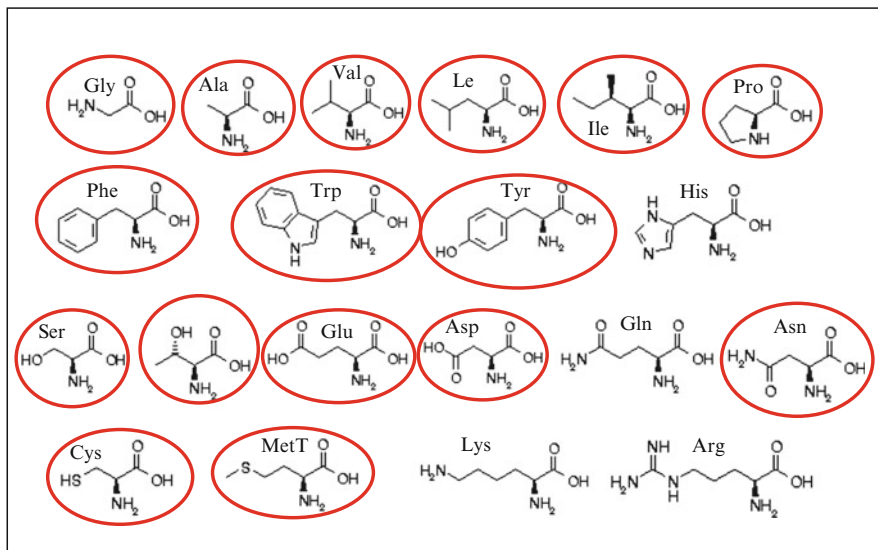


Fig. 8 Coded α -amino acids. The amino acids encircled have been studied by rotational spectroscopy

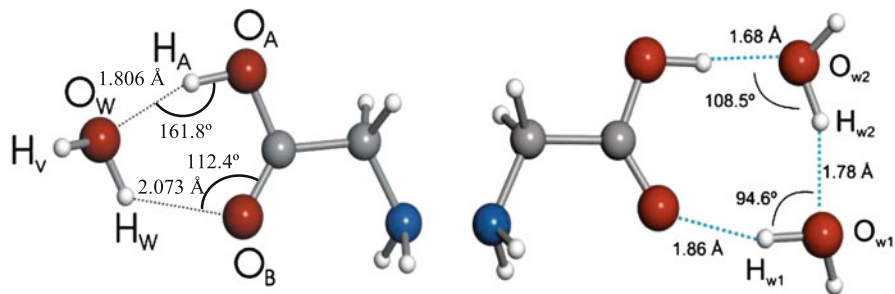


Fig. 9 Structure of the glycine- $(\text{H}_2\text{O})_n$ complexes (From [100, 101])

The large flexibility of amino acids, which makes folding and unfolding of proteins possible, produces the appearance of a significant number of low-energy conformers. Whereas covalent forces determine the molecular skeleton, conformational isomerism is also controlled by weaker nonbonded interactions within the molecule, especially hydrogen bonding. Hydrogen bonding between the amino and carboxylic moieties is expected to control the configuration in non-polar side chain amino acids. The presence of a polar side chain brings about new sets of intramolecular interactions between the lateral side chain and the amino or acid groups, which greatly increase the number of low-energy forms. In β - or γ -amino acids, the balance of forces may change with the length of the chain between the amino and acid groups and, as shown below, new interacting forces may contribute to

conformer stabilization. In the next sections, the conformational behavior of the different α -amino acids as observed from LA-MB-FTMW spectroscopy is described.

4.2 Proteinogenic Amino Acids with Non-polar Side Chains

The conformational behavior of α -amino acids with non-polar side chains is essentially determined by the three possible intramolecular hydrogen bonds established between the amino and carboxylic functional groups. Configuration I is stabilized by a hydrogen bond between the hydrogen atom of the amino group and the carbonyl oxygen of the carboxylic group ($\text{N-H} \cdots \text{O}=\text{C}$), and a *cis* arrangement of the $-\text{COOH}$ group. In configuration II, the hydrogen bond links the hydrogen atom of the hydroxyl group with the electronic lone pair at the nitrogen atom ($\text{N} \cdots \text{H}-\text{O}$) while the $-\text{COOH}$ group exhibits a *trans* arrangement. Configuration III bears a hydrogen bond between the amino group and the oxygen atom of the hydroxyl group ($\text{N-H} \cdots \text{O}-\text{H}$) and presents a *cis*- COOH . On this basis, the conformational behavior of glycine, alanine, valine, leucine, and isoleucine are presented in the next sections.

4.2.1 Glycine and Its Hydrates

Glycine ($\text{R}=\text{H}$, m.p. = 240°C) is the smallest α -amino acid, and has been the object of different studies, so its structural properties in the gas phase are well understood. Following the simultaneous discovery by Brown et al. [43] and Suenram et al. [44] of glycine II, a subsequent search led to the detection of the weaker glycine I [89, 90]. Further work using jet expansions provided additional information on isotopic species, electric dipole moment, and ^{14}N nuclear quadrupole coupling hyperfine structure [92–93].¹ The rotational data are consistent with ground-state structures of C_s symmetry for glycines I and II. Recently, the rotational data have been extended into the millimeter-wave region [94] to improve radioastronomical searches.

These studies provided fundamental pieces of information about the dominant role of intramolecular hydrogen bonding between the polar moieties of the amino acid skeleton in molecular conformation, concluding that glycine rotamers are stabilized by intramolecular hydrogen bonding from either amine to carbonyl $\text{N-H} \cdots \text{O}=\text{C}$ (type I conformer) or hydroxyl to the amine nitrogen lone pair $\text{N} \cdots \text{H}-\text{O}$ (type II conformer). The third low-energy conformer III, with an intramolecular hydrogen bond $\text{N-H} \cdots \text{O}-\text{H}$, could not be detected. The experimental evidence, including a recent reexamination of the glycine spectrum in our laboratory giving a conformational ratio $\text{N}_\text{I}/\text{N}_\text{II} \sim 6/1$ (see footnote 1), concludes that conformer I is the global minimum, a result also supported by multiple theoretical studies [95–99].

¹ Investigation of the parent, $^{15}\text{NH}_2\text{CH}_2\text{COOH}$ and $^{15}\text{NH}_2\text{CH}_2^{13}\text{COOH}$ glycine species, partially reported as supplementary material of [100].

Hydration is known to trigger the transformation of amino acids from the neutral form observed in the gas phase to the charged zwitterion present in condensed media. The glycine–water complexes provide the simpler molecular models of the biologically important amino acid–water interaction, representing the initial steps of the hydration process. Although their generation is difficult in experiments where laser ablation is involved because of the formation of a hot plasma, the improvements in the LA-MBFTMW instruments by modifying the nozzles and incorporating picosecond (20–35 ps) laser technology [59, 61, 67] has rendered possible the observation of glycine microsolvates [100, 101]. In both cases, glycine–H₂O and glycine–(H₂O)₂, only conformer I was observed (see Fig. 9). As for bare glycine, conformer III was not detected.

On the basis of the excellent agreement between experimental and theoretical spectroscopic constants and the use of enriched samples of ¹⁵N and H₂¹⁸O, it was possible to determine the structure of the complexes shown in Fig. 9. The structures provide relevant chemical information on the nature of the hydrogen bond interactions which stabilize the adducts. Binding of one or two water molecules to glycine proceeds through the carboxylic group and gives rise to a closed ring structure in which the water bridges benefit from enhancing cooperative effects [102–104]. The structures of glycine–H₂O and glycine–(H₂O)₂ retain the preferred conformation I (NH $\cdot\cdot$ O=C, *cis*-COOH) of unsolvated glycine, so the water molecules in the complexes do not enter into competition with the intramolecular hydrogen bonds in glycine. Comparisons of the O–H $\cdot\cdot$ O hydrogen bond distances and angles show that the acidic character of the OH carboxylic group dominates the interactions between water and glycine.

4.2.2 Alanine

Alanine (R=CH₃) (m.p. = 315°C) was first studied in the gas phase by Godfrey et al. [93] from free-jet millimeter-wave absorption spectroscopy and observed two conformers of five possible low energy conformers. They were identified as I and II according to the type of their intramolecular hydrogen bonds. However, for conformer II, the torsion of the –COOH group gives rise to two close energy conformations. Because of the similar mass distribution of both structures, it was not possible to discriminate between them on the basis of the rotational constants alone. A later investigation of the rotational spectrum by high resolution LA-MB-FTMW spectroscopy [105] allowed one to resolve fully the ¹⁴N quadrupole coupling hyperfine structure of the two observed conformers and confirmed their assignment as alanine I and IIa. The derived mean ratio N_I/N_{IIa} = 3.7(5) confirmed that conformer I is the most stable species in the jet.

In addition, the rotational spectra of ten different isotopologues (one parent, one ¹⁵N, three ¹³C and five deuterated species) for the two most stable conformers of alanine were detected. The extensive isotopic data were analyzed to derive the substitution [106, 107] and effective structures for both I and IIa conformers. Unlike glycine, the amino acid skeleton in alanine is non-planar. Deuteration at the amino and

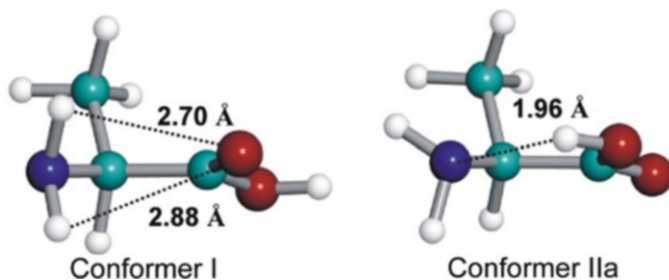


Fig. 10 The structures of conformers I and IIa of alanine determined in [105], showing the hydrogen bond distances

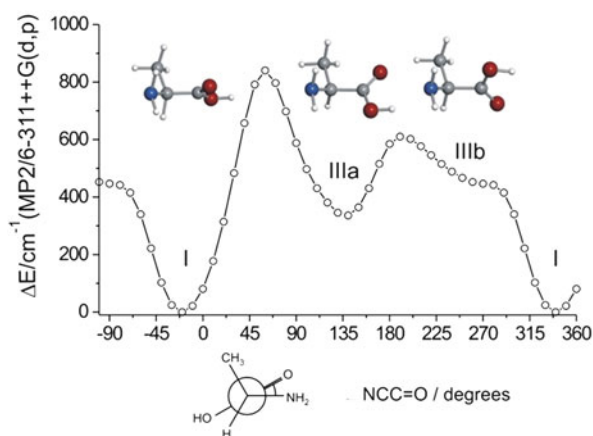


Fig. 11 Calculated MP2/6-311++G(d,p) energy profile of alanine along the torsional coordinate defined by dihedral angle $\angle\text{NCC}=\text{O}$. (From [105])

hydroxyl hydrogen atoms rendered information on intramolecular interactions: the data are consistent with the formation of an intramolecular hydrogen bond $\text{N}\cdots\text{H}-\text{O}$ in conformer IIa and with a non-symmetrically bifurcated hydrogen bond $\text{N}-\text{H}\cdots\text{O}=\text{C}$ established by the two amino hydrogen atoms in conformer I (see Fig. 10).

The absence of type III conformers which are predicted to fall in the group of lowest energy forms deserves some comment. This problem of “missing conformers” first observed in glycine was attributed by Godfrey et al. [93] to conformational relaxation, plausibly a $\text{III} \rightarrow \text{I}$ interconversion. It is accepted that this event takes place by collisions with the noble carrier gas in the adiabatic expansion. This phenomenon has been observed in systems involving only one degree of freedom, such as torsional isomerism [108–110] and axial/equatorial relaxation in hydrogen-bonded complexes [30, 111], when energy barriers are less than about 400 cm^{-1} . The ab initio calculation of a section of the potential energy surface for the interconversion $\text{III} \rightarrow \text{I}$ along the rotation of the $-\text{COOH}$ group predicts a low barrier (see Fig. 11) [105], which is consistent with the relaxation hypothesis.

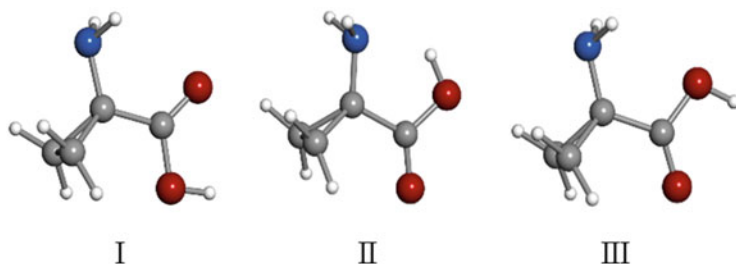


Fig. 12 Observed conformers for 1-aminocyclopropane carboxylic acid (Ac3c). (From [112])

The hypotheses of the III \rightarrow I interconversion was confirmed by the study of the 1-aminocyclopropanecarboxylic acid (Ac3c) (R=cyclopropyl, m.p. = 231°C) [112], a natural α -amino acid in which the rotation of the $-\text{COOH}$ group is hindered by the π -electron-donating capacity of the cyclopropane ring, able to conjugate with the COOH group [113, 114]. The LA-MB-FTMW investigation of Ac3c [30, 111] reveals the presence of the conformers I, II, and III of Fig. 12 with relative abundances I > III > II. The observation of the III conformer in Ac3c is highly remarkable but not completely unexpected according to the above considerations. Relaxation from the III to the I form involves rotation around the $\text{C}_\alpha\text{-C}(\text{O})$ bond and this conformational change should be highly restricted in Ac3c because of conjugation between the cyclopropane ring and the carboxylic acid. A predicted barrier of $2,000\text{ cm}^{-1}$ to interconversion from I and III conformers is high enough [108–110] to preclude conformational relaxation III \rightarrow I.

4.2.3 Valine, Isoleucine and Leucine

Valine (R=isopropyl, m.p. = 295–300°C) [115], isoleucine (R=*sec*-butyl, m.p. = 288°C) [116], and leucine (R=iso-butyl, m.p. > 300°C) [117] complete the study of aminoacids with aliphatic non-polar side chains.

Because of the hydrophobic character of the side chain, these amino acids are usually involved in protein or enzyme construction but rarely in protein function [86–88]. The main issues in the study of such aliphatic α -amino acids are the increase of conformational possibilities from the multiple configurations associated to torsion about single bond in the lateral chain. Two conformers of types I and II shown in Fig. 13 were ultimately detected in the supersonic jet for valine, isoleucine, and leucine, and conclusively identified through comparison of the experimental rotational and ^{14}N nuclear quadrupole coupling constants with the predicted values *ab initio*, as described in Sect. 3.

In the three amino acids, the conformational preferences for the intramolecular interactions reproduced those of glycine and alanine, with the amino-to-carbonyl interaction (I) most stable. Interestingly, the two conformers observed for each amino acid present the same arrangement of the side chain (valine Ia and IIa, isoleucine Ia₁

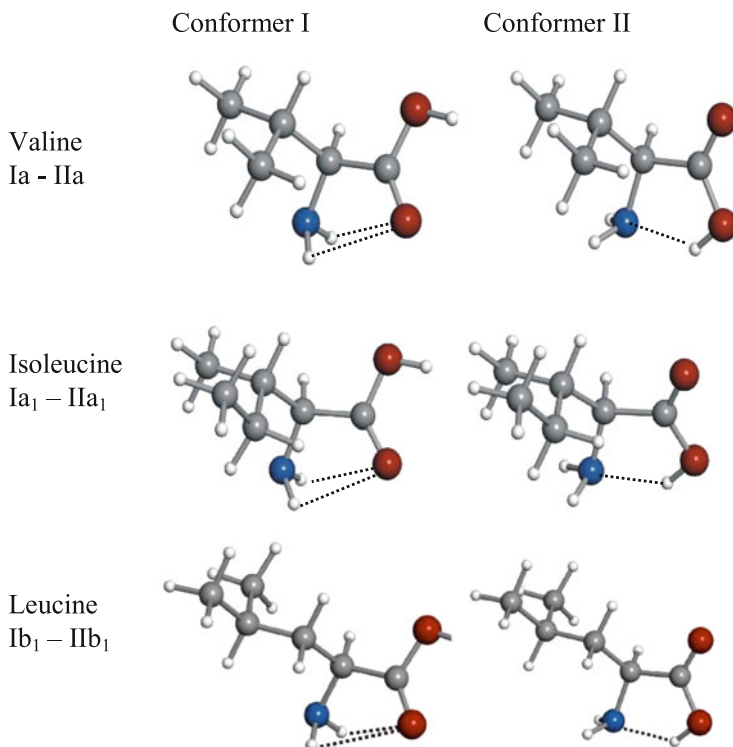


Fig. 13 The observed conformers for valine, leucine and isoleucine. (From [115–117])

and IIa₁, and leucine Ib₁ and IIb₁). This fact seems to indicate that the non-polar side chains do not interact significantly with the polar groups of the amino acid.

4.3 Imino Acids Proline and Hydroxyproline

Proline (m.p. = 228°C) is the only proteinogenic α -amino acid bearing a secondary amine in which the alpha carbon and the nitrogen atom are bound together to form a five-membered pyrrolidine ring. Proline has a unique role in protein formation because of its cyclic structure, acting at the end of α -helices or as structure disruptor in turns and loops [86–88]. Proline is also one of the major constituents of collagen, the most abundant protein in vertebrates, where it is accompanied by 4(*R*)-hydroxyproline (m.p. = 273°C), formed in a post-translational modification [118–124]. Collagen is a triple helix made of three super-coiled polyproline II-like chains with repetitive tripeptide sequences X-Y-Gly, where usually X=proline and Y=4(*R*)-hydroxyproline. However, the diastereoisomer 4(*S*)-hydroxyproline (m.p. = 243°C)

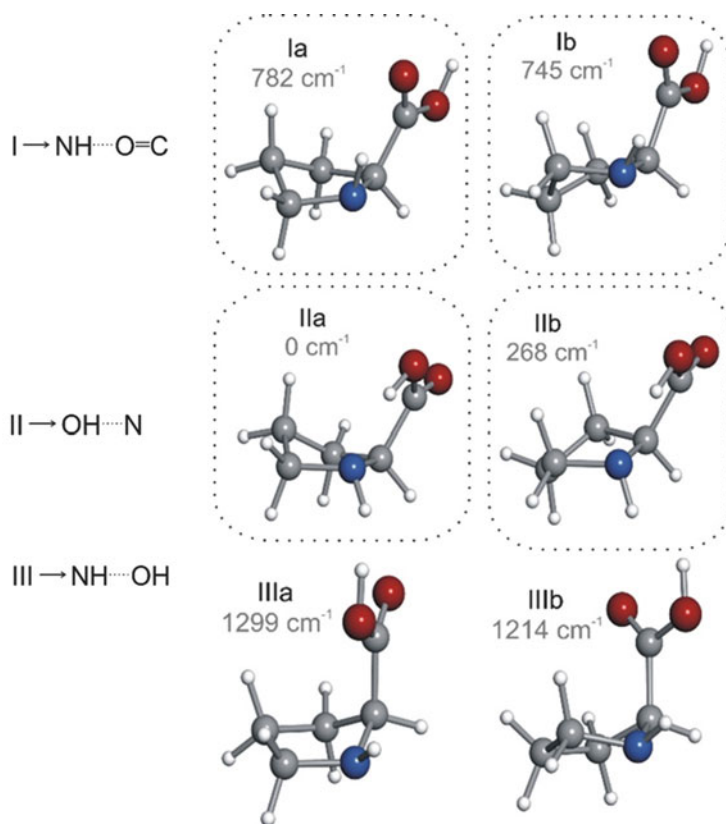
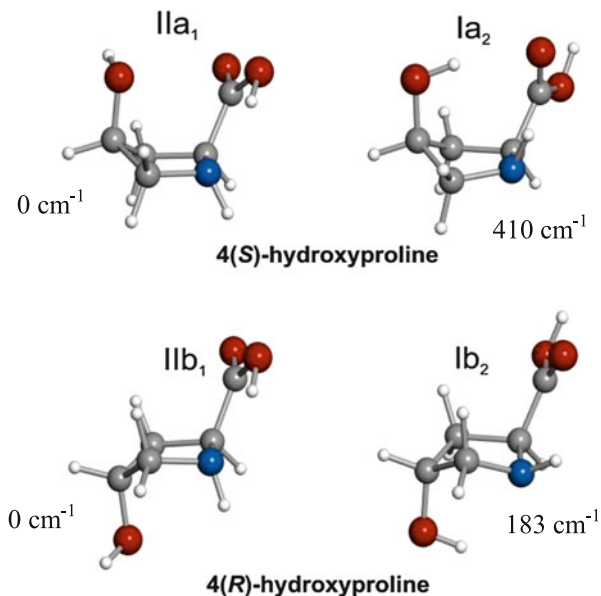


Fig. 14 Predicted low-energy conformers of proline and relative energies (MP2/6-311+G(d,p)) with respect to the global minimum. Conformers are labeled as *I*, *II*, or *III* depending on the hydrogen bond established between the amino group and carboxylic groups. Labels *a* or *b* indicate the configuration *endo* (*a*) or *exo* (*b*) adopted by the ring. The detected conformers are encircled. (From [59, 128])

inhibits the proper folding of the triple helix in either X or Y position [125, 126]. The fact that the diastereoisomer 4(*S*)-hydroxyproline produces a reverse destabilizing effect in the collagen triple helix makes a comparison of the three molecules interesting.

The low energy conformers of proline [127] are shown in Fig. 14 where they are labeled according to the hydrogen bond type I, II or III and the labels *a* or *b* to denote the C_{γ} -*endo* or C_{γ} -*exo* bent ring configurations, respectively. The first LA-MB-FTMW study on the rotational spectrum of an amino acid carried out in our laboratory was that of proline [59]. In this first study, only two conformers, *IIa* and *IIb*, were detected. Conformer *IIa* exhibits a *trans*-COOH arrangement, with the pyrrolidine ring adopting a C_{γ} -*endo* bent configuration. The distance between the nitrogen atom and the carboxylic hydrogen supports the existence of an N...H-O hydrogen bond, similar to those found in configuration II of aliphatic α -amino

Fig. 15 The observed conformers of 4(*S*)-hydroxyproline and 4(*R*)-hydroxyproline (From [129])



acids. For conformer IIb, analysis of the rotational data yielded a similar N \cdots H–O structure, but differing in a C_{γ} -*exo* ring. A later investigation taking advantage of the improvements in the laser ablation experiment [59] allowed the identification of two new conformers Ia and Ib [128]. These two new conformers are stabilized by interactions that link the hydrogen atom on the imino group to the oxygen atom of the carboxylic group (N–H \cdots O=C) and differ in the pyrrolidine ring configurations *endo*-like (a) and *exo*-like (b). From the intensities of the lines, the population of the observed conformers was found to follow the order IIa > IIb > Ia ~ Ib, with the conformer IIa as the global minimum in good agreement with ab initio calculations. For the most abundant conformer, IIa, all five monosubstituted ^{13}C isotopomers could be observed in their natural abundance. This isotopic information, together with the data from enriched samples of ^{15}N and two deuterated species, was used to derive its effective structure. The preference for the N \cdots H–O interaction in proline should be attributed to the geometrical constraints imposed by the pyrrolidine ring.

Two conformers were detected for each diastereoisomer of 4-hydroxyproline [129] (Fig. 15), which were labeled according to their intramolecular hydrogen bond pattern (II \equiv N \cdots H–O; I \equiv N–H \cdots O=C), ring puckering (a \equiv C_{γ} -*endo*; b \equiv C_{γ} -*exo*), and hydroxyl orientation (using arbitrary subscripts of the conformer type). In both 4-hydroxyprolines, the most stable conformer was found to exhibit an N \cdots H–O hydrogen bond (4(*R*): IIb₁; 4(*S*): IIa₁), as occurred for proline. However, unlike proline, the second stable conformers (4(*R*): Ib₂; 4(*S*): Ia₂) displayed an N–H \cdots O=C hydrogen bond. Significantly, the two forms detected in each case had the same ring puckering orientation (C_{γ} -*exo* for 4(*R*) and C_{γ} -*endo* in 4(*S*)). The propensity of 4(*R*)-hydroxyproline to give C_{γ} -*exo* puckerings would represent a form of preorganization of the collagen helix. Inspection of the most stable

structures gives information on the intramolecular interactions causing this effect. Apart from easily recognizable hydrogen bond interactions, the arrangement of the 4-hydroxyl group in the most stable conformer IIa₁ of 4(*S*)-hydroxyproline suggests an $n-\pi^*$ interaction produced by a hyperconjugative delocalization of a non-bonding electron pair (n) of the oxygen atom in the hydroxyl group into the π^* orbital at the carboxylic group carbon. This interaction is reminiscent of the preferred Bürgi–Dunitz trajectory for a nucleophilic addition to a carbonyl group [130, 131] because the predicted $O\cdots C=O$ distance of 2.97 Å and angle $\angle O\cdots C=O < 108.5^\circ$ are optimal for this approach. In this context, the possibility of further stereoelectronic contributions, as the so-called *gauche* effect [132, 133], were considered.

4.4 Proteinogenic Amino Acids with Polar Side Chain

The presence of polar functional groups in the side chains of α -amino acids is expected to increase dramatically the number of low-energy conformers. The functional group can establish additional interactions, which do not occur in other α -amino acids previously studied, and thus may affect the conformational preferences giving rise to a rich conformational space. The conformational behavior of serine [60], cysteine [134], threonine [61], aspartic acid [135], glutamic acid [67], and asparagine [136] have been revealed using LA-MB-FTMW spectroscopy. A summary of the results is collected in Table 3. The analysis of threonine

Table 3 The polar side chain (R) of α -amino acids with the number of observed conformers in each case

R	Amino acid	Observed conformers
–CH ₂ –OH	Serine	7
–CH ₂ –SH	Cysteine	6
–CH(OH)–CH ₃	Threonine	7
–CH ₂ –COOH	Aspartic acid	6
–CH ₂ –CH ₂ –COOH	Glutamic acid	5
–CH ₂ –CONH ₂	Asparagine	(1)

Source: From [136]

was described in detail in Sect. 3. The results for the other are presented in the next pages.

4.4.1 Serine, Cysteine, and Threonine

Serine [60] [$R=(CH_2OH)$, m.p. = 240°C] is the simplest proteinogenic amino acid with a polar group ($R=CH_2OH$) in its side chain. The hydroxyl group can interact as proton donor to the amino or carboxyl groups, or as proton acceptor through the non-bonding electron pair at its oxygen atom. Ab initio calculations predicted a significant number of conformers within 1,000 cm^{-1} (see Fig. 16). Up to seven rotamers were observed in the rotational spectrum following the procedure described in Sect. 3. The experiment also provided information on the relative stability of the serine conformers which follow the order $Ia > I Ib > I' b > IIc > III_{\beta} b \approx III_{\beta} c \approx IIa$, Ia, with an $N-H \cdots O=C$ hydrogen bond, being the global minimum. All observed conformers are stabilized by a network of

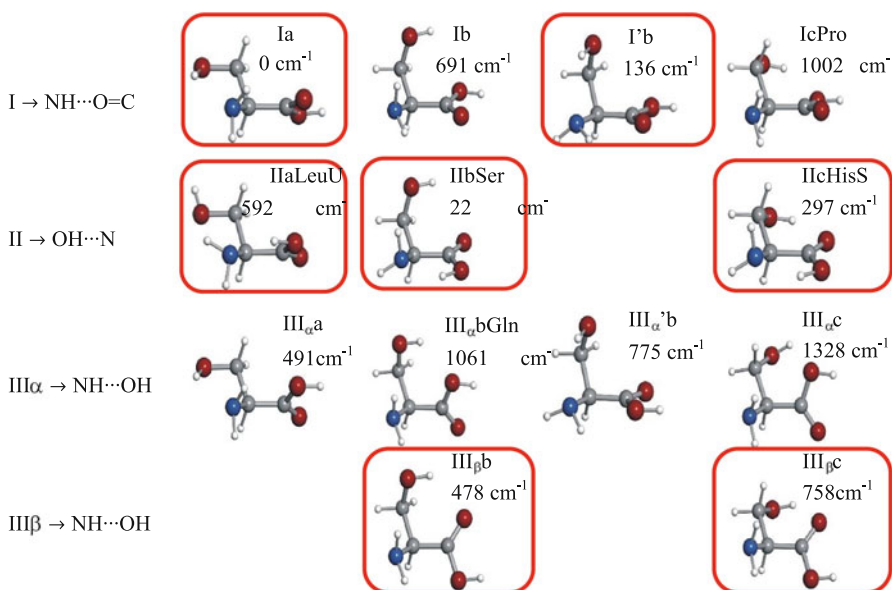


Fig. 16 The predicted low energy conformers of serine and relative energies (MP4/6-311++G(d, p)) with respect to the global minimum (in cm^{-1}). *I*, *II*, or *III* refer to the intramolecular hydrogen bonds between the $-NH_2$ and $-COOH$ groups. *a* corresponds to a +*sc* configuration for the $-OH$ and $-NH_2$ groups viewed along the $C_{\beta}-C_{\alpha}$ bond according to IUPAC terminology, *b* corresponds to *-sc* and *c* to *ap* configurations. The prime label indicates a down orientation of the $-NH_2$ along with an $O-H \cdots N$ intramolecular interaction with the side chain. In conformers *III* an α or β subscript indicates the H atom of the amino group bonded to the hydroxyl or carboxylic groups. The detected conformers are encircled. (From [60])

hydrogen bonds established between the different polar groups. Conformers Ia and I'b bear, in addition to the $\text{N-H} \cdots \text{O}=\text{C}$ amino acid backbone interaction, an $\text{O-H} \cdots \text{N}$ bond between the hydroxyl group in the side chain and the lone pair of the nitrogen atom. The side chain hydroxyl group interacts in a different way in each of the type II conformers. In IIa, the $-\text{OH}$ acts as a proton acceptor of one of the amino hydrogen atoms. In conformer IIc, the $-\text{OH}$ interacts as a proton donor to the carbonyl oxygen of the $-\text{COOH}$ group, and in conformer IIb the $-\text{OH}$ acts both as proton donor, to the $-\text{COOH}$ group, and as proton acceptor of one of the $-\text{NH}_2$ group hydrogen atoms. The amphoteric character of the hydroxyl group is also manifested in conformer III $_{\beta}$ b, where it interacts with both amino and carboxylic groups. In conformer III $_{\beta}$ c, the hydroxyl group only interacts with the carbonyl oxygen of the carboxylic group.

The observation of III $_{\beta}$ b and III $_{\beta}$ c conformers in serine is in direct contrast with the behavior observed in non-polar side chain amino acids. As explained above, in those cases the non-observation of type III conformers is attributed to the collisional relaxation from III to the most stable I conformers. For serine, conformers III $_{\beta}$ b and III $_{\beta}$ c are predicted to be more stable than the corresponding Ib and Ic forms. A low barrier of $\sim 140 \text{ cm}^{-1}$ was predicted for the interconversion from conformer Ib to conformer III $_{\beta}$ b. A similar energy profile was found for the interconversion from Ic to III $_{\beta}$ c conformer. The microwave work on serine [60] concluded that the $\text{I} > \text{III}$ preference of α -amino acids with non-polar side chains no longer holds for the "b" and "c" side chain arrangements because of the cumulative effects of the intramolecular hydrogen bonds established between all polar groups. The non-detection of conformer III $_{\alpha}$ a or conformer III $_{\alpha}$ b can be explained by a relaxation to the corresponding more stable type I conformers (see Fig. 11). In those cases, the side-chain hydroxyl group can only establish an $\text{O-H} \cdots \text{N}$ hydrogen bond with the amino group. The difference in stability is completely attributable to the interaction between the $-\text{NH}_2$ and $-\text{COOH}$ groups, and the $\text{I} > \text{III}$ preference of α -amino acids with non-polar side chains is maintained.

Similar behavior to that of serine was found for cysteine [134], [$\text{R}=\text{CH}_2\text{SH}$, m.p. = 300°C], a natural amino acid which bears an $-\text{SH}$ group in the side chain instead of the $-\text{OH}$ group. The microwave spectra reveals six rotamers identified as conformers Ia, IIa, Ib, IIb, III $_{\beta}$ b, and III $_{\beta}$ c. In cysteine, conformer Ib has been observed instead of I'b observed in serine, while conformed IIc observed in serine has not been observed for cysteine. The main differences between both molecules come from the marked less acidic character of the $-\text{SH}$ group compared to the $-\text{OH}$ group.

Threonine ($\text{R}=-\text{CH}(\text{CH}_3)\text{OH}$, m.p. = 256°C) [61] is related to serine as it only differs on the C_γ methyl group. Figure 7 shows the seven low energy conformers detected for threonine. While form III $_{\beta}$ b observed in serine [60] and cysteine [134] was also observed in threonine, conformer III $_{\beta}$ c was not detected. This can be attributed to the raising of the energy of this conformer caused by steric interaction of the C_γ methyl group with the $-\text{COOH}$ group. Conformer III $_{\alpha}$ a has been observed, in contrast to serine and cysteine, where this conformer relaxes in the supersonic

expansion to the Ia form. The only explanation in this case is that the steric interactions between the C_γ methyl group and the $-\text{COOH}$ group would increase the barrier of interconversion between those forms, predicted to be $\sim 1,000\text{ cm}^{-1}$ in threonine [61] which cannot be surmounted via collisions with the carrier gas.

4.4.2 Aspartic and Glutamic

Aspartic acid ($\text{R}=-\text{CH}_2-\text{COOH}$, m.p. $>300^\circ\text{C}$) is a natural amino acid with a carboxyl group in the side chain. Ab initio searches in the potential energy surface predict up to 14 low energy conformers to lie within 800 cm^{-1} [135]. Investigation of the microwave spectrum of laser ablated aspartic acid yields the identification of six conformers (see Fig. 17). The relative population yields follow the order $\text{Ia-I} > \text{Ib-I} \approx \text{IIa-I} > \text{Ia-II} \approx \text{IIIb-I} > \text{III}_\beta\text{b-I}$. Conformer Ia-I is the most abundant conformer in the molecular beam, in poor agreement with the ab initio calculations which predict Ia-I to be ca. 300 cm^{-1} above the global minimum. This fact can be attributed to relaxation processes between conformers [108–110, 135].

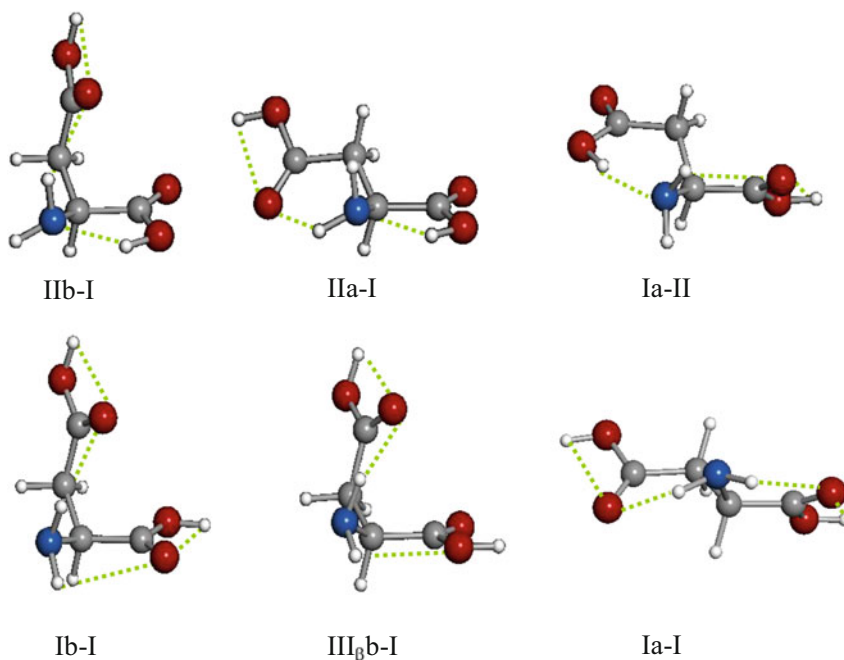


Fig. 17 The observed conformers of aspartic acid. Labeling is the same as used in serine, cysteine or threonine. An additional label indicating the hydrogen bond type (*I*, *II*, or *III*) between the $-\text{NH}_2$ and the $\beta\text{-COOH}$ groups is used to avoid ambiguities. (From [135])

It was confirmed again that the presence of a nitrogen nucleus makes the ^{14}N quadrupole coupling constants a unique tool to establish the nature of the intramolecular hydrogen bonding. All the conformers detected are stabilized through a network of intramolecular hydrogen bonds, in which all functional groups participate (see Fig. 19). The six conformers observed belong to the “a” or “b” families, the β -COOH group being synclinal to the amino group, which allows these two functional groups to interact. This indicates that the β -COOH prefers to interact with the amino group rather than with the COOH group in α . The $-\text{NH}_2$ group acts as a bridge connecting the $-\text{COOH}$ groups in the hydrogen bond network. Furthermore, all observed conformers except one are of type I, indicating that the β -COOH acts essentially as proton acceptor rather than as proton donor.

Aspartic acid follows the trends observed by all but one of the aliphatic amino acids studied: type I conformers are more populated than type II conformers. The most populated form of aspartic acid is conformer Ia-I, which is stabilized by two $\text{N}-\text{H}\cdots\text{O}$ hydrogen bonds established between the amino group and each of the carboxylic groups, both of them in a *cis* configuration. The amino group disposition in Ia-I gives rise to considerably shorter $\text{N}-\text{H}\cdots\text{O}$ hydrogen bonds than in other conformers, which is usually associated with greater hydrogen bond strength and could explain the stability of Ia-I. In the case of aspartic acid, the observation of the conformer $\text{III}_{\beta}\text{b-I}$ can be ascribed to the high potential barrier (around $1,000\text{ cm}^{-1}$), as predicted from ab initio calculations [135] for relaxation to the Ib-I form.

Glutamic acid ($\text{R}=-\text{CH}_2-\text{CH}_2-\text{COOH}$, m.p. = 205°C) is a clear example of the great torsional flexibility originating from multiple torsional degrees of freedom because of its large side chain. As in aspartic acid, the polar group in the side chain is a $-\text{COOH}$ group. Figure 18 shows the five observed conformers [67] for glutamic acid which show intramolecular hydrogen bonds of type I or II between the carboxylic group in α and the amino group and have the terminal COOH group in

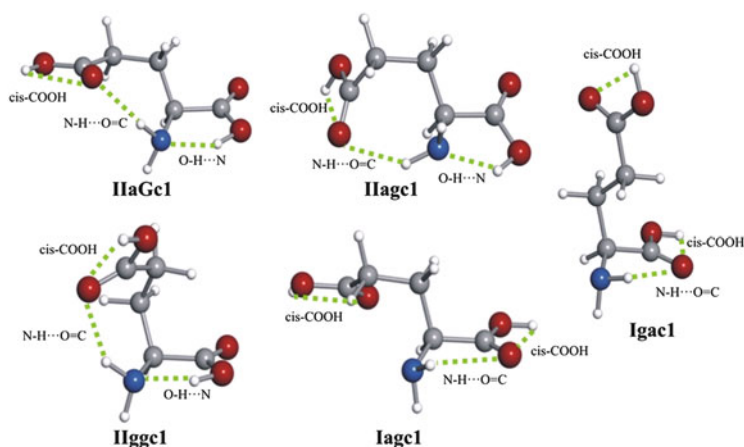


Fig. 18 The five observed conformers of glutamic acid showing their intramolecular interactions. (From [67])

a *cis* configuration. No type III conformers have been observed in this case. Conformer Iagc1 presents a type I hydrogen bond, i.e., an N–H···O=C hydrogen bond between the –NH₂ and the α-COOH which adopts a *cis* arrangement. Glutamic acid thus follows the behavior displayed by the overwhelming majority of the aliphatic α-amino acids studied so far with a type I conformer as global minimum. Conformer Igac1 is also stabilized by an N–H···O=C hydrogen bond between the carboxylic group in alpha (in a *cis* disposition) and the amino group. Similar to the most abundant conformer Iagc1, it displays an extended backbone configuration where the acid group in the side chain does not establish any additional interactions with the other polar groups in the amino acid. The other conformers identified, IIggc1, IIagc1, and IIaGc1, present a similar intramolecular hydrogen bond network. As for aspartic acid, the interactions of the terminal –COOH group with the amino acid backbone are dominated by the type I interaction NH···O in which the terminal –COOH acts as proton acceptor.

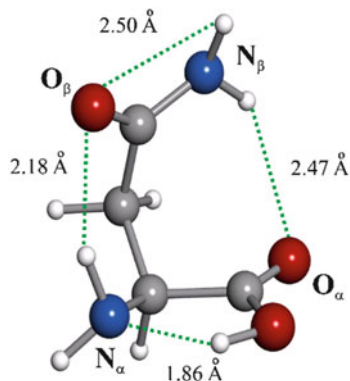
The γ-COOH group of Iagc1 and Igac1 does not establish hydrogen bonds with any of the other functional groups. This is in sharp contrast with what was observed for other α-amino acids with polar side chains, described above, where all conformers flaunted interactions between all functional groups. For example, the most abundant conformer of aspartic acid [135], the polar amino acid more closely related to glutamic acid, was stabilized by two N–H···O=C hydrogen bonds between the amino group and the carboxylic groups in alpha and beta. In comparison with aspartic acid, glutamic acid has a longer side chain which can adopt a larger number of dispositions. For some of these dispositions the γ-COOH group is unable to establish any interactions with the other polar groups in the amino acid. These conformations, with less intramolecular interactions, are entropically favored and Iagc1 is the most populated conformer in the supersonic jet.

Type III conformers have only been observed in those α-amino acids where rotation of the α-COOH group is hindered or in those for which the I/III relative energies are reversed. Glutamic acid has a polar side chain –CH₂–CH₂–COOH with the same functional group as that of aspartic acid (–CH₂–COOH) but possessing just one more methylene group. However, this apparently small increase in chain length produces a strong impact on the conformational preferences.

4.4.3 Asparagine: Conformational Locking

Asparagine (R=–CH₂–CONH₂, m.p. = 236°C) is similar to aspartic acid but with an amide group, –CONH₂, in the lateral chain instead of a –COOH group. In this case, only one structure has been identified (see Fig. 19) in the jet cooled rotational spectrum [136]. This is in sharp contrast to the multiconformational behavior observed in other proteinogenic amino acids with polar side chains. This conformational locking to a single conformer IIa is caused by a network of three cooperative intramolecular hydrogen bonds, N_α···H–O_α, N_α–H···O_β, and N_β–H···O_α, forming the intramolecular sequence shown in Fig. 19. Such arrangement of functional groups into hydrogen bond networks leads to the phenomenon of

Fig. 19 Observed conformer, atom labeling for asparagine and hydrogen bond distances taken from the ab initio structure (From [136])



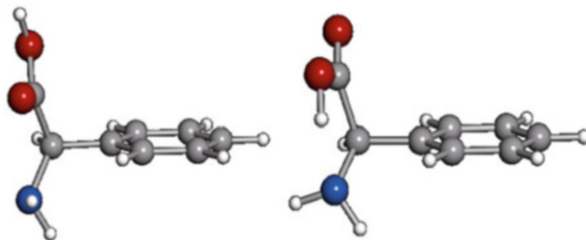
cooperativity [102–104] under which the strength of individual hydrogen bond interactions is notably enhanced. In addition, the simultaneous formation of $N_{\beta} \cdots H-O_{\alpha}$ and $N_{\alpha}-H \cdots O_{\beta}$ may contribute to a further stabilization of the amide group by favoring the resonance form which confers a double bond character to the C–N amide bond.

To conclude, in serine ($R=-CH_2-OH$) [60], threonine ($R=CH(OH)-CH_3$) [61], cysteine ($R=-CH_2-SH$) [134], and aspartic acid ($R=-CH_2-COOH$) [135], the side chain has been found to change the conformational preferences with respect to aliphatic α -amino acids without polar side chains. The number of possible conformations increased and conformers with type III hydrogen bonds between the α -COOH and $-NH_2$ groups were observed and sometimes found to be more stable than their analogues with type I hydrogen bonds. The longer side chains of glutamic acid ($R=-CH_2-CH_2-COOH$) [67] confer larger flexibility and multiply the conformational possibilities, but also changes the influence of the interactions involving the γ -COOH group and the other polar groups in the amino acid. The increased length of the side chain in glutamic acid makes entropic contributions arising from the presence of intramolecular interactions more significant than in other polar amino acids. This is reflected in the fact that the most populated conformer of glutamic acid is an extended conformer where the side chain polar group does not interact with the $-NH_2$ or α -COOH groups. The rich conformational behavior of these amino acids contrasts with that of asparagine [136], which collapses to only one conformer, highly stabilized by a network of intramolecular hydrogen bonds, involving all polar groups in the molecule.

4.5 Amino Acids with Aromatic Side Chain

As already explained in the introduction of this review, the pioneering studies of the proteinogenic amino acids in the gas phase were carried out using laser spectroscopy techniques. However, since these techniques can be applied only to molecules

Fig. 20 The two observed conformers, I (*left*) and II (*right*), of the unnatural amino acid phenylglycine in the gas phase (From [152])



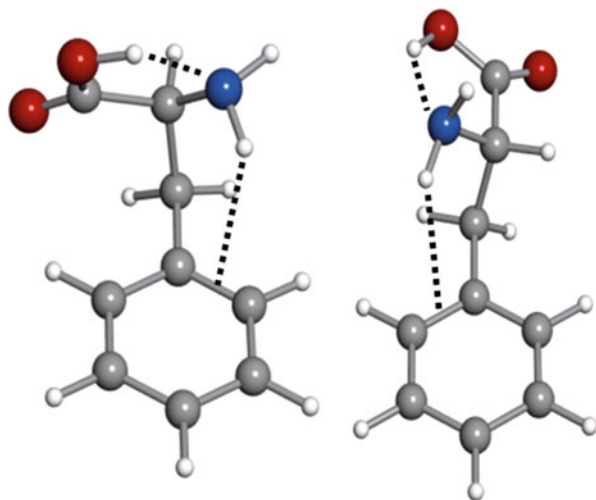
bearing chromophore groups, their application to proteinogenic amino acids has been limited to phenylalanine (R=benzyl) [11, 137–147], tyrosine (R=*p*-hydroxybenzyl) [3, 11, 138, 148], and tryptophan (R=3-methylene indolyl) [149–151] which have been the subject of a large number of investigations. While the studies of non-aromatic amino acids by microwave spectroscopy using laser ablation techniques has been successful, the same techniques applied to the aromatic ones have been shown to have problems related to photofragmentation/ionization of these molecules taking place during the laser ablation process.

4.5.1 Phenylglycine and Phenylalanine

Phenylglycine (R=–C₆H₅, m.p. = 290°C) is the simplest aromatic α -amino acid, with the phenyl group directly attached to the α -carbon. It was the first aromatic amino acid studied by LA-MB-FTMW and thus provided a suitable test for the application of laser ablation chromophoric systems [152]. The analysis of the rotational spectrum leads to the identification of two conformers shown in Fig. 20. The most abundant conformer Ia exhibits a hydrogen bond interaction N–H \cdots O=C and a *cis*-COOH arrangement reminiscent of glycine I. The second amino group hydrogen atom points to the phenyl ring, indicative of an N–H \cdots π hydrogen bond interaction. The conformer II presents an N \cdots H–O interaction with a *trans*-COOH configuration. The population ratio derived for the phenylglycine forms in the supersonic jet is $N_{\text{Ia}}/N_{\text{II}} \sim 4$, which demonstrates the predominance of the type I conformer, as occurs in all aliphatic α -amino acids. The amino acid skeleton of phenylglycine reproduces the primary conformational preferences of aliphatic amino acids.

The conformational landscape of phenylalanine (R=CH₂–C₆H₅, m.p. = 270–275°C) has been widely investigated [11, 137–147]. Six conformational species were identified using laser induced fluorescence LIF, hole burning UV–UV, and ion dip IR–UV spectroscopy coupled with *ab initio* calculations [11, 137–141]. Lee et al. [141] carried out a definitive identification of the conformers of phenylalanine, based upon comparisons between the partially resolved ultraviolet band contours and that simulated by *ab initio* computations. The study of the rotational spectrum of phenylalanine by LA-MB-FTMW [153] showed rather weak spectra of only two conformers, IIa and IIb (see Fig. 21). Both conformers exhibit a *trans* configuration in the COOH group, being stabilized by O–H \cdots N and N–H \cdots π

Fig. 21 The two observed conformers, IIa (*left*) and IIb (*right*), of phenylalanine (From [153])



intramolecular hydrogen bonds. The low intensity of the observed spectra and the non-observation of the conformers, detected with other techniques, can be caused by ionization and photofragmentation processes during laser ablation taking place with different rates for the different conformers. Kim and co-workers [145–147], from their measurements of the ionization energies (IE) of the low-energy conformers of phenylalanine, affirm that the observed conformers IIa and IIb have higher IEs than the other conformers.

4.6 Non-proteinogenic Amino Acids

Rotational studies extended to non-coded α -amino acids are also of biochemical relevance [154]. The effects on the conformational behavior of enlarging the amino acid backbone chain have been analyzed on β -alanine [155, 156] and γ -amino butyric acid (GABA) [157]. Both are neurotransmitters which bind to the same sites as glycine [158–160]. These are also the simplest β -amino and γ -amino acids and so are the natural starting point to analyze the conformational panorama of this type of amino acids. Other studies include α -aminobutyric acid [161], the *N*-alkylated species sarcosine [162], *N,N*-dimethylglycine [163], and taurine [164]. Brief results on β -alanine and GABA are presented.

4.6.1 β -Alanine

The rotational spectrum of β -alanine ($\text{NH}_2\text{-CH}_2\text{-CH}_2\text{-COOH}$, m.p. = 202°C) was studied by Godfrey et al. [155] using heating methods of vaporization.

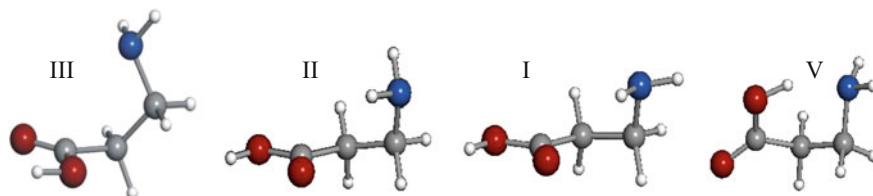


Fig. 22 Observed conformers of β -alanine. (From [156])

Two conformers (I and V shown in Fig. 22) were identified, stabilized by $\text{N-H}\cdots\text{O}=\text{C}$ and $\text{N}\cdots\text{H}-\text{O}$ intramolecular bonds. An LA-MB-FTMW study [156] led to the characterization of two new conformers, II and III. Full resolution of the ^{14}N quadrupole coupling structure was invaluable to distinguish between the most abundant conformers I and II, which exhibit an $\text{N-H}\cdots\text{O}=\text{C}$ hydrogen bond with different orientation of the amino group. Conformer III is not stabilized by a hydrogen bond, but by an $n\rightarrow\pi^*$ interaction between the electron lone pair at the nitrogen atom and the π^* orbital of the $-\text{COOH}$ carbonyl group which produces electronic delocalization by hyperconjugation to the π^* orbital [130, 131]. Such interaction is described in more detail below.

4.6.2 γ -Aminobutyric Acid (GABA)

In γ -aminobutyric acid (GABA) ($\text{NH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2\text{-COOH}$, m.p. = 204°C) the five hindered rotations around the single bonds generate a plethora of conformational species. An overall picture of the conformational landscape obtained from theoretical predictions [157] confirms the richness of GABA: up to 30 feasible conformers shown in Fig. 23 were localized on the ab initio potential energy surface with relative energies below 900 cm^{-1} . Thorough analysis of the rotational spectra finally led to the assignment of nine different rotamers of GABA [157] encircled in Fig. 23. A close look at the detected conformers indicates that in *GG2*, *aG1*, *aal*, and *gal*, the two polar groups are far apart and no intramolecular interactions are apparent, apart from a stabilizing *cis*- COOH functional group interaction. In folded configurations, non-covalent interactions can be established between the two polar groups. Conformers *gG2* and *GG1* are stabilized by type II, $\text{O-H}\cdots\text{N}$, and conformer *GG3* by type I $\text{N-H}\cdots\text{O}=\text{C}$ intramolecular hydrogen bonds similar to those observed in non-polar aliphatic α -amino. As occur for the III form of β -alanine, conformers *gG1* and *gG3* of GABA show an arrangement of the amino and carboxyl groups which resembles the Bürgi–Dunitz trajectory [130, 131] which describes the most favorable approach of a nucleophile nitrogen to a carbonyl group carbon in an addition reaction. This is a signature of the existence of an $n\rightarrow\pi^*$ interaction arising from the hyperconjugative delocalization of the non-bonding

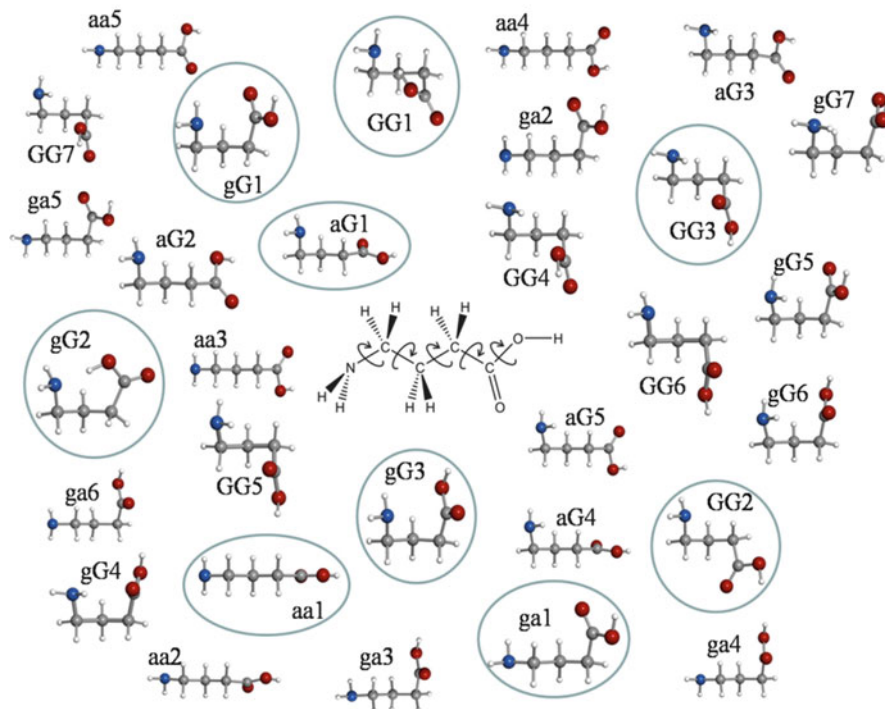


Fig. 23 Predicted low energy conformers of GABA. The nine observed conformers are encircled. (From [157])

electron pair of the nitrogen atom to the π^* orbital at the carbonyl group (see Fig. 24). In the Bürgi–Dunitz trajectory, the approach path of the nucleophile ($N:$) is described to lie in the plane bisecting the $\angle R-C-R'$ angle with an angle α of about $105^\circ \pm 5^\circ$ for sort N to C distances [130, 131]. The ab initio structures confirm that gG1 and gG3 conformers of GABA show an optimal geometrical arrangement for this interaction.

The relative populations of the GABA conformers $GG2 > aG1 > gG1 > aal > gal$ can be taken as proof of the coexistence of GABA conformers having intramolecular interactions (folded) with those free of them (extended). Conformers of GABA free of intermolecular interactions are the most abundant. Intramolecular interactions contribute to decrease entropy and to increase the Gibbs energy, thus diminishing number density. The works on β -alanine [156] and GABA [157] show that an increment in the number of methylene groups between the amino and carboxylic groups gives rise to the existence of stabilizing interactions, such as $n \rightarrow \pi^*$, different from the hydrogen bond. In addition, the presence of low energy extended conformers free of intramolecular interactions becomes significant.

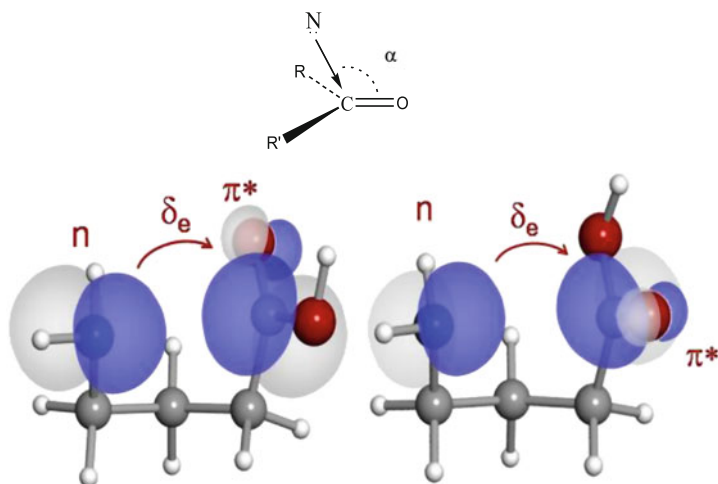


Fig. 24 Scheme of the Bürgi-Dunitz trajectory of addition of a nucleophile N to a carbonyl carbon atom and representation of the $n \rightarrow \pi^*$ interaction in conformers gG1(a) and gG3(c) of GABA. (From [157])

5 Nitrogen Bases

Much effort has been devoted to the identification of preferred tautomers of nucleobases since the structure of nucleic acid and its base pairs were first reported [165]. The best experimental approach to address the structural preferences of nucleobases is to place them under isolation conditions in the gas phase, cooled in a supersonic expansion. Under these conditions, the various tautomers/conformers can coexist and are not affected by the bulk effects of their native environments, which normally mask their intrinsic molecular properties. The main restriction to the gas-phase study of these building blocks is the difficulty in their vaporization owing to their high melting points (ranging from 316°C for guanine to 365°C for adenine) and associated low vapor pressures. The success of LA-MB-FTMW experiments to the study of coded amino acids prompted their application to nucleic acids uracil [166], thymine [167], guanine [168], and cytosine [169], as well as the monohydrates of uracil and thymine [170]. This technique gives a precise interpretation of the structure and relative energies of the different forms of nucleic acid bases to solve apparent discrepancies between the previous studies. The rotational constants, which are the main tool to identify the different forms of a biomolecule, had a minor role when trying to discern between the different tautomers of a nucleobase. It is the quadrupole coupling hyperfine structure caused by the presence of ^{14}N nuclei which constitute authentically the fingerprints of every tautomer. In the next sections, the results on nucleobases are discussed.

5.1 Uracil, Thymine, and Their Monohydrates

Uracil and thymine have similar structures and only differ in the presence of a methyl group in thymine (5-methyluracil). These can exist in various tautomeric forms which differ by the position of the hydrogen atoms, which may be bound to either nitrogen or oxygen atoms (keto-enolic equilibrium). The first observations of the rotational spectra of uracil [171] and thymine [172] were made using Stark modulation free-jet absorption millimeter-wave spectroscopy. The solid samples were vaporized under carefully controlled heating conditions to avoid decomposition. Only the diketo tautomer was observed in each case. The identification was based on the agreement between predicted and experimental rotational constants value. Uracil and thymine have been probed in the gas phase under high-resolution conditions using LA-MB-FTMW spectroscopy. Both nucleobases bear two ^{14}N atoms with nonzero quadrupole moments ($I=1$), which interact with the electric field gradient at the nucleus, resulting in a complicated hyperfine structure. As mentioned in Sect. 3, the experimental values of the quadrupole coupling constants provide an independent approach to identify the tautomeric forms.

In the case of thymine, the rotational spectrum was complicated not only by the hyperfine structure of two ^{14}N atoms but also by a further doubling arising from the coupling of the internal rotation of the methyl group to the overall rotation (see Fig. 25). The resulting structure was completely resolved and analysis of the spectrum yielded the quadrupole coupling constants for both ^{14}N nuclei and the barrier to internal rotation of the methyl group, $V_3 = 1.502(9)$ kcal/mol obtained from the A–E doublet (see Fig. 25).

For uracil, the intensity of the observed spectrum led to the observation of the spectra of the $^{15}\text{N}(1)\text{-}^{14}\text{N}(3)$ and $^{14}\text{N}(1)\text{-}^{15}\text{N}(3)$ isotopomers in natural abundance. In a subsequent step, a $^{15}\text{N}\text{-}^{15}\text{N}$ enriched sample was used to observe the

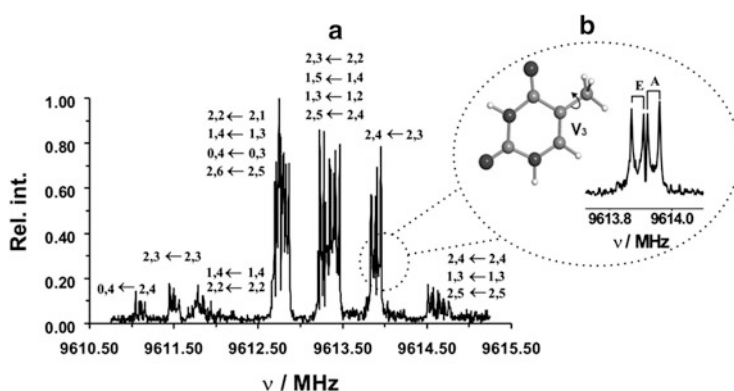


Fig. 25 (a) The $4_{1,4}\text{-}3_{0,3}$ rotational transition of thymine showing the ^{14}N quadrupole components labeled with the quantum numbers $I', F' \leftarrow I'', F''$. (b) Detail of the $I, F = 2, 4 \leftarrow 2, 3$ quadrupole component, which is split into four lines because of the internal rotation (A–E doublet) and Doppler effects (\square). (From [167])

spectra of all ^{13}C and ^{18}O monosubstituted isotopologues. The inertial defect ($\Delta_c = I_c - I_a - I_b = \sum_i m_i c_i^2$) measures the mass extension out of the ab inertial plane and is close to zero for planar molecules. The values of parameter for the observed isotopologues range from -0.129 to $-0.134 \text{ u}\text{\AA}^2$, allowing one to establish the planarity of uracyl. The substitution and equilibrium structures of this molecule [166] are shown in Table 4. Recently, the computational composite scheme [173] has been applied to the first study of the rotational spectrum of 2-thiouracil. The joint experimental–computational study allowed the determination of the accurate molecular structure and spectroscopic properties of this nucleobase.

The uracil–water and thymine–water complexes provide the simplest molecular models of the interactions between biologically important nitrogen bases and water. These monohydrates have been the subject of many theoretical studies [174 and references therein, 175, 176] and experimental studies [177–179]. In the microwave work [170] by LA-MBFTMW spectroscopy, only one conformer of each complex has been observed. Investigation of the structure of the adducts from the rotational constants of the different isotopologues shows that the observed conformers correspond to the most stable forms in which water closes a cycle with the nucleic acid bases forming $\text{N}-\text{H}_{\text{NB}} \cdots \text{O}_{\text{w}}$ and $\text{H}_{\text{w}} \cdots \text{O}=\text{C}_2$ hydrogen bond (see Fig. 26). Both adducts present similar hydrogen bond structures which are also comparable to those observed for related complexes, such as formamide–water [180–182] *N*-methylformamide–water [183], or 2-pyridone–water [184], as could be expected given the similar natures of the hydrogen bonds in all these systems.

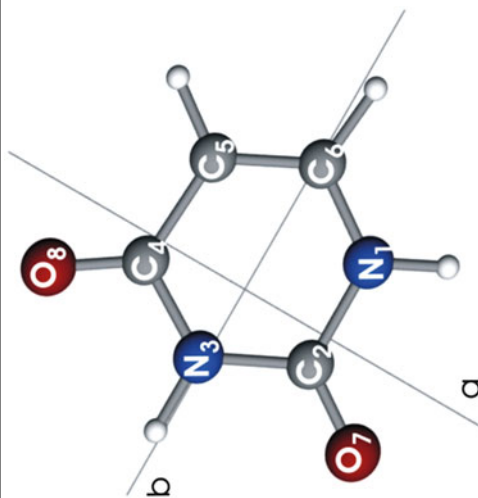
5.2 Guanine

Theoretical calculations predict the existence of four low-energy forms [185–187]: keto N7H, keto N9H, and enol N9H *cis* and *trans* (Fig. 27). The features observed from UV laser spectroscopy of guanine [188–193], IR techniques in He nanodroplets experiments [194], and electron diffraction study [195] led to some controversy. The four more stable forms were first identified by de Vries and co-workers in the R2PI spectrum [188, 191]. However, in successive works by Mons and co-workers [189, 192, 193] and Seefeld et al. [190] it was concluded that the spectrum is dominated by the less stable tautomers N7H enol and the two keto imine tautomers. This was attributed to the occurrence of a fast non-radiative relaxation of the excited states of the N7H keto, N9H keto, and N9H enol *trans* which prevents their observation in the R2PI spectrum. On the other hand, the fourth most stable form was identified by IR techniques in He nanodroplet experiments [194], while an electron diffraction study confirmed the existence of N9H keto form [195].

The difficulties in detecting conclusively the most stable tautomers of guanine showed the need for a gas-phase high sensitivity structural probe such as microwave spectroscopy insensitive to excited-state dynamics. However, guanine has the drawback of having five ^{14}N atoms in its structure so their quadrupole coupling

Table 4 Atom coordinates in the principal inertial axis system and substitution (r_s) structure for the diketo tautomer of uracil. Distances are given in Å and angles in degrees

	a	b	c ^a
N ₁	1.1884(10) ^b	-1.0115(12)	0
C ₂	1.1975(10)	0.375(3)	0
N ₃	-0.091(13)	0.9640(12)	0
C ₄	-1.2925(9)	0.295(4)	0
C ₅	-1.1992(10)	-1.1535(10)	0
C ₆	0.05(2)	-1.7396(7)	0
O ₇	2.2323(5)	1.0189(12)	0
O ₈	-2.3273(5)	0.9432(13)	0
$r(\text{N}_1-\text{C}_2)$	1.386(5) ^c	$\angle \text{N}_1\text{C}_6\text{C}_5$	122.3(6)
$r(\text{N}_3-\text{C}_4)$	1.38(2)	$\angle \text{C}_2\text{N}_1\text{C}_6$	123.0(11)
$r(\text{C}_4-\text{C}_5)$	1.451(4)	$\angle \text{C}_5\text{C}_4\text{N}_3$	115.4(16)
$r(\text{C}_5-\text{C}_6)$	1.379(4)	$\angle \text{C}_6\text{C}_5\text{C}_4$	118.8(12)
$r(\text{C}_6-\text{N}_1)$	1.352(14)	$\angle \text{N}_1\text{C}_2\text{O}_7$	122.3(8)
$r(\text{C}_2-\text{O}_7)$	1.219(4)	$\angle \text{C}_5\text{C}_4\text{O}_8$	118.8(7)
$r(\text{C}_4-\text{O}_8)$	1.22(2)		



Source: From [166]

^aAssumed value for all atoms, since uracil is a planar molecule lying on the ab inertial plane

^bDerived errors in parentheses in units of the last digit

^cDerived errors in parentheses in units of the last digit

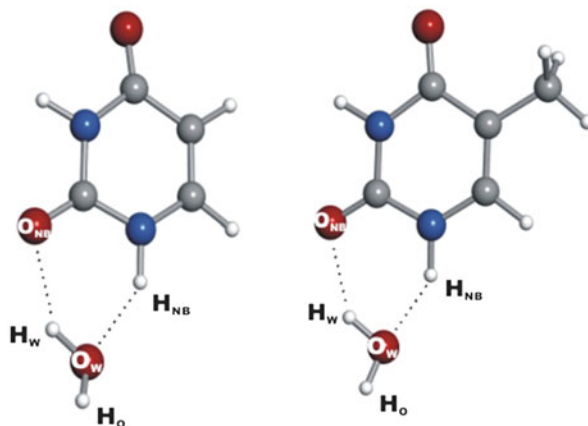


Fig. 26 The uracil-water (*left*) and thymine water (*right*) observed complexes. (From [170])

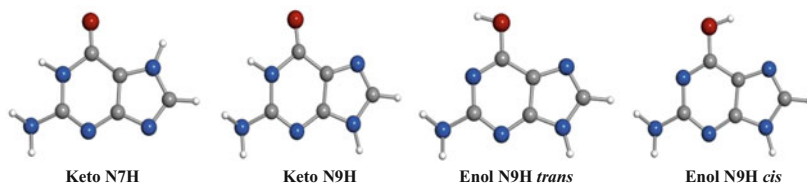


Fig. 27 The four observed conformers of guanine. (From [168])

hyperfine structure gives very complex patterns in the rotational spectra. With this background the LA-MB-FTMW rotational spectra was investigated and four different rotamers were recognized [168]. All observed transitions were split into many components, confirming that they belong to guanine. No attempt was made to assign the quadrupole hyperfine components, and the rotational frequencies were measured as the intensity-weighted mean of the line clusters. The values of the rotational constants and the electric dipole moment components were used to identify the observed rotamers as the four most stable forms of guanine in Fig. 27. The post-expansion abundances measured from relative intensity measurements point to a higher stability of the N9H and N7H keto forms. Finally, the values of the inertial defect, ranging from 0.48 to 0.68 $\text{u}\text{\AA}^2$, show that the tautomers of guanine are slightly non-planar.

5.3 Cytosine

The molecular system of cytosine (CY) is more complex than that of guanine. Figure 28a shows the five most stable species according to theoretical calculations

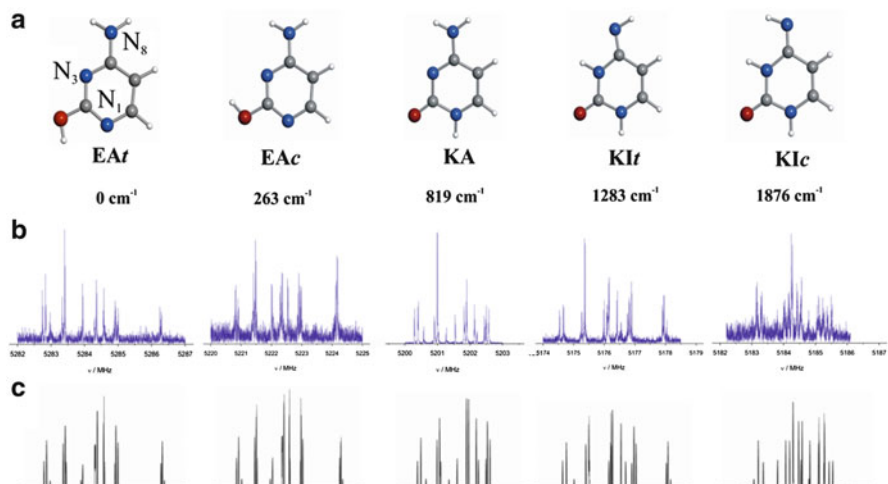


Fig. 28 (a) The five more stable species of cytosine: enol-amino *trans* (EAt), enol-amino *cis* (EAc), keto-amino (KA), keto-imino *trans* (KItr), keto-imino *cis* (KIc) given in order of stability according to MP2/6-311++G(d,p) ab initio calculations. (b) LA-MB-FTMW spectra for the $1_{1,1}-0_{0,0}$ rotational transition of the five species. (c) Theoretical simulation of the nuclear quadrupole hyperfine structure for the $1_{1,1}-0_{0,0}$ rotational transition. The differences among the various patterns act as fingerprints for tautomeric/conformational assignment. (From [169])

[169]: enol-amino *trans* (EAt), enol-amino *cis* (EAc), keto-amino (KA), keto-imino *trans* (KItr), and keto-imino *cis* (KIc). Different experiments have been conducted to reveal the tautomerism of cytosine. The infrared spectra in inert gas matrices were interpreted in terms of a mixture of KA and EA [196]. IR laser spectroscopy in helium nano-droplets [197] characterized the EAt, EAc, and KA species. Two features observed in the vibronic spectra were attributed to the KA and EA forms [198]. The electron diffraction pattern [199, 200] was interpreted in terms of a conformation mixture dominated by the EA forms. In very recent experiments on Ar-matrix, photoisomerization processes [201, 202] were interpreted in terms of the coexistence of various tautomers of cytosine. The free jet millimeter-wave absorption spectra [203] of the three detected species were tentatively assigned to KA, Eat, and KI forms. Five rotamers were observed in the high resolution rotational spectrum of cytosine investigated by LA-MB-FTMW spectroscopy [169]. Rotational transitions exhibit a very complex hyperfine structure caused by the presence of three ^{14}N nuclei. Analysis of this hyperfine structure yields the nuclear quadrupole coupling constants, which are extremely sensitive to the electronic distribution around the quadrupolar nuclei $^{14}\text{N}_1$, $^{14}\text{N}_3$, and $^{14}\text{N}_8$. Comparison between experimental and predicted spectroscopic constants leads to a conclusive identification of the detected rotamers.

Species EA_t and EA_c have very similar values for the three ¹⁴N nuclei, being the differences between the two conformers only caused by the different orientation of the hydroxyl group. In those conformers, the positive values of χ_{cc} for ¹⁴N₁ and ¹⁴N₃, in the range 1.0–1.8 MHz, indicate that these are pyridinic nitrogen atoms, while the negative values for ¹⁴N₈ indicate this is an amino nitrogen atom. A conclusive discrimination of both species was achieved by using the trend of the change on the rotational constants. In tautomer KA, the χ_{gg} ($g = a, b, c$) values associated with atoms ¹⁴N₃ and ¹⁴N₈ do not change very much with respect to the previous tautomers, but those for atom ¹⁴N₁ change radically, indicating that it is an imino N atom. In passing to conformers KI, the quadrupole coupling constants associated with atom ¹⁴N₃ change and the corresponding value of χ_{cc} reveals that it is now an imino N atom. Finally, χ_{cc} of ¹⁴N₈ of KI_t and KI_c are “chemically” different from all other χ_{cc} values and were finally identified as being part of a C=N–H group (see Table 5). ¹⁴N nuclear quadrupole patterns make it possible to obtain the spectral signatures for each individual tautomer in the complex sample and thus act as a sort of fingerprint (Fig. 28b, c). The relative intensity measurements indicate that the EA forms are more abundant in the gas phase than the canonical KA form. The values of the inertial defects show that all species are effectively planar.

Table 5 Spectroscopic constants for the observed tautomers and conformers of cytosine

	EA _t	EA _c	KA	KI _t	KI _c
A/MHz	3,951.85325(32) ^a	3,889.46510(38)	3,871.54618(31)	3,848.18174(41)	3,861.2966(12)
B/MHz	2,008.95802(12)	2,026.31804(12)	2,024.97804(11)	2,026.31068(31)	2,011.41032(62)
C/MHz	1,332.47228(08)	1,332.86951(10)	1,330.33627(08)	1,327.99167(10)	1,323.19999(22)
N ₁					
χ_{aa} /MHz	−2.6373(13)	−2.8007(18)	1.6211(19)	1.8518(69)	1.898(23)
χ_{bb} /MHz	1.1672(28)	1.0340(27)	1.4772(34)	2.0545(40)	2.104(28)
χ_{cc} /MHz	1.4701(28)	1.7667(27)	−3.0983(34)	−3.9063(40)	−4.002(28)
N ₃					
χ_{aa} /MHz	2.2619(20)	2.2371(23)	2.5217(12)	2.1383(33)	2.105(20)
χ_{bb} /MHz	−3.6570(22)	−3.3890(25)	−3.5140(16)	1.6064(42)	1.764(28)
χ_{cc} /MHz	1.3951(22)	1.1519(25)	0.9923(16)	−3.7448(42)	−3.870(28)
N ₈					
χ_{aa} /MHz	2.2167(17)	2.2237(17)	2.1802(17)	1.8033(87)	−2.091(15)
χ_{bb} /MHz	1.9511(20)	1.9832(20)	1.8429(26)	−3.1572(58)	0.940(14)
χ_{cc} /MHz	−4.1678(20)	−4.2069(20)	−4.0231(26)	1.3539(58)	1.151(14)
$\Delta^b_c/\mu\text{Å}^2$	−0.1676(3)	−0.1767(4)	−0.2212(3)	−0.1789(5)	−0.2023(13)
σ^c /kHz	1.1	1.4	1.1	1.3	3.9
N ^d	71	74	84	54	45

Source: From [169]

^aErrors in parenthesis are expressed in units of the last digit

^b $\Delta_c = I_c - I_a - I_b$ is the inertial defect. Conversion factor: 505,379.1 MHz $\mu\text{Å}^2$

^cStandard deviation of the fit

^dNumber of fitted lines

6 Monosaccharides

Carbohydrates are one of the most versatile biochemical building blocks, widely acting in energetic, structural, or recognition processes [204, 205]. The importance of its structure has been the driving force behind the development of methods for elucidating the shape of their building blocks, monosaccharides. Thus, it comes as no surprise that 3D structures and relative stability of conformers of monosaccharides continue to be an area of great research interest [204, 205]. The subtle variation in hydroxyl arrangement is thought to account for differences in chemical and physical properties of the sugars. This is also relevant to distinguish between different conformers. Additionally, monosaccharides are also of interest in the field of astrophysics. The availability of rotational data has been the main bottleneck for examining the presence of these building blocks in the interstellar medium (ISM) [206]. Based on the rotational spectra identification, the simplest C_2 sugar of glycolaldehyde [207, 208] has been identified, but has yet to detect the C_3 sugar of glyceraldehyde [209].

The experimental results obtained in condensed phases [210–217] seem to indicate that a subtle balance between intrinsic and environmental effects governs the conformational preferences of monosaccharides; the structure and relative stability of isolated sugars are different from their counterparts in solution. To separate these contributions, it is crucial to obtain data on the isolated monosaccharides in the gas phase. This highlights the importance of generating sugars in isolated conditions, free from the influence of environmental effects to determine its intrinsic conformational properties relevant to understand its biological activity [218]. In the particular case of biomolecular building blocks, the group of Prof. Simons in Oxford, one of the pioneers in the field of laser spectroscopy, has heavily contributed to the study of carbohydrates [219, 220 and references therein] as can be seen in the chapter dedicated to this subject in this book [221].

Presently, Fourier transform microwave spectroscopy techniques in supersonic jets, combined with laser ablation techniques [61–63], can bring intact monosaccharides into the gas phase for structural investigation. The low-temperature environment of a supersonic expansion provides the ideal medium for preparing individual conformers of sugars in virtual isolation conditions, ready to be interrogated by a short burst of microwave radiation. To date, rotational investigations of monosaccharides have been carried out for C_4 sugars [222], C_5 sugars [62, 223, 224], and C_6 sugars [225–227]. All factors contributing to stabilization of the observed species are given in the next sections.

6.1 C_4 Sugars: *D*-Erythrose

D-Erythrose ($C_4H_8O_4$, see Fig. 29) may be present in linear or cyclic furanose forms. Aqueous solution NMR studies [228, 229] have shown that furanose forms

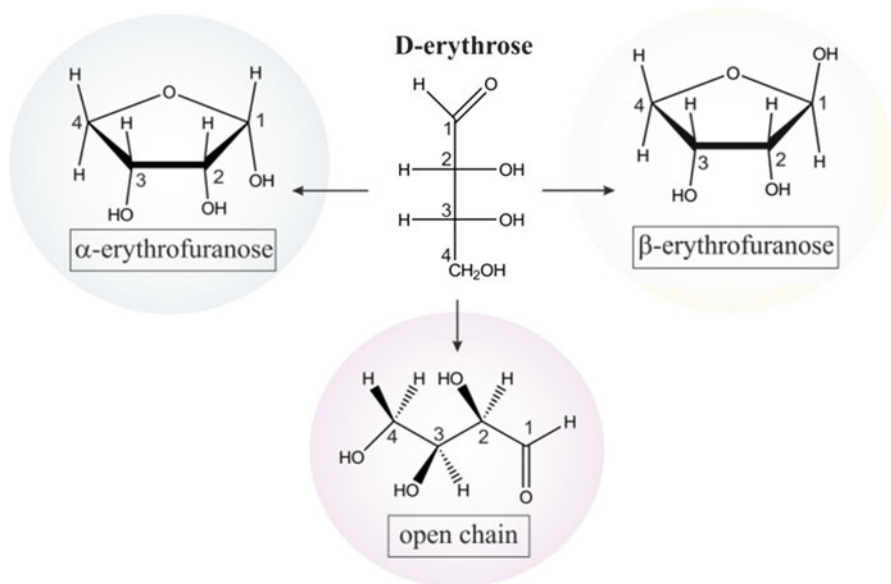


Fig. 29 Fisher projection of D-erythrose (*center*). Haworth projections of the α and β anomers and sketch of the open chain form. (From [222])

(25% and 65% of α and β furanose forms, respectively) are in equilibrium with an appreciable amount of acyclic forms. The five-membered ring structures contain an asymmetric carbon C_1 which leads to the appearance of two stereochemical α and β anomeric species, according to the position of the anomeric OH group (see Fig. 32). Puckering of the ring gives rise to envelope (E) and twist (T) configurations which are interconvertible by rotation of single bonds. This adds complexity, as each of these furanose rings may give rise to many conformers because of the relative arrangement of the OH groups.

Erythrose is a syrup at room temperature and is thermally decomposed using conventional heating methods, so it has been vaporized using laser ablation of solid NaCl doped with D-erythrose [222]. In the experimental procedure, some drops of D-erythrose were mixed with finely powdered NaCl and a small amount of a commercial binder. The laser vaporized products were probed by CP-FTMW spectroscopy [62]. The broadband spectrum (see Fig. 30) shows, apart from the strong NaCl rotational transitions, additional weak lines (inset in Fig. 30) which were attributed to two rotamers A and B of D-erythrose. The derived experimental rotational constants were contrasted with those from *ab initio* calculations on the lowest lying conformations [229]. Rotamer A was unequivocally identified as conformer α - 2 E-cc (predicted as global minimum) and rotamer B as conformer β - 1 T $_2$ -cc. Conformer α - 2 E-cc (Fig. 31) has the three hydroxyl groups on the same side of the furanose ring forming a cyclic cooperative intramolecular hydrogen

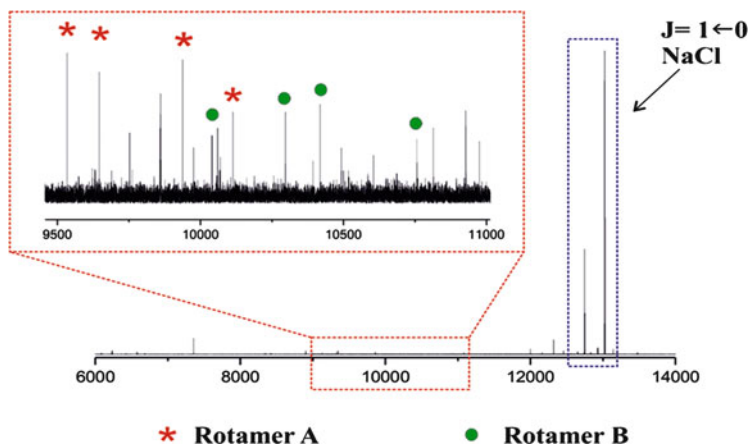


Fig. 30 Broadband CP-FTMW rotational spectrum of D-erythrose and NaCl in the 6–14 GHz frequency region. *Top inset*: details of the CP-FTMW spectrum showing the feature ascribed to several transitions for both detected rotamers. (From [222])

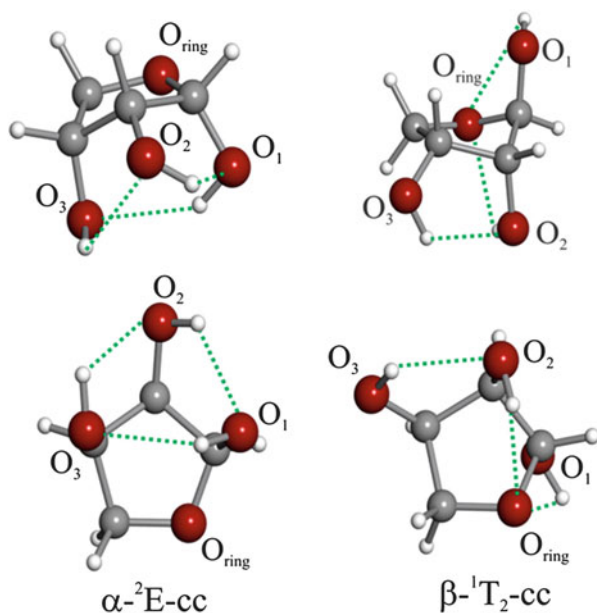


Fig. 31 The three-dimensional structures of the two observed conformers of D-erythrose showing the intramolecular hydrogen bond networks. The notation used to label the conformers includes the symbols α and β to denote the anomer type, E and T with lower and upper subscripts indicate the ring puckerings, and the symbols “c” or “cc” to indicate the clockwise or counterclockwise configuration of the adjacent OH bonds, respectively. (From [222])

bond network, $\text{OH}_1 \cdots \text{OH}_3 \cdots \text{OH}_2 \cdots \text{OH}_1$, with a counterclockwise arrangement. Such cyclic network, explains the stability associated with this conformer.

Conformer $\beta\text{-}^1\text{T}_2\text{-cc}$ has two hydroxyl groups on the same side of the furanose ring. Hence, $\text{OH}_3 \cdots \text{OH}_2 \cdots \text{O}_{\text{ring}}$ and $\text{OH}_1 \cdots \text{O}_{\text{ring}}$ hydrogen bond motifs together with the axial position of the hydroxyl group OH_1 (anomeric effect) are the main stabilizing factors for this conformer.

6.2 2-Deoxy-D-Ribose and Ribose

2-Deoxy-D-ribose (2DR, $\text{C}_5\text{H}_{10}\text{O}_4$) (Fig. 32a) is an important naturally occurring monosaccharide, present on nucleotides' structures, known as the building blocks of DNA [230]. In DNA, 2DR is present in the furanose (five-membered) ring form, whereas in aqueous solution it is present as five- and six-membered rings species, with the latter being dominant [231, 232]. In the six-membered ring, the C_1 carbon atom is an asymmetric centre, yielding two possible stereochemical α and β anomeric species (Fig. 32b). In aqueous solution, 2DR primarily exists as a mixture of nearly equal amounts of α - and β -pyranose forms, present in their low energy chair conformations, $^4\text{C}_1$ and $^1\text{C}_4$ (Fig. 32c) [210, 233–235]. Former experiments to determine the conformation of monosaccharides based on X-ray and NMR

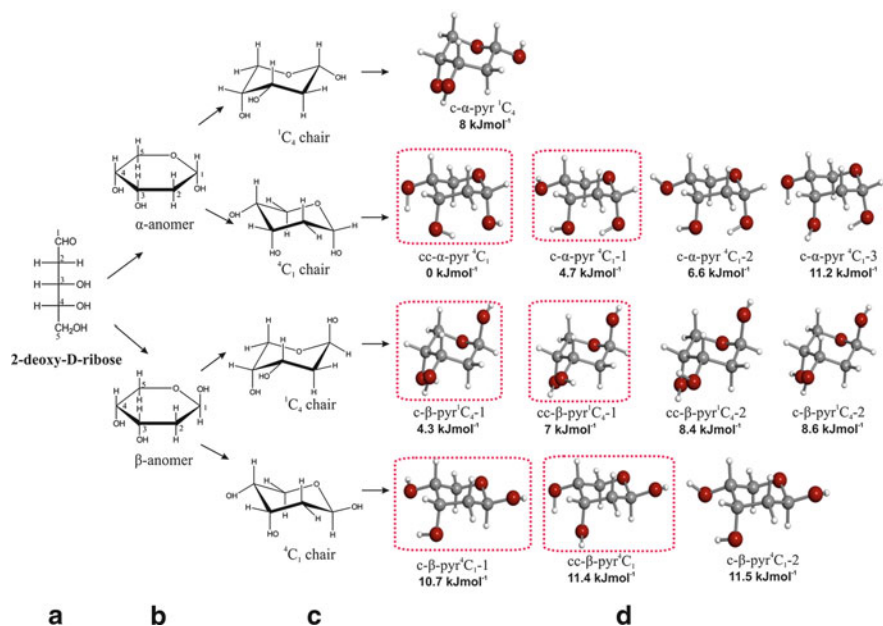


Fig. 32 (a) Fisher projection of 2-deoxy-D-ribose. (b) Haworth projections of α and β anomers. (c) $^1\text{C}_4$ and $^4\text{C}_1$ chair conformations. (d) Predicted conformers within 12 kJ mol^{-1} from MP2(full)/6-311++G(d,p) ab initio computations; the observed conformers are encircled. (From [224])

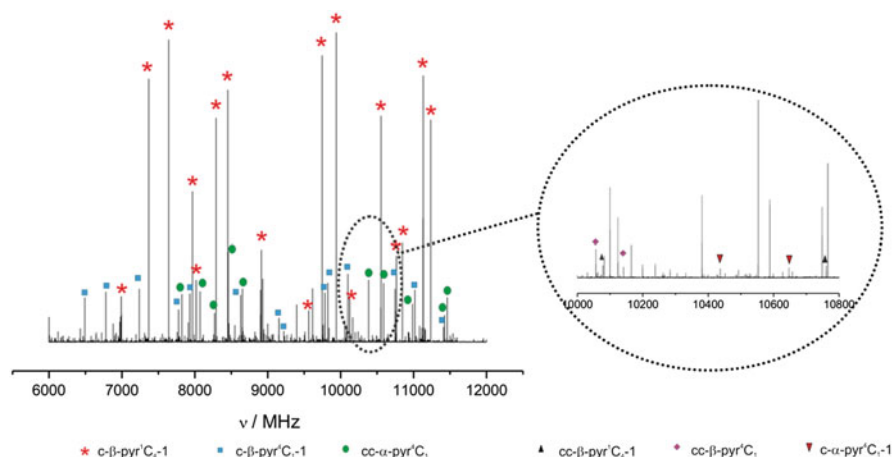


Fig. 33 Broadband microwave spectrum of 2-deoxy-D-ribose. (From [224])

measurements [233, 234, 236, 237] are influenced by environmental effects associated with the solvent or crystal lattice. An IR spectrum of 2DR isolated in an inert matrix [238] has been interpreted by summing the modeled spectra for several α and β conformers.

The conformational panorama of isolated 2-deoxy-D-ribose (m.p. = 89–90°C) has been recently unveiled [224] using CP-FTMW spectroscopy in conjunction with a picosecond laser ablation LA source. The broadband spectra of Fig. 33 allowed the assignment of six different rotameric species labeled I to VI. The rotational constant values were found to be consistent with those predicted ab initio for the conformers shown in Fig. 33d. In addition, spectral measurements have been extended to all five monosubstituted ^{13}C species and the endocyclic ^{18}O species in their natural abundance ($\sim 1.1\%$ and $\sim 0.2\%$) for the most abundant $c\text{-}\beta\text{-pyr-}^1\text{C}_4\text{-1}$ conformer using laser ablation combined with MB-FTMW. The isotopic information was used to derive its structure [224]. The population ratios for α and β conformers estimated from transition intensities indicate that 2DR exists in the gas phase as a mixture of approximately 10% of α - and 90% of β -pyranose forms, thus displaying the dominant $\beta\text{-}^1\text{C}_4$ pyranose form, as found in the previous X-ray crystalline study [236]. No evidence has been found of either α/β -furanoses or any linear forms in gaseous 2DR (Fig. 33).

The detected conformers of 2DR, depicted in Fig. 34, can be rationalized in terms of factors that may contribute to their stabilization. The two observed α conformers, $cc\text{-}\alpha\text{-pyr-}^4\text{C}_1$ and $c\text{-}\alpha\text{-pyr-}^4\text{C}_1\text{-1}$, are stabilized by anomeric effects; they have a $^4\text{C}_1$ ring configuration, thus leading the anomeric OH group towards the axial position. The hydroxy groups of both conformers are located at the same side of the ring, and are able to form chains of hydrogen bonds, which, in turn, are strongly reinforced by sigma hydrogen bond cooperativity [102–104]. The most abundant α form $cc\text{-}\alpha\text{-pyr-}^4\text{C}_1$ presents a counterclockwise arrangement of the OH

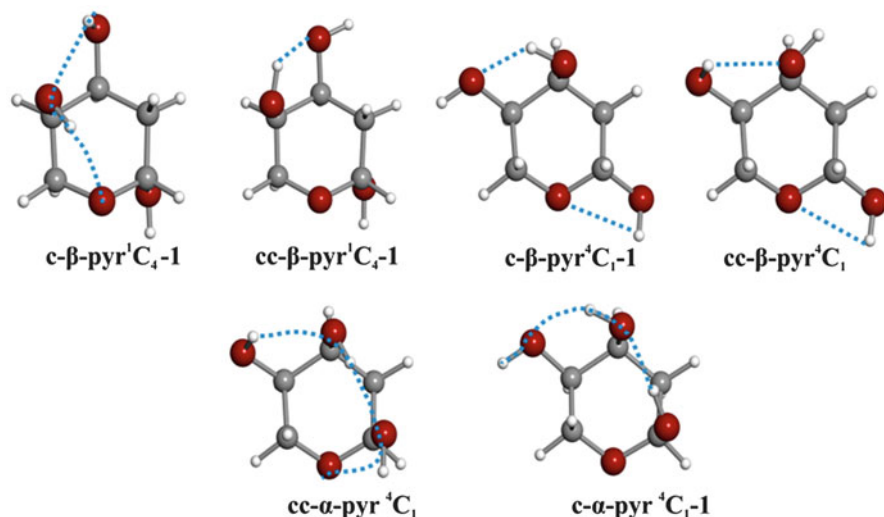


Fig. 34 The six observed conformers of 2-deoxy-D-ribose showing the intramolecular hydrogen bond arrangements. (From [224])

groups with a chain of three hydrogen bonds $O_{(4)}H \cdots O_{(3)}H \cdots O_{(1)}H \cdots O_{ring}$, while the less abundant $c\text{-}\alpha\text{-pyr-}^4C_1-1$ shows a chain of two $O_{(1)}H \cdots O_{(3)}H \cdots O_{(4)}H$. The anomeric effect in the most abundant β form $c\text{-}\beta\text{-pyr-}^1C_4-1$ is reinforced by the intramolecular hydrogen bond network $O_{(3)}H \cdots O_{(4)}H \cdots O_{ring}$. Conformers $c\text{-}\beta\text{-pyr-}^4C_1-1$ and $cc\text{-}\beta\text{-pyr-}^4C_1$, with the anomeric hydroxy group in equatorial position, are stabilized by two non-cooperative intramolecular hydrogen bonds.

Similar to 2DR, ribose ($C_5H_5O_5$) is one of the most important monosaccharides since it constitutes a subunit of the backbone of RNA. NMR studies have shown that ribose in solution is a mixture of α - and β -pyranose and α - and β -furanose forms, the β -pyranose form being predominant. The recently settled crystal structures have shown that the α - and β -pyranose forms are present in the solid phase [239–243]. The structure in the gas phase has been experimentally investigated using a laser ablation molecular beam Fourier transform microwave spectroscopy (LA-MBFTMW) technique [62]. The high resolution rotational spectrum has provided structural information on a total of six rotamers of ribose, three belonging to the α -pyranose forms and other three to the β -pyranose forms. Recently, D-ribose (m.p. 95°C) has been submitted to a laser ablation broadband (CP-FTMW) spectroscopic study and eight conformers (two new α -pyranose forms) have been identified. A broadband section of the spectra is shown in Fig. 35 and the detected conformers depicted in Fig. 36.

Compared to ribose, the absence of the hydroxy group at C_2 in 2-deoxyribose limits the possibility of forming hydrogen bonds and in practice leads to weakening of the cooperative hydrogen-bond network, altering the relative abundances. For example, the most stable α -pyranose form $c\text{-}\alpha\text{-pyr-}^1C_4$ of ribose has not been

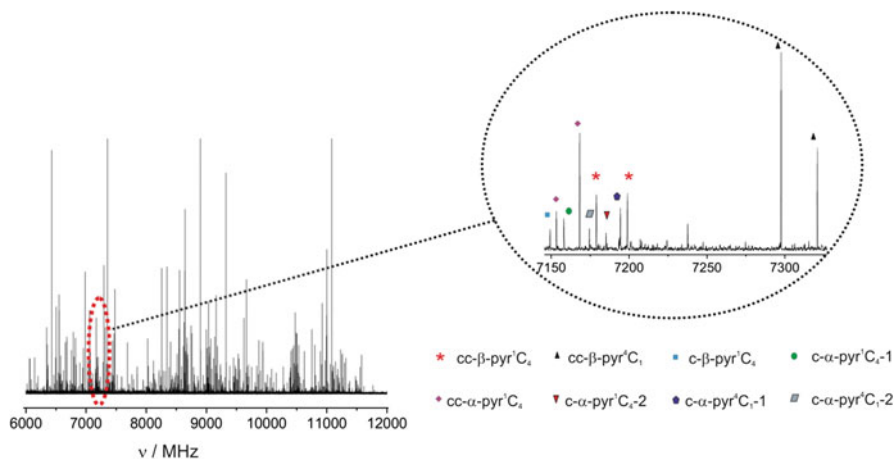


Fig. 35 Broadband spectrum of D-ribose

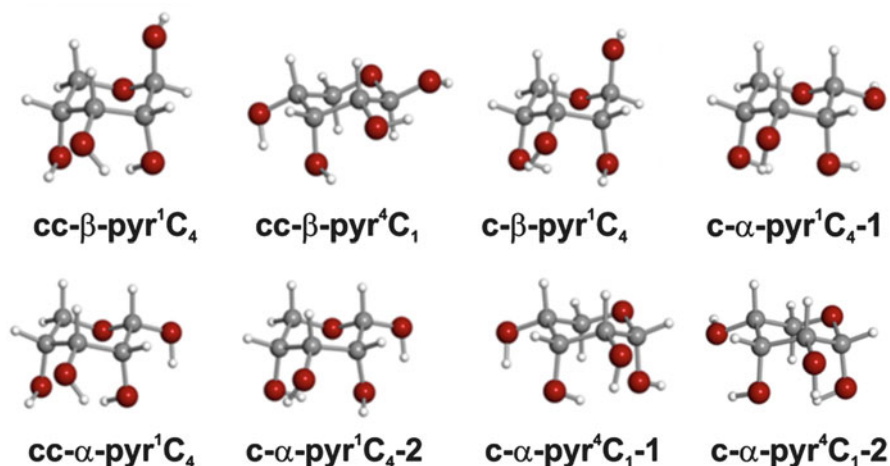


Fig. 36 The eight detected conformers of D-ribose, adopting ¹C₄ or ⁴C₁ -chair structures

detected in 2DR. The absence of an O₍₂₎H group reverses the arrangement of the OH groups in the most stable β-pyranose forms (from clockwise in c-β-pyr¹C₄ of 2DR to counterclockwise in ribose cc-β-pyr¹C₄) to maximize the number of hydrogen bonds (two in cc-orientation vs one in the clockwise arrangement). The evidence collected so far supports that pyranose forms of ribose and deoxyribose are more stable both in gas phase and solution, so the biological pathway to the insertion of furanose forms of ribose and deoxyribose in RNA or DNA cannot be merely attributed to a preference for the furanose forms in the physiological medium.

6.3 D-Glucose and D-Xylose

D-Glucose ($C_6H_{12}O_6$, see Fig. 37a) is one archetypical monosaccharide representing the major building block for many carbohydrate systems [244–247]. Similar to other six-carbon monosaccharides, it may exist in either linear or cyclic, pyranose or furanose, forms. In aqueous solution, NMR studies have shown that the pyranose form is dominant [210, 212, 213]. The cyclization leads to the occurrence of two anomeric species, α and β , according to the position of the OH group (see Fig. 37b). It is commonly believed that the α anomer is more stable than the β anomer because of the stereoelectronic anomeric effect [248, 249]. However, when D-glucose is dissolved in water, the α and β anomers are present in a 40:60 ratio in the 4C_1 ring conformation (Fig. 37c). The observed abundance of the β anomer in water could only be explained by taking into consideration strong solvation effects, which overcome the preference for the α anomer [250–253].

The glucopyranose's hydroxymethyl group configuration must be considered in the conformational analysis of D-glucose in terms of three staggered conformers, designated G+, G–, and T (see Fig. 37d), associated with the C6–O6–C5–O5 torsional angle, which assumes values of *ca.* 60° , -60° or 180° , respectively. Experimental observations in both the solid phase [211, 214–217, 254] and solution [212, 213, 233] display approximately equal populations of G+ and G– conformers, with an almost complete absence of the T conformer. This propensity in glucopyranosides to adopt gauche conformations is known as gauche effect [255, 256 and references therein]. Finally, the structural analysis of D-glucose requires the consideration of intramolecular hydrogen bond networks involving adjacent OH groups. The orientation of the hydroxyl groups is relevant to distinguish between the different conformers of D-glucose and is thought to account for

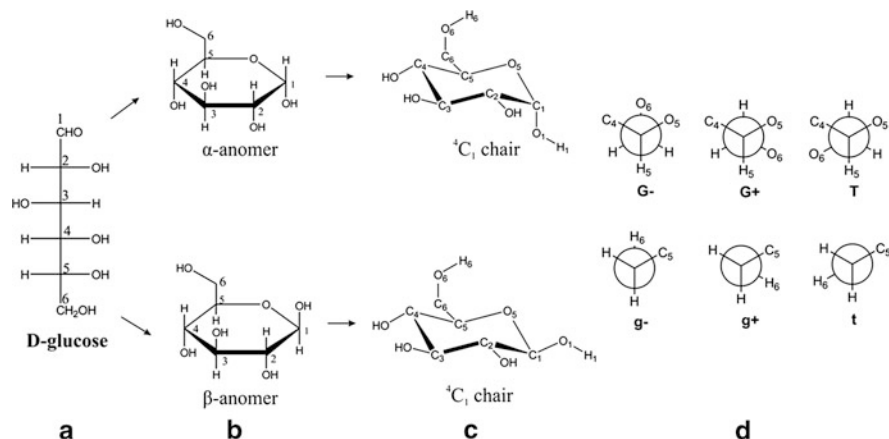


Fig. 37 (a) Fisher projection of D-glucose. (b) Haworth projections of α and β anomers. (c) 4C_1 chair conformations. (d) Newman projections of the plausible conformations of the hydroxymethyl group around C5–C6 and C6–O6 bonds. (From [225])

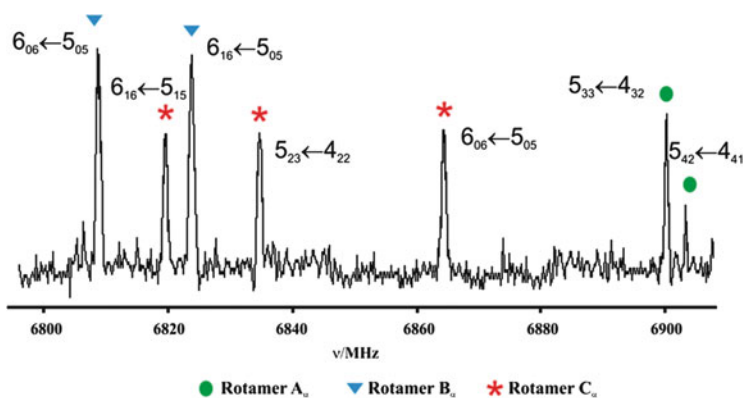


Fig. 38 A section of the LA-MB-FTMW spectrum of α -D-glucopyranose showing the rotational transitions for three of the four observed rotamers. (From [224])

differences in chemical and physical properties of this biologically relevant biomolecule [244–247].

Apart from the number of theoretical studies on D-glucopyranose [257, 258], only one vibrational spectroscopic study of α -D-glucopyranose isolated in Ar matrix has been reported [259]. Laser spectroscopy through UV–UV and IR–UV double-resonance techniques has contributed to the description of the conformations of some β -phenylglucopyranosides and their hydrates [219, 220, 260, 261] but these studies are limited to vibrational resolution and the structural conclusions are not totally transferable to D-glucose because of the electronic chromophore at the anomeric position.

The gas-phase structures of α - and β -D-glucopyranose (m.p = 153°C and 157°C, respectively) have been examined using LA-MB-FTMW spectroscopy [225] (see Fig. 38). After completing a wide frequency scan, it became possible to assign the rotational spectra of four different rotamers of α -D-glucopyranose. Their identification was based on the agreement between the experimental and theoretical values of the rotational constants and their trends of variation upon subtle structural changes. In the same way, the agreement between the electric dipole moment s and the observed rotational selection rules allow one to confirm the assignments. Relative populations of the identified conformers (see Fig. 39) G-g+/cc/g+ : G-g-/cc/g+ : Tg+/cc/g+ : G-g+/cl/g- = 1:0.9(2):0.5(1):0.4(2), estimated by relative intensity measurements of rotational transitions, were found to be in reasonable agreement with those calculated from the ab initio Gibbs free energies of 1:0.90:0.30:0.26. Gauche 4C_1 glucopyranose forms with a counterclockwise arrangement of OH groups dominating the conformational panorama of α -D-glucopyranose. The four observed conformers of α -glucopyranose, depicted in Fig. 42, are stabilized by anomeric effect; they have a 4C_1 ring configuration with the anomeric OH group towards the axial position. The hydroxyl groups located at equatorial positions are able to form chains of hydrogen bonds, strongly reinforced by sigma-hydrogen bond cooperativity [102–104]. The most abundant α conformers, G-g+/cc/t and G

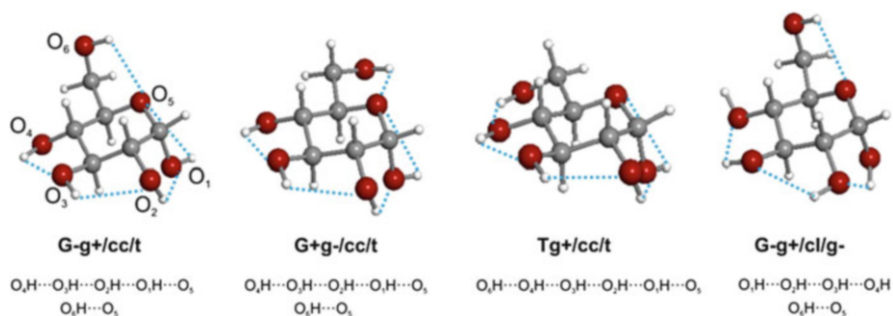


Fig. 39 The four observed conformers of α -D-glucopyranose showing the intramolecular hydrogen bond networks. The symbol in capital letters, $G+$, $G-$, or T , describe the torsion angle $O_6-C_6-C_5-O_5$ (see Fig. 1) of about 60° , -60° or 180° , respectively, which describe the configuration of the hydroxymethyl group. The lower case symbol $g+$, $g-$, or t , describe in the same way the torsion angle $H_6-O_6-C_6-C_5$ (see Fig. 37d). These symbols are followed by a slash and the symbol cl or cc describing, respectively, the clockwise (cl) or counterclockwise (cc) arrangement of the cooperative network of intramolecular hydrogen bonds. Then, after the slash, the last symbol, $g+$, $g-$ or t , gives the value of the torsion angle $H_1-O_1-C_1-C_2$ which describes the orientation of the anomeric hydroxyl group hydrogen atom. (From [225])

$+g-/cc/t$, present a counterclockwise arrangement of the OH groups with a chain of four cooperative hydrogen bonds $O_4H \cdots O_3H \cdots O_2H \cdots O_1H \cdots O_5$ and additional $O_6H \cdots O_5$ interaction with $G-$ or $G+$ configurations of hydroxymethyl side chains, respectively. The least abundant $G-g+/cl/g-$ conformer presents a clockwise arrangement of three cooperative hydrogen bonds $O_1H \cdots O_2H \cdots O_3H \cdots O_4H$ and one non-cooperative $O_6H \cdots O_5$. The $O_1H \cdots O_5$ interaction does not take place in the clockwise oriented network explaining its low abundance. Gauche 4C_1 glucopyranose forms with a counterclockwise arrangement of OH groups dominate the conformational panorama of α -D-glucopyranose.

In β -D-glucopyranose, rotational spectra revealed the presence of three conformers $G-g+/cc/t$, $G+g-/cc/t$ and $Tg+/cc/t$ with relative abundances 0.9(2):1:0.2 (1), respectively. The two most abundant β conformers $G-g+/cc/t$ and β - $G+g-/cc/t$ shown in Fig. 40, exhibit the same conformational shape as observed in α forms with the obvious differences in the anomeric OH group.

The observation of conformers with an *anti* orientation of the dihedral angle ($O_6-C_6-C_5-O_5$) in α - and β -D-glucopyranose constitutes a remarkable fact. Numerous experimental studies on α - and β -glucopyranosides, both in solid [211, 214–217, 254] and solution phases [262–264], have shown that the dihedral angle ($O_6-C_6-C_5-O_5$) displays a preference for $G-$ and $G+$ gauche configurations, which has been attributed to the gauche effect [255]. This feature was exemplified in a statistical analysis of X-ray structures of glucopyranosyl derivatives [265], yielding a rotamer population of 40:0:60 ($G+/T/G-$). In contrast to previous results, our gas-phase experiment revealed the existence of *trans* configurations in α - $Tg+/cc/t$ and β - $Tg+/cc/t$ conformers. In agreement with *ab initio* calculations, these forms have a higher energy and are less abundant in the jet. Both conformers exhibit a

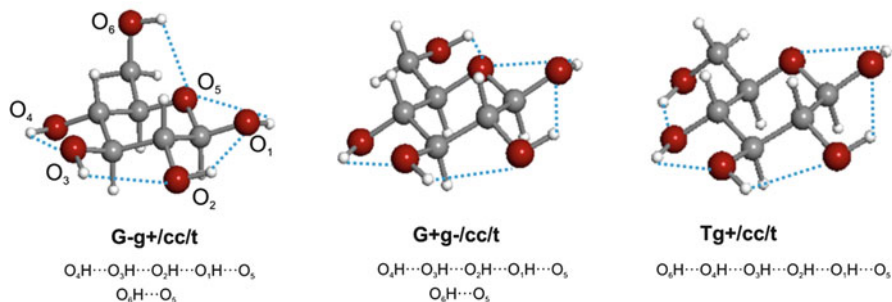


Fig. 40 The three observed conformers of β -D-glucopyranose showing the intramolecular hydrogen bond networks. (From [225])

chain of five cooperative hydrogen bonds $O_6H \cdots O_4H \cdots O_3H \cdots O_2H \cdots O_1H \cdots O_5$ oriented counterclockwise involving the hydroxymethyl group. Therefore, the structure and relative stability of isolated α - and β -D-glucopyranose are different from their counterparts in condensed phases.

D-Xylose ($C_5H_{10}O_5$) is the aldopentose analogue to the aldohexose D-glucose, lacking the hydroxymethyl group. Xylose exists predominantly in the pyranose form in the condensed phase [210, 233, 234, 266], and it is believed to maintain this structure in the gas phase [267]. The pyranose structures have two enantiomers designated as α and β depending on the OH group position at the chiral C_1 . Crystalline samples of D-xylose (m.p. = 162°C) have been vaporized by laser ablation and probed by CP-FTMW [222] using a new parabolic reflector system [268]. The recorded broadband spectrum, shown in Fig. 41, allows identification of the conformers $cc\text{-}\alpha\text{-}^4C_1$ and $c\text{-}\alpha\text{-}^4C_1$ (see Fig. 42). The population ratio estimated from the relative intensity measurements $cc\text{-}\alpha\text{-}^4C_1 : c\text{-}\alpha\text{-}^4C_1 \approx 1:0.03$ is in good agreement with the theoretical predictions. The isotopic information corresponding to the five monosubstituted ^{13}C species of the most abundant conformer $cc\text{-}\alpha\text{-}^4C_1$ was used to derive its structure [222]. Despite the sensitivity reached in the experiment, no traces belonging to β -pyranose forms have been detected. This is in accordance with the fact that the rotational spectra of laser ablated crystalline D-xylose should reflect the α form found in the crystalline sample [266]. The interconversion between α and β anomers is usually a solvent-mediated reaction and would not occur that easily during the laser ablation process or in the gas phase [269].

The conformers of α -D-xylose, depicted in Fig. 42, present the most favorable chair configuration 4C_1 , with the largest substituent in equatorial position and the anomeric hydroxyl group in axial position (anomeric effect) [248, 249]. They correlate with the corresponding G-g+/cc/g+ and G-g+/cl/g- of α -D-glucopyranose. Both conformers show arrangements of hydroxyl groups into intramolecular H-bond networks which can lead to a phenomenon known as cooperativity [102–104]. Under cooperativity, directionally arranged H-bonds that form an H-bond network can increase the strength of an individual H-bond donor or acceptor.

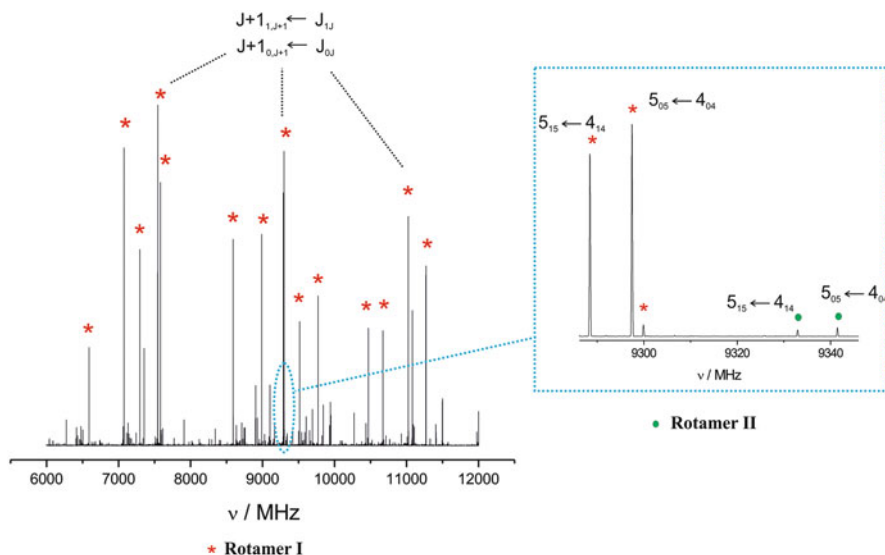


Fig. 41 The broadband rotational spectrum of D-xylose showing the intense rotational transitions for rotamer I. The *inset* shows the characteristic μ_a -R-branch progressions for rotamer I and rotamer II. (From [223])

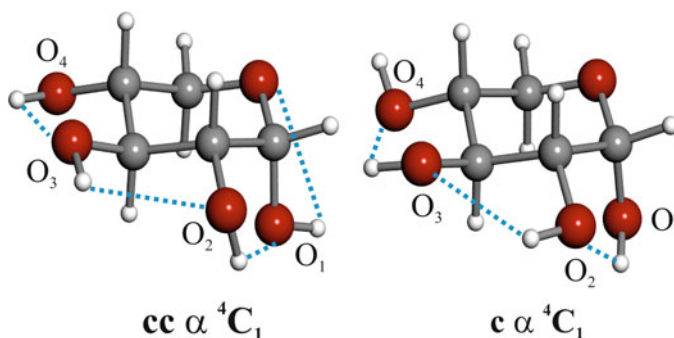


Fig. 42 The three-dimensional structures of the two observed conformers of α -D-xylose showing the intramolecular hydrogen bond arrangements. (From [223])

The most abundant conformer $cc\text{-}\alpha^4C_1$ exhibits a counterclockwise arrangement with a chain of four hydrogen bonds $O_{(4)}H \cdots O_{(3)}H \cdots O_{(2)}H \cdots O_{(1)}H \cdots O_{\text{ring}}$. The less abundant $c\text{-}\alpha^4C_1$ species presents a three intramolecular hydrogen bond network $O_{(1)}H \cdots O_{(2)}H \cdots O_{(3)}H \cdots O_{(4)}H$ orientated clockwise. The $O_{(1)}H \cdots O_{\text{ring}}$ hydrogen bond found in conformer $cc\text{-}\alpha^4C_1$ might be the cause of the over-stabilization of this species. These results demonstrate the pivotal role that

intramolecular hydrogen-bonding network plays in the conformational behavior of free monosaccharides.

Gas phase structures of phenyl α - and β -D-xylopyranoside have been investigated by UV–UV and IR–UV laser spectroscopic techniques coupled with theoretical calculations [270]. The authors hypothesized that the substitution of the anomeric OH group for a phenoxy group compatible with the UV excitation scheme has little effect in the conformational behavior. Present results clearly show that the substitution of the anomeric OH group by the chromophore phenoxy affects the intramolecular hydrogen bond network and, consequently, the conformational behavior and the structural conclusions are transferable to D-xylose. Indeed, the related $c\text{-}\alpha\text{-}^4C_1$ conformer has not been observed in the phenyl derivative.

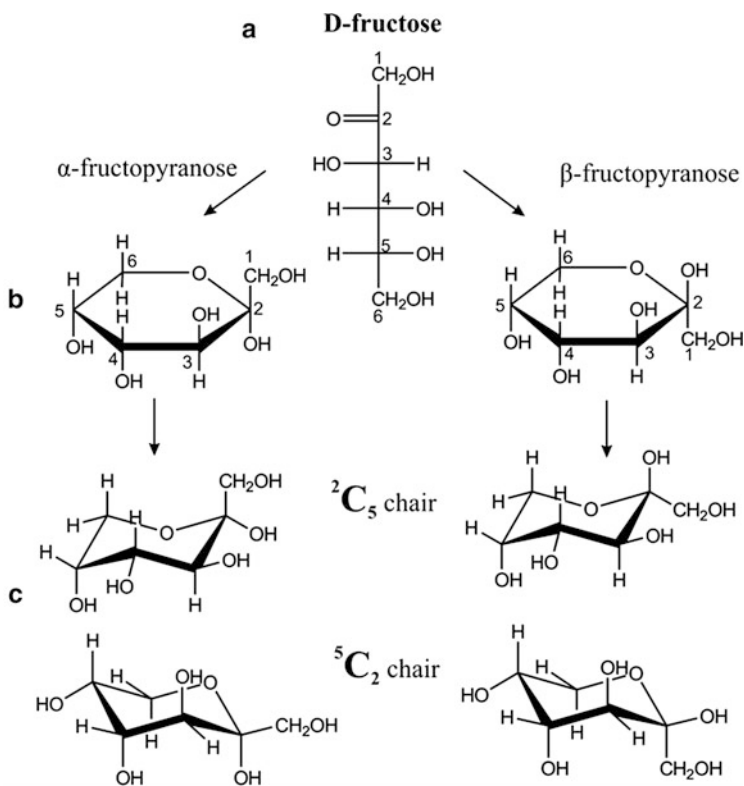


Fig. 43 (a) Fisher projection of D-fructose. (b) Haworth projections of α and β anomers. (c) 2C_5 and 5C_2 chair configurations. (From [226])

6.4 D-Fructose

D-Fructose ($C_6H_{12}O_6$) is a six-carbon polyhydroxyketone (Fig. 43a). Although ketohexoses such as fructose can exhibit a linear form, D-fructose rapidly cyclizes in aqueous solution to form mixtures of pyranose and furanose forms [246, 247]. The cyclization reaction converts C_2 in a chiral carbon, yielding two enantiomers designated α and β (Fig. 43b). For D-fructose, the equilibrium concentrations in water are around 82% of pyranose forms and 12% of furanose forms [271]. However, in its crystalline form, the unique species found is the β -D-fructopyranose [272, 273]. Besides other higher energy forms, pyranoses preferably adopt a rigid chair backbone with the conformations 5C_2 and 2C_5 shown in Fig. 43c.

The most stable structures of D-fructose in isolated conditions of gas phase have been unveiled by CP-FTMW [226] spectroscopy, bringing crystalline D-fructopyranose (m.p.120°C) into the gas phase by laser ablation. Once the lines from known photofragmentation species are removed from the broadband spectrum (see Fig. 44), rotational spectra of two rotamers I and II could be identified. The match between the experimental and calculated ab initio values of the rotational constants leads to the irrefutable identification of rotamer I as conformer cc β 2C_5 g– and rotamer II as conformer cc β 2C_5 t. Both conformers (see Fig. 45) are present in the most favorable chair configuration 2C_5 , with the largest substituent in an equatorial position and the anomeric hydroxyl group ($OH_{(2)}$) in an axial position (anomeric effect) [248, 249]. It is stabilized by a five cooperative intramolecular hydrogen bond network (Fig. 2) $OH_{(5)} \cdots OH_{(4)} \cdots OH_{(3)} \cdots OH_{(2)} \cdots OH_{(1)} \cdots O_{(ring)}$, with a counterclockwise arrangement of the OH groups. This hydrogen bond cooperative interaction is a form of intramolecular solvation which reinforces the

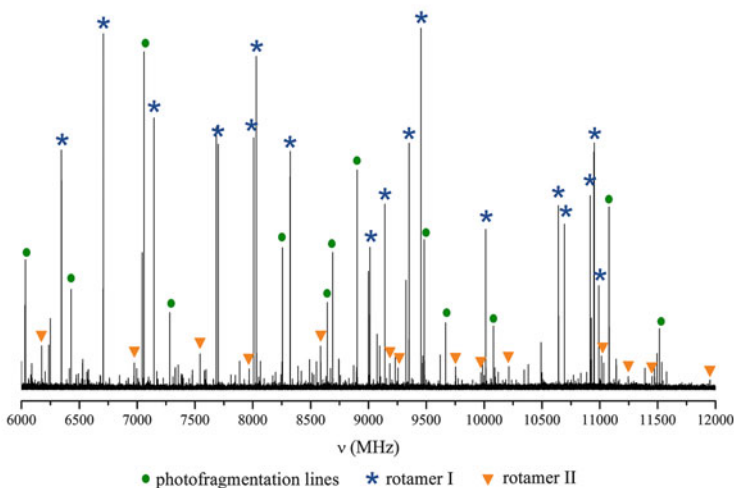


Fig. 44 Broadband microwave spectrum of D-fructose. (From [226])

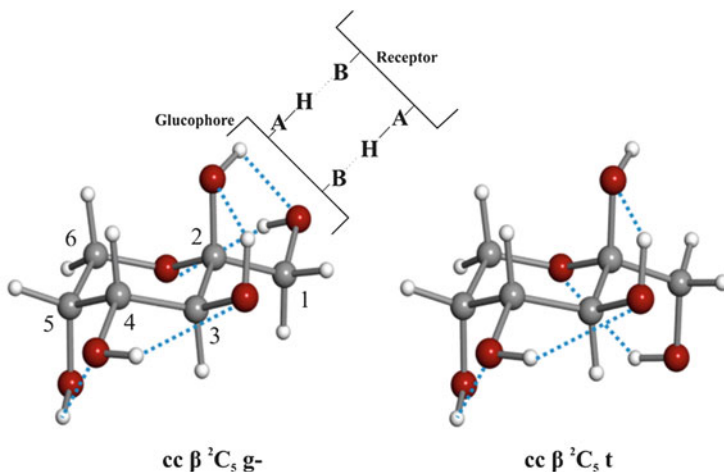


Fig. 45 The three-dimensional structures of the two observed conformers of β -D-fructopyranose showing the intramolecular hydrogen-bonding networks and the glucophore unit. (From [226])

stability. Conformer $cc \beta^2C_5 t$ shows a chain of three cooperative $OH_{(5)} \cdots OH_{(4)} \cdots OH_{(3)} \cdots OH_{(2)}$ and one non-cooperative $OH_{(1)} \cdots O_{(ring)}$ hydrogen bonds. The estimated population ratio $\beta^2C_5 g^- : cc \beta^2C_5 t$ as 1: 0.02 is in excellent accordance with the computed energies; conformer $cc \beta^2C_5 t$ is predicted 700 cm^{-1} above the global minimum $cc \beta^2C_5 g^-$. The most abundant conformer $cc \beta^2C_5 g^-$ has also been examined by LA-MB-FTMW spectroscopy [227] and its structure has been determined from the rotational spectra of parent and monosubstituted isotopic species.

D-Fructose is the sweetest naturally occurring carbohydrate. Its sweetness in solution is directly related to the D-fructopyranose proportion [274–279]. A simple rationalization of the structure–sweetness relationship is based on the existence of a basic structural unit formed by proton donating A-H and proton accepting B electronegative groups [274–279]. The concept of a tripartite glucophore AH-B (“sweetness triangle”) has its merit as a unifying criterion proved useful in rationalizing the sweetness in diverse classes of compounds. The assignment of the AH-B tripartite in detected conformers in Fig. 45 is very complicated as each OH group can function as AH and/or B. After examination of the structure of $cc \beta^2C_5 g^-$ conformer, one can state that $OH_{(1)}$ and $O_{(2)}$ can be considered as the most likely AH-B glucophore causing the sweet response via interaction with a complementary hydrogen bond donor and acceptor in the taste receptor. Hence, the most abundant conformer $cc \beta^2C_5 g^-$ might be responsible for the sweetness of D-fructose.

7 Summary and Outlook

The main objective of this chapter is to summarize the advances attained in the knowledge of the structure of isolated biomolecular building blocks by the combination of the experimental methods that combine laser ablation with Fourier transform microwave spectroscopy techniques in supersonic jets.

The results described confirm that these methods are formidable tools to investigate the conformational landscape of solid biomolecules in the gas phase.

The high accuracy of frequency measurements, the resolution achieved, and the supersonic expansion cooling are clear advantages for the study of the complex conformational behavior of such molecules. They allow unambiguous discrimination between conformers and provide rich information about the intermolecular forces at play. The molecular properties extracted from the analysis of the rotational spectrum, such as the rotational and quadrupole coupling constants, can be directly compared with those predicted *ab initio* to achieve a conclusive identification of the most abundant conformers in the supersonic expansion. ^{14}N quadrupole coupling interaction has been shown to be an invaluable alternative tool to identify the nature and most stable forms of these compounds. The characteristic hyperfine structure pattern produced by this coupling has been found to be a fingerprint for each observed species. Moreover, rotational data provide a benchmark against which quantum theory calculations can be checked.

Thanks to the LA-MB-FTMW technique, knowledge of the structural properties of non-aromatic neutral amino acids in the gas phase, unavailable through other techniques, has expanded considerably. The investigation of individual neutral amino acids in the gas phase reveals a complex and subtle network of forces that condition the conformations adopted by amino acids. Conformational stabilization is dictated, in all cases, by hydrogen bonding. The presence of a polar side chain has been found to play an influential role in all α -amino acids studied, reflected in the increased number of low-energy conformers, with non-polar side-chain amino acids exhibiting a limited conformational variety. The hydrogen bond interaction of the side chain with the other polar groups, in the amino acid backbone, may reverse the relative stabilities of conformers or even increase the interconversion barriers. Interestingly, in asparagine, a conformational locking occurs because of the interplay of several hydrogen bond interactions. The increase of the amino acid backbone chain in the series of α -, β -, and γ -amino acids also has important consequences for conformational behavior, because of the presence of other stabilization forces such as $n-\pi^*$ interactions.

The LA-MB-FTMW technique has also contributed to extending our knowledge on the tautomer equilibria in important nucleobases such as guanine or cytosine. Again, ^{14}N quadrupole coupling interaction has been shown to be instrumental in identifying the nature of the different nitrogen atoms and thus to establish unambiguously the most stable tautomers/conformers of these compounds. Moreover, characteristic hyperfine structure patterns produced by this coupling have been found to be fingerprints for each observed species.

The future perspectives of this research area depend heavily on the application of CP-FTMW spectrometers, which have changed the scope of rotational spectroscopy in recent years. Chirped-pulse Fourier-transform microwave spectroscopy, combined with laser ablation, opens a new era in the investigation of isolated biomolecules, as can be shown by its application to carbohydrates as discussed in this chapter. The broad frequency and large dynamic range make it possible to extend the range of detectable conformers to less stable forms and to access structural determinations in molecular systems of increasingly larger sizes from measurements of heavy atom (^{13}C , ^{15}N , ^{18}O) isotopes detected in natural abundance. This opens promising perspectives for the structural determination of key systems, and the study of heavier systems such as new amino acids, sugars, dipeptides or tripeptides, glycosides, and other biomolecules of increasing complexity.

References

1. Schermann JP (2008) Spectroscopy and modelling of biomolecular building blocks. Elsevier, Amsterdam
2. Robertson EG, Simons JP (2001) Getting into shape: conformational and supramolecular landscapes in small biomolecules and their hydrated clusters. *Phys Chem Chem Phys* 3:1
3. Zwier TS (2001) Laser spectroscopy of jet-cooled biomolecules and their water-containing clusters: water bridges and molecular conformation. *J Phys Chem A* 105:8827
4. Weinkauff RP, Schermann J-P, de Vries MS, Kleinermanns K (2002) Molecular physics of building blocks of life under isolated or defined conditions. *Eur Phys J D* 20:309
5. Simons JP (2004) Bioactive molecules in the gas phase. *Phys Chem Chem Phys* 6:E7
6. Simons JP (2007) Spectroscopic probes of molecular recognition. *Phys Chem Chem Phys* 32:4429–4608
7. Kistenmacher TJ, Rand GA, Marsh RE (1974) Refinements of the crystal structures of DL-serine and anhydrous L-serine. *Acta Crystallogr B* 30:2573
8. Görbitz C, Dalhus HB (1996) Redetermination of L-leucine at 120K. *Acta Crystallogr C* 52:1754
9. Levy DH (1980) Laser spectroscopy of cold gas-phase molecules. *Annu Rev Phys Chem* 31:197
10. Scoles G (ed) (1988) Atomic and molecular beam methods, vol 1. Oxford University Press, New York/Oxford
11. Martinez SJ, Alfano JC, Levy DH (1992) *J Mol Spectrosc* 156:421
12. Gordy W, Cook RL (1984) Microwave molecular spectra. Wiley, New York
13. Structure data of free polyatomic molecules, vol 7 (1976), vol 15 (1987), vol 21 (1992), vol 23 (1995), vol 25A (1998), vol 25B (1999), vol 25C (2000), vol 25D (2003), vol 28A (2006), vol 28B (2006), vol 28C (2007), vol 28D (2007). Springer, Berlin
14. Graner G, Hirota E, Iijima T, Kuchitsu K, Ramsay DA, Vogt J, Vogt N (1998) Structure data of free polyatomic molecules. Basic data. Springer, Berlin, 214 pp
15. Balle TJ, Flygare WH (1981) *Rev Sci Instrum* 52:33
16. Grabow J-U, Stahl W (1990) *Z Naturforsch A* 45:1043
17. Andresen U, Dreizler H, Grabow JU, Stahl W (1990) *Rev Sci Instrum* 61:3694
18. Brown GG, Dian BC, Douglass KO, Geyer SM, Shipman ST, Pate BH (2008) *Rev Sci Instrum* 79:053103

19. Grabow J-U, Mata S, Alonso JL, Peña I, Blanco S, López JC, Cabezas C (2011) *Phys Chem Chem Phys* 13:21063
20. Leopold KR, Fraser GT, Novick SE, Klemperer W (1994) *Chem Rev* 94:1807
21. Bauder A (1996) In: Fausto R (ed) *Low temperature molecular spectroscopy*. Kluwer, The Netherlands, pp 271–289
22. Legon AC, Millen DJ (1986) *Chem Rev* 86:635
23. Endo Y, Kohguchi H, Ohshima Y (1994) *Faraday Discuss* 97:341
24. McCarthy MC, Thaddeus P (2001) *Chem Soc Rev* 30:177
25. Sutter DH, Dreizler H (2001) *Z Naturforsch A Phys Sci* 56:425
26. Gerry MCL (1992) *Chem Phys Lett* 188:213
27. Brupbacher B, Brupbacher T (1999) *J Chem Phys* 111:6300
28. Antolínez S, López JC, Alonso JL (1999) *Angew Chem Int Ed* 38:1772
29. Sanz ME, López JC, Alonso JL (1999) *Chem Eur J* 5:3293
30. Sanz ME, Lesarri A, López JC, Alonso JL (2001) *Angew Chem Int Ed* 40:935
31. Antolínez S, López JC, Alonso JL (2001) *ChemPhysChem* 2:114
32. Alonso JL, Antolínez S, Blanco S, Lesarri A, López JC, Caminati W (2004) *J Am Chem Soc* 126:3244
33. Caminati W, López JC, Alonso JL, Grabow J-U (2005) *Angew Chem Int Ed* 44:2
34. Ottaviani P, Caminati W, Favero LB, Blanco S, López JC, Alonso JL (2006) *Chem Eur J* 12:915
35. López JC, Caminati W, Alonso JL (2006) *Angew Chem Int Ed* 45:290
36. Blanco S, Sanz ME, Mata S, Lesarri A, López JC, Dreizler H, Alonso JL (2003) *Chem Phys Lett* 375:355
37. Blanco S, Sanz ME, Lesarri A, López JC, Alonso JL (2004) *Chem Phys Lett* 397:379
38. Blanco S, López JC, Sanz ME, Lesarri A, Dreizler H, Alonso JL (2004) *J Mol Spectrosc* 227:202
39. Sanz ME, Alonso JL, Blanco S, Lesarri A, López JC (2004) *Astrophys J* 621:L157
40. Suenram RD, Lovas FJ, Fraser GT (1988) *J Mol Spectrosc* 127:472
41. Welzel A, Stahl W (1999) *Phys Chem Chem Phys* 1:5109
42. Blanco S, López JC, Alonso JL, Ottaviani P, Caminati W (2003) *J Chem Phys* 119:880
43. Brown RD, Godfrey PD, Storey JWV, Bassez MPJ (1978) *Chem Soc Chem Commun* 547–548
44. Suenram RD, Lovas FJ (1978) *J Mol Spectrosc* 72:372
45. Levis RJ (1994) *Annu Rev Phys Chem* 45:483
46. Simard B, Mitchell SA, Humphries MR, Hackett PA (1988) *J Mol Spectrosc* 129:186
47. Cable JR, Tubergen MJ, Levy DH (1989) *J Am Chem Soc* 111:9032
48. Barnes M, Fraser MM, Hajigeorgiou PG, Merer AJ, Rosner SD (1995) *J Mol Spectrosc* 170:449
49. Elam JW, Levy DH (1998) *J Phys Chem B* 102:8113
50. Suenram RD, Lovas FJ, Fraser GT, Matsumura K (1990) *J Chem Phys* 92:4724
51. Walker KA, Gerry MCL (1997) *J Mol Spectrosc* 182:178
52. Oshima Y, Endo Y (1993) *Chem Phys Lett* 213:95
53. Low RJ, Varberg TD, Sonnely JP, Auty AR, Howard BJ, Brown JM (1993) *J Mol Spectrosc* 161:499
54. Kretschmer U, Consalvo D, Knaack A, Schade W, Stahl W, Dreizler H (1996) *Mol Phys* 87:1159
55. Lovas FJ, Kawashima Y, Grabow J-U, Suenram RD, Fraser GT, Hirota E (1995) *Astrophys J Lett* 455:L201
56. Lesarri A, Mata S, López JC, Alonso JL (2003) *Rev Sci Instrum* 74:4799
57. Antolínez S, Lesarri A, Mata S, Blanco S, López JC, Alonso JL (2002) *J Mol Struct* 612:125
58. Lesarri A, Mata S, Blanco S, López JC, Alonso JL (2004) *J Chem Phys* 120:6191
59. Lesarri A, Mata S, Cocinero EJ, Blanco S, López JC, Alonso JL (2002) *Angew Chem Int Ed* 41:4673

60. Blanco S, Sanz ME, López JC, Alonso JL (2007) *Proc Natl Acad Sci U S A* 104:20183
61. Alonso JL, Pérez C, Sanz ME, López JC, Blanco S (2009) Seven structures of threonine in the gas phase: a LA-MB-FTMW study. *Phys Chem Chem Phys* 11:617
62. Cocinero EJ, Lesarri A, Ecija P, Basterretxea FJ, Grabow J-U, Fernández JA, Castaño F (2012) *Angew Chem Int Ed* 51:3119–3124
63. Mata S, Peña I, Cabezas C, López JC, Alonso JL (2012) *J Mol Spectrosc* 280:91–96
64. Peña I, Daly AM, Cabezas C, Mata S, Bermúdez C, Niño A, López JC, Grabow J-U, Alonso JL (2013) *J Phys Chem Lett* 4:65–69
65. Cabezas C, Alonso JL, López JC, Mata S (2012) *Angew Chem Int Ed* 51:1375–1378
66. Atkinson DB, Smith MA (1995) *Rev Sci Instrum* 66:4434
67. Peña I, Sanz ME, López JC, Alonso JL (2012) Preferred conformers of proteinogenic glutamic acid. *J Am Chem Soc* 134:2305
68. 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) *Gaussian 09*, revision B.01. Gaussian, Wallingford
69. Becke AD (1993) *J Chem Phys* 98:5648–5652
70. Lee C, Yang W, Parr RG (1988) *Phys Rev B* 37:785–789
71. Vosko SH, Wilk L, Nusair M (1980) *Can J Phys* 58:1200–1211
72. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) *J Phys Chem* 98:11623–11627
73. Ditchfield R, Hehre WJ, Pople JA (1971) Self-consistent molecular orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J Chem Phys* 54:724
74. Møller C, Plesset MS (1934) *Phys Rev* 46:618
75. Raghavachari K, Binkley JS, Seeger R, Pople JA (1980) Self-consistent molecular orbital methods. 20. Basis set for correlated wave-functions. *J Chem Phys* 72:650–654
76. Townes CH, Schawlow AL (1975) *Microwave spectroscopy*. Dover, New York
77. Krotto HW (1975) *Molecular rotation spectra*. Wiley, London
78. Albrecht G, Corey RB (1939) *J Am Chem Soc* 61:1087
79. Marsh RE (1958) *Acta Crystallogr* 11:654
80. Levy HA, Corey RB (1941) *J Am Chem Soc* 63:2095
81. Donohue J (1950) *J Am Chem Soc* 72:949
82. Hirokawa S (1955) *Acta Crystallogr* 8:637
83. Gaffney JS, Pierce RC, Friedman L (1977) *J Am Chem Soc* 99:4293
84. Edsall JT, Blanchard MH (1933) *J Am Chem Soc* 55:2337
85. Ellzy MW, Jensen JO, Hameka HF, Kay JG (2003) *Spectrochim Acta A* 59:2619
86. Meister A (1965) *Biochemistry of the amino acids*, vol 2, 2nd edn. Academic, New York
87. Barrett GC (ed) (1985) *Chemistry and biochemistry of the amino acids*. Chapman and Hall, Glasgow
88. Barrett GC (1998) *Amino acids and peptides*. Cambridge University Press, Cambridge
89. Schafer L, Sellers HL, Lovas FJ, Suenram RDJ (1980) *Am Chem Soc* 102:6566–6568
90. Suenram RD, Lovas FJ (1980) *J Am Chem Soc* 102:7180
91. Godfrey PD, Brown RD (1995) *J Am Chem Soc* 117:2019
92. McGlone SJ, Elmes PS, Brown RD, Godfrey PD (1999) *J Mol Struct* 485:225
93. Godfrey PD, Firth S, Hatherley LD, Brown RD, Pierlot AP (1993) *J Am Chem Soc* 115:9687

94. Ilyushin VV, Alekseev EA, Dyubko SF, Motiyenko RA, Lovas FJ (2005) Millimeter wave spectrum of glycine. *J Mol Spectrosc* 231:15
95. Császár AG (1992) *J Am Chem Soc* 114:9568
96. Hu C-H, Shen M, Schaefer HF (1993) *J Am Chem Soc* 115:2923
97. Barone V, Adamo C, Lejl F (1995) *J Chem Phys* 102:364
98. Császár AG (1995) *J Mol Struct* 346:141
99. Bludský O, Chocholoušová J, Vacek J, Huisken F, Hobza P (2000) *J Chem Phys* 113:4629
100. Alonso JL, Cocinero EJ, Lesarri A, Sanz ME, López JC (2006) The glycine-water complex. *Angew Chem Int Ed* 45:3471
101. Alonso JL, Peña I, Sanz ME, Vaquero V, Mata S, Cabezas C, López JC (2013) *Chem Commun* 49:3443
102. Saenger W (1979) *Nature* 279:343
103. Jeffrey GA (1997) *An introduction to hydrogen bonding*. Oxford University Press, Oxford
104. Jeffrey GA, Saenger W (1991) *Hydrogen bonding in biological structures*. Springer, New York
105. Blanco S, Lesarri A, López JC, Alonso JL (2004) *J Am Chem Soc* 126:11675
106. Kraitichman J (1953) *Am J Phys* 21:17
107. Van Eijck BP (1982) *J Mol Spectrosc* 91:348
108. Ruoff RS, Klots TD, Emilsson T, Gutowsky HS (1990) *J Chem Phys* 93:3142
109. Florio GM, Christie RA, Jordan KD, Zwier T (2002) *J Am Chem Soc* 124:10236
110. Zwier T (2006) *J Phys Chem A* 110:4133
111. Sanz ME, Lopez JC, Alonso JL (2002) *Chem Eur J* 8:4265
112. Jimenez AI, Vaquero V, Cabezas C, Lopez JC, Cativiela C, Alonso JL (2011) The singular gas-phase structure of 1-aminocyclopropanecarboxylic acid (Ac(3)c). *J Am Chem Soc* 133:1062
113. de Meijere A (1979) *Angew Chem Int Ed Engl* 18:809–826
114. Rademacher P (2003) *Chem Rev* 103:933–975
115. Lesarri A, Cocinero EJ, López JC, Alonso JL (2004) *Angew Chem Int Ed* 43:605
116. Lesarri A, Sánchez R, Cocinero EJ, López JC, Alonso JL (2005) *J Am Chem Soc* 127:12952
117. Cocinero EJ, Lesarri A, Grabow J-U, López JC, Alonso JL (2007) *ChemPhysChem* 8:599
118. Ramachandran GN, Reddi AH (eds) (1976) *Biochemistry of collagen*. Plenum, New York
119. Nimmi ME (ed) (1988) *Collagen*, vol 1–4. CRC, Boca Raton
120. Adams E, Frank L (1980) *Annu Rev Biochem* 49:1005
121. Ramachandran GN, Kartha G (1955) *Nature* 176:593
122. Rich A, Crick FHC (1961) *J Mol Biol* 3:483
123. Fraser RD, McRae TP, Suzuki E (1979) *J Mol Biol* 129:463
124. Beck K, Brodsky J (1998) *Struct Biol* 122:17
125. Rosenbloom J, Prockop DJ (1971) *J Biol Chem* 246:1549
126. Inouye K, Sakakibara S, Prockop D (1976) *J Biochim Biophys Acta* 420:133
127. Czinki E, Császár AG (2003) *Chem Eur J* 9:1008
128. Mata S, Vaquero V, Cabezas C, Peña I, Pérez C, López JC, Alonso JL (2009) Observation of two new conformers of neutral proline. *Phys Chem Chem Phys* 11:4141
129. Lesarri A, Cocinero EJ, López JC, Alonso JL (2005) The shape of 4(S)- and 4(R)-hydroxyproline in gas phase. *J Am Chem Soc* 127:2572
130. Bürgi HB, Dunitz JD, Shefter E (1973) *J Am Chem Soc* 95:5065
131. Bürgi HB, Dunitz JD, Lehn JM, Wipff G (1974) *Tetrahedron* 30:1563
132. Deslongchamps P (1983) *Stereoelectronic effects in organic chemistry*. Pergamon, New York
133. Tatcher GRJ (1993) *The anomeric effect and associated stereoelectronic effects*. American Chemical Society, Washington, DC
134. Sanz ME, Blanco S, López JC, Alonso JL (2008) Rotational probes of six conformers of neutral cysteine. *Angew Chem Int Ed* 47:6216
135. Sanz ME, Lopez JC, Alonso JL (2010) Six conformers of neutral aspartic acid identified in the gas phase. *Phys Chem Chem Phys* 12:3573

136. Cabezas C, Varela M, Peña I, Mata S, López JC, Alonso JL (2012) The conformational locking of asparagine. *RSC Chem Commun* 48:5934
137. Snoek LC, Robertson EG, Kroemer RT, Simons JP (2000) Conformational landscapes in amino acids: infrared and ultraviolet ion-dip spectroscopy of phenylalanine in the gas phase. *Chem Phys Lett* 321:49
138. Lee KT, Sung J, Lee KJ, Kim SK, Park YD (2002) *J Chem Phys* 116:8251
139. Lee KT, Sung J, Lee KJ, Park YD, Kim SK (2002) *Angew Chem Int Ed* 41:4114
140. Lee KT, Sung J, Lee KJ, Kim SK, Park YD (2003) *Chem Phys Lett* 368:262
141. Lee YH, Jung J, Kim B, Butz P, Snoek LC, Kroemer RT, Simons JP (2004) *J Phys Chem A* 108:69
142. Lindinger A, Toennies JP, Vilesov AF (1999) *J Chem Phys* 110:3
143. Cohen R, Brauer B, Nir E, Grace L, de Vries MS (2000) *J Phys Chem A* 104:6351
144. Grace LI, Cohen R, Dunn TM, Lubman DM, de Vries MS (2002) *J Mol Spectrosc* 215:204
145. Kaczor A, Reva ID, Proniewicz LM, Fausto R (2006) *J Phys Chem A* 110:2360
146. von Helden G, Compagnon I, Blom MN, Frankowski M, Erlekam U, Oomens J, Brauer B, Gerbercd RB, Meijera G (2008) *Phys Chem Chem Phys* 10:1248
147. Huang Z, Yu W, Lin Z (2006) *J Mol Struct THEOCHEM* 758:195
148. Inokuchi Y, Kobayashi Y, Ito T, Ebata T (2007) *J Phys Chem A* 111:3209
149. Rizzo TR, Park YD, Peteanu LA, Levy DH (1986) *J Chem Phys* 84:2534
150. Snoek LC, Kroener RT, Hockridge MR, Simons JP (2001) *Phys Chem Chem Phys* 3:1819
151. Bakker JM, Aleese LM, Meijer G, von Helden G (2003) *Phys Rev Lett* 91:203003
152. Sanz ME, Cortijo V, Caminati W, Lopez JC, Alonso JL (2006) The conformers of phenylglycine. *Chem Eur J* 12:2564
153. Pérez C, Mata S, Blanco S, López JC, Alonso JL (2011) *J Phys Chem A* 115:9653
154. Wang L, Schultz PG (2005) *Angew Chem Int Ed* 44:34
155. McGlone SJ, Godfrey PD (1995) *J Am Chem Soc* 117:1043
156. Sanz ME, Lesarri A, Peña I, Vaquero V, Cortijo V, López JC, Alonso JL (2006) The shape of beta-alanine. *J Am Chem Soc* 128:3812
157. Blanco S, López JC, Mata S, Alonso JL (2010) The conformations of γ -aminobutyric acid (GABA): the role of $n \rightarrow \pi^*$ interaction. *Angew Chem Int Ed* 49:9187
158. Choquet D, Korn H (1988) *Neurosci Lett* 84:329
159. Wu FS, Gibbs TT, Farb DH (1993) *Eur J Pharmacol* 246:239
160. Horikoshi T, Asanuma A, Yanagisawa K, Anzai K, Goto S (1988) *Mol Brain Res* 4:97
161. Cocinero EJ, Lesarri A, Sanz ME, López JC, Alonso JL (2006) Conformations of β -aminobutyric acid in the gas phase. *ChemPhysChem* 7:1481
162. Cocinero EJ, Villanueva P, Lesarri A, Sanz ME, Blanco S, Mata S, López JC, Alonso JL (2007) The shape of neutral sarcosine in gas phase. *Chem Phys Lett* 435:336
163. Lesarri A, Cocinero EJ, López JC, Alonso JL (2005) Gas-phase structure of N, N-dimethylglycine. *ChemPhysChem* 6:1559
164. Cortijo V, Sanz ME, López JC, Alonso JL (2009) *J Phys Chem A* 113:14681–14683
165. Watson D, Crick FH (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 1953(17):737–738
166. Vaquero V, Sanz ME, López JC, Alonso JL (2007) The structure of uracil: a laser ablation rotational study. *J Phys Chem A* 111:3443
167. López JC, Peña I, Sanz ME, Alonso JL (2007) Probing thymine with laser ablation molecular beam Fourier transform microwave spectroscopy. *J Chem Phys* 126:191103
168. Alonso JL, Peña I, López JC, Vaquero V (2009) Rotational spectral signatures of four tautomers of guanine. *Angew Chem Int Ed* 48:6141
169. Alonso JL, Vaquero V, Peña I, López JC, Mata S, Caminati W (2013) All five forms of cytosine revealed in the gas phase. *Angew Chem* 125:3287
170. López JC, Alonso JL, Peña I, Vaquero V (2010) Hydrogen bonding and structure of uracil–water and thymine–water complexes. *Phys Chem Chem Phys* 12:14128
171. Brown RD, Godfrey PD, McNaughton D, Pierlot AP (1988) *J Am Chem Soc* 110:2329

172. Brown RD, Godfrey PD, McNaughton D, Pierlot AP (1989) *J Chem Soc Chem Commun* 37–38
173. Puzzarini C, Barone V (2011) *Phys Chem Chem Phys* 13:7158
174. Kabeláč M, Hobza P (2008) *Phys Chem Chem Phys* 9:903
175. Yoshikawa A, Matsika S (2008) *Chem Phys* 347:393
176. Wang FF, Zhao DX, Yang ZZ (2009) *Chem Phys* 360:141
177. Casaes RN, Paul JB, McLaughlin RP, Saykally RJ, van Mourik T (2004) *J Phys Chem A* 108:10989
178. Graindourze M, Smets J, Zeegers-Huyskens T, Maes G (1990) *J Mol Struct* 222:345
179. Choi MY, Miller RE (2005) *Phys Chem Chem Phys* 7:3565
180. Blanco S, López JC, Lesarri A, Alonso JL (2006) *J Am Chem Soc* 128:1211
181. Lovas FJ, Suenram RD, Fraser GT, Gillies CW, Zozom JJ (1988) *Chem Phys* 88:722
182. Fraser GT, Suenram RD, Lovas FJ (1988) *J Mol Struct* 189:165
183. Caminati W, López JC, Blanco S, Mata S, Alonso JL (2010) How water links to cis and trans peptidic groups: the rotational spectrum of N-methylformamide–water. *Phys Chem Chem Phys* 12:10230
184. Maris A, Ottaviani P, Caminati W (2002) *Chem Phys Lett* 360:155
185. Hanus M, Ryjacek F, Kubar T, Bogdan TV, Trygubenko SA, Hobza P (2003) *J Am Chem Soc* 125:7678
186. Elshakre M (2005) *Int J Quant Chem* 104:1–15
187. Chen H, Li S (2006) *J Phys Chem A* 110:12360
188. Nir E, Janzen CH, Imhof P, Kleinermanns K, de Vries MS (2001) *J Chem Phys* 115:4604
189. Mons M, Dimicoli I, Piuze F, Tardivel B, Elhanine M (2002) *J Phys Chem A* 106:5088
190. Seefeld K, Brause R, Häber T, Kleinermanns K (2007) *J Phys Chem A* 111:6217
191. Nir E, Grace L, Brauer B, de Vries MS (1999) *J Am Chem Soc* 121:4896
192. Chin W, Mons M, Piuze F, Tardivel B, Gorb IL, Leszczynski J (2004) *J Phys Chem A* 108:8237
193. Mons M, Piuze F, Dimicoli I, Gorb L, Leszczynski J (2006) *J Phys Chem A* 110:10921
194. Choi MY, Miller RE (2006) *J Am Chem Soc* 128:7320
195. Gahlmann A, Park ST, Zewail AH (2009) *J Am Chem Soc* 131:2806
196. Szczesniak M, Szczepaniak K, Kwiatkowski JS, KuBulat K, Person WB (1988) *J Am Chem Soc* 110:8319
197. Dong F, Miller RE (2002) *Science* 298:1227
198. Nir E, Müller M, Grace LI, de Vries MS (2002) *Chem Phys Lett* 355:59
199. Shorokhov D (2000). Ph.D. dissertation, University of Oslo
200. Feyer V, Plekan O, Richter R, Coreno M, de Simone M, Prince KC, Trofimov AB, Zaytseva IL, Schirmer J (2010) *J Phys Chem A* 114:10270
201. Lapinski L, Nowak MJ, Reva I, Rostkowska H, Fausto R (2010) *Phys Chem Chem Phys* 12:9615–9618
202. Lapinski L, Reva I, Nowak MJ, Fausto R (2011) *Phys Chem Chem Phys* 13:9676–9684
203. Brown RD, Godfrey PD, McNaughton D, Pierlot AP (1989) *J Am Chem Soc* 111:2308
204. Colins P, Ferrier R (1995) *Monosaccharides: their chemistry and their roles in natural products*. Wiley, New York
205. Pigman W, Horton D (1972) *The carbohydrates: chemistry and biochemistry*. Academic, New York
206. Herbst E, van Dishoeck EF (2009) *Annu Rev Astron Astrophys* 47:427
207. Hollis JM, Lovas FJ, Jewell PR (2000) *Astrophys J* 540:L107
208. Hollis JM, Vogel SN, Snyder LE, Jewell PR, Lovas FJ (2001) *Astrophys J* 554:L81
209. Lovas FJ, Suenram RD, Plusquellic DF, Møllendal H (2003) *J Mol Spectrosc* 222:263
210. Angyal SJ (1969) *Angew Chem Int Ed* 8:157–226
211. Sponsler OL, Dore WH (1931) *J Am Chem Soc* 53:1639
212. Mackie W, Perlin AS (1966) *Can J Chem* 44:2039
213. Perlin AS (1966) *Can J Chem* 44:539

214. McDonald TRR, Beevers CA (1952) *Acta Crystallogr* 5:654
215. Ferrier WG (1960) *Acta Crystallogr* 13:678
216. Ferrier WG (1963) *Acta Crystallogr* 16:1023
217. Brown GM, Levy HA (1965) *Science* 147:1038
218. (2002) See for example the special issue of *Chem. Rev.* (issue 102, 2002) dedicated to glycosylation
219. Simons JP, Jockusch RA, Carçabal P, Hung I, Kroemer RT, Mcleod NA, Snoeck LC (2005) *Int Rev Phys Chem* 24:489–531
220. Simons JP (2009) *Mol Phys* 107:2435–2458
221. Cocinero EJ, Carçabal P (2014) *Carbohydrates*. *Top Curr Chem*. doi:[10.1007/128_2014_596](https://doi.org/10.1007/128_2014_596)
222. Cabezas C, Peña I, Daly AM, Alonso JL (2013) *Chem Commun* 49:10826
223. Peña I, Mata S, Martín A, Cabezas C, Daly AM, Alonso JL (2013) *Phys Chem Chem Phys* 15:18243
224. Peña I, Cocinero EJ, Cabezas C, Lesarri A, Mata S, Ecija P, Daly AM, Cimas A, Bermúdez C, Basterretxea FJ, Blanco S, Fernández JA, López JC, Castaño F, Alonso JL (2013) *Angew Chem Int Ed* 52:11840
225. Alonso JL, Lozoya MA, Peña I, López JC, Cabezas C, Mata S, Blanco S (2014) *Chem Sci* 5 (2):515
226. Bermúdez C, Peña I, Cabezas C, Daly AM, Alonso JL (2013) *ChemPhysChem* 14:893–895
227. Cocinero EJ, Lesarri A, Écija P, Cimas A, Davis BJ, Basterretxea FJ, Fernández JA, Castaño F (2013) *J Am Chem Soc* 135:2845
228. Risley JM, Van Etten RL (1982) *Biochemistry* 21:6360
229. Serianni AS, Pierce J, Huang SG, Barker R (1982) *J Am Chem Soc* 104:4037
230. Azofra LM, Alkorta I, Elguero J, Popelier PLA (2012) *Carbohydr Res* 358:96
231. Watson JD, Crick FHC (1953) *Nature* 171:737–738
232. Saenger W (1984) *Principles of nucleic acid structure*. Springer, New York, pp 1–556
233. Rudrum M, Shaw DF (1965) *J Chem Soc* 52–57
234. Lemieux RU, Stevens JD (1966) *Can J Chem* 44:249
235. Cortes SJ, Mega TL, Van Etten RL (1991) *J Org Chem* 56:943–947
236. Furberg S (1960) *Acta Chem Scand* 14:1357–1363
237. Breitmaier E, Hollstein U (1976) *Org Magn Reson* 8:573
238. Nikolaenko TY, Bulavin LA, Govorun DN (2011) *J Appl Spectrosc* 78:751–754
239. Furberg S, Hordvik A, Taulbol K (1956) *Acta Chem Scand* 10:135
240. Sisak D, McCusker LB, Zandomenighi G, Meier BH, Bläser D, Boese R, Schweizer WB, Gylmour R, Dunitz JD (2010) *Angew Chem* 122:4605
241. Sisak D, McCusker LB, Zandomenighi G, Meier BH, Bläser D, Boese R, Schweizer WB, Gylmour R, Dunitz JD (2010) *Angew Chem Int Ed* 49:4503
242. Saenger W (2010) *Angew Chem* 122:6633
243. Saenger W (2010) *Angew Chem Int Ed* 49:6487
244. Lemieux RU (1990) *Explorations with sugars: how sweet it was*. American Chemical Society, Washington, DC
245. Collins P, Ferrier R (1995) *Monosaccharides: their chemistry and their roles in natural products*. Wiley, New York
246. Lehninger AL (2004) *Principles of biochemistry*, 4th edn. W. H. Freeman, New York
247. Sinnott ML (2007) *Carbohydrate chemistry and biochemistry*. RSC, Cambridge
248. Juaristi E, Cuevas G (1992) *Tetrahedron* 48:5019–5087
249. Perrin CL, Armstrong KB, Fabian MA (1994) *J Am Chem Soc* 116:715–722
250. Cramer CJ, Truhlar DG (1993) *J Am Chem Soc* 115:5745–5753
251. Corchado JC, Sánchez ML, Aguilar MA (2004) *J Am Chem Soc* 126:7311–7319
252. Wladkowski BD, Chenoweth SA, Jones KE, Brown JW, Phys J (1998) *J Phys Chem A* 102:5086–5092
253. Barrows SE, Storer JW, Cramer CJ, French AD, Truhlar DG (1998) *J Comput Chem* 19:1111

254. Hough E, Neidle S, Rogers D, Troughton PGH (1973) *Acta Crystallogr Sect B: Struct Crystallogr Cryst Chem* 29:365
255. Wolfe S (1972) *Acc Chem Res* 5:102–111
256. Kirschner KN, Woods RJ (2001) *Proc Natl Acad Sci U S A* 98:10541–10545
257. Mihura N, Taniguchi T, Monde K, Nishimura SI (2006) *Chem Phys Lett* 419:326
258. Hoffmann M, Rychelewski J (2001) *J Am Chem Soc* 123:2308–2316
259. Kovács A, Ivanov AY (2009) *J Phys Chem B* 113:2151
260. Çarçabal P, Jockusch RA, Hünig I, Snoek LC, Kroemer RT, Davis BG, Gamblin DP, Compagnon I, Oomens J, Simons JP (2005) *J Am Chem Soc* 127:11414–11425
261. Çarçabal P, Cocinero EJ, Simons JP (2013) *Chem Sci* 4:1830–1836
262. Ohru H, Nishida Y, Watanabe M, Hori H, Meguro H (1985) *Tetrahedron Lett* 26:3251–3254
263. Nishida Y, Ohru H, Meguro H (1984) *Tetrahedron Lett* 25:575–1578
264. Bock K, Duus JØ (1994) *J Carbohydr Chem* 13:513–543
265. Marchessault RH, Pérez S (1979) *Biopolymers* 18:2369–2374
266. Hordvik A (1971) *Acta Chem Scand* 25:2175
267. Guler LP, Yu Y-Q, Kenttämäa HI (2002) *J Phys Chem A* 106:6754
268. Jahn MK, Dewald DA, Wachsmuth D, Grabow J-U, Mehrotra SC (2012) *J Mol Spectrosc* 280:54
269. Finch P (1999) *Carbohydrates: structures, syntheses and dynamics*. Kluwer, The Netherlands
270. Hünig I, Painter AJ, Jockusch RA, Çarçabal P, Marzluff EM, Snoek LC, Gamblin DP, Davis BG, Simons JP (2005) *Phys Chem Chem Phys* 7:2474–2480
271. Doddrell D, Allerhan A (1971) *J Am Chem Soc* 93:2779–2781
272. Kanters JA, Roelofsen G, Alblas BP, Meinders I (1977) *Acta Crystallogr B* 33:665–672
273. Takagi S, Jeffrey GA (1977) *Acta Crystallogr B* 33:3510–3515
274. Shallenberger RS, Acree TE (1967) *Nature* 216:480–482
275. Kier LB (1972) *J Pharm Sci* 61:1394–1397
276. van der Heijden A (1997) *Pure Appl Chem* 69:667–674
277. Woods RJ, Szarek WA, Smith VH (1990) *J Am Chem Soc* 112:4732–4741
278. Shallenberger RS (1978) *Pure Appl Chem* 50:1409–1420
279. Deutsch EW, Hansch C (1966) *Nature* 211:75