J. Indian Inst. Sci.
A Multidisciplinary Reviews Journal
ISSN: 0970-4140 Coden-JIISAD

REVIEW ARTICLE

Exploring Hydrogen Bond in Biological Molecules

José A. Fernández[*]

Abstract | Life makes extensive use of non-covalent interactions, as they are a convenient way to build complex structures that can be assembled or disassembled quickly, with a minimum energy consumption. Among the inter-molecular interactions, hydrogen bond plays a central role, and it is the main responsible of the structure of proteins, DNA, and several other superstructures in the cell. Characterization of hydrogen bond in biologic environment is not an easy task, and several complex and imaginative techniques have been developed to circumvent the technical challenges of such studies. We present here an overview of the field of mass-resolved laser spectroscopy applied to nucleobases, peptides, and monosaccharides to demonstrate that despite the different environment the molecules encounter in the jet, such experiments yield important structural information that helps understanding the role played by hydrogen bond in biology.

1 Introduction

Life is a complex and fascinating phenomenon, in part due to the exotic combination of factors required for its appearance. It is usually taken as a complex combination of chemical reactions in delicate equilibrium that demand very strict conditions. In this model, even small alterations of the environment necessarily cause life disappearance. However, life has shown to have exceptional resilience. Living organisms have demonstrated to be able to model the environment at planetary scale. Certainly, it is well known that originally, our planet had a reductive atmosphere and that the appearance in the Neoarchean era of photosynthetic organisms that released oxygen as a by-product of their metabolism changed that primitive atmosphere[1] and most important, life was able to survive such dramatic change, adapted to it, and finally expanded.

Part of the plasticity of life comes from its use of inter-molecular interactions:[2, 3] sticky forces that come into play whenever two (or more) atoms, molecules, or a combination of them approaches one each other. The electronic clouds of the interacting entities dance a complex choreography known as "electron dynamic correlation" that results in the appearance of attractive forces.[4] Although the module of these forces, usually divided into van der Waals and hydrogen bond,[5] is small, they present interesting characteristics, such as cooperativity, that reinforce their importance.[6] Certainly, inter-molecular interactions shape the environment around us and perhaps the similarity between their module and kT is one of the reasons that helped the appearance of life. For example, the difference in interaction energy between ammonia and water molecules is the reason why ammonia is a gas at room temperature, while water is a liquid. Furthermore, cooling water below 0 °C is enough to transform it into a solid able to sink a ship. Indeed, the strong propensity of water towards formation of hydrogen bonds makes it a fascinating element and it is probably one of the key factors behind the existence of life on earth.

Another example is the barrier that isolates (and protects) cells and bacteria from the environment, the well-known lipid membrane. This barrier is mainly composed of amphiphilic molecules called phospholipids, which are glycerol derivatives with a polar head group and two aliphatic chains; in contact with water, the

[1] *Department of Physical Chemistry, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Barrio Sarriena s/n, 48940 Leioa, Spain.*
**josea.fernandez@ehu.es*

phospholipids auto-organize to hide the lipophilic side and to expose the (hydrophilic) head group to the surrounding water.[7] Thus, the lipid bilayer is not composed of covalently bonded molecules but by an assembly of molecules, held together by non-covalent interactions. This gives the lipid membrane a fluid consistency and the cells need to use cholesterol to create domains of increased rigidity in the membrane, where membrane proteins can anchor and maintain a stable conformation. These domains are known as lipid rafts and move around the lipid membrane, like "rafting" the interphase.[8]

The design of proteins is also based on favorable/unfavorable water–amino acid (AA) and AA–AA interactions. Certainly, among the collection of amino acids (20 in the case of human cells), some present OH, NH or a combination of both groups, SH, or even charged groups, which enable a favorable interaction with water. Conversely, amino acids with aliphatic side chains or containing aromatic groups will present unfavorable interactions with water. A smart combination of the amino acids induces the protein to fold in such a way that the exposure of the hydrophobic amino acids is minimized, while formation of intramolecular hydrogen bonds (H bonds) is optimized.[9]

Formation of H bonds is also an essential aspect in DNA: the two strands of DNA are held together mainly by the formation of hydrogen bonds between partner bases (cytosine with guanine and thymine with adenine, C-G and A-T). The whole machinery of DNA storage in the form of chromatin and the translation of the genome into proteins is also based on very specific interactions between nucleobases and proteins. The protein machinery that enables information storage and retrieval has an incredible complexity and involves around ~ 5–7% of the human coding genome.[9] The whole process is regulated by a subtle interplay of protein–DNA interactions in which formation of H bonds plays a dominant role. Actually, gene expression may be promoted/silenced by methylation: a process that blocks formation of H bonds in key places. This mechanism is so relevant that a new discipline, the Epigenetics, has appeared devoted to understand its principles.

All the examples above highlight the importance of having a deep knowledge of the non-covalent forces and specially of hydrogen bond. However, its intrinsic weak nature makes construction of accurate models a difficult task.[2] Despite their importance in biological environments, the stability added by these interactions to, for example, a protein is a small fraction of the total stability of the protein, and very often, it is of the order of the calculation error.[10] Thus, very accurate experimental data are required to adjust the theoretical models.[11]

One of the main sources of data is the spectroscopy in supersonic expansions,[10, 12–14] as the extensive literature published in the last decades demonstrate.[2] The expansion is usually created using a pulsed valve that releases a short (microseconds) pulse of gas into a vacuum chamber. The gas is seeded with the molecules to be explored and maintained at pressures between 1 and 50 bar. Thus, when the valve opens, an adiabatic expansion is created, transforming the ro-vibrational energy of the molecules into translational energy. Cooling of the rotational energy is substantially more efficient, leaving the molecules at ~ 3–5 K. Depending on the molecules, the vibrational temperature is reduced to ~ 50–100 K: the very energetic vibrational levels of the small molecules will require of hard, direct collisions with the buffer gas to transfer their vibrational energy to translational energy. On the other hand, very large systems are difficult to cool. Also, when laser desorption systems are used, the geometry of the nozzle is usually less efficient, producing hotter expansions. In any case, the temperatures reached enable molecules to aggregate, forming clusters of sizes containing up to dozens of molecules. Once the aggregates leave the collisions region, they travel isolated from the environment forming a dense molecular beam. As it can be seen, the expansion has three advantages: it cools the molecules simplifying their spectroscopy, enables formation of molecular aggregates, and enables achieving a relatively high density of species in a confined space that can be probed using a combination of spectroscopic techniques. An excellent review on the subject may be found in Ref. [15].

We will arbitrarily divide here the techniques available into two categories: with and without mass selectivity, due to the extreme advantage that the use of mass spectrometers for the discrimination between species introduces in the process of obtaining physical observables of the aggregates.[17] Certainly, in the case of electronic spectroscopy, formation of molecular aggregates results in a modest perturbation of the electronic transition, and therefore, they present overlapping absorption spectra. Transition from a traditional laser-induced fluorescence (LIF) experiment to a resonance-enhanced multiphoton ionization (REMPI) technique enables segregation of the spectra of each species in

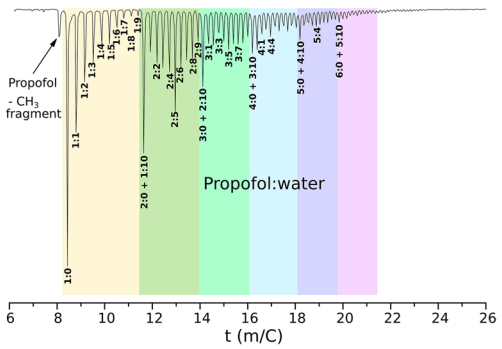

Figure 1: Mass spectrum obtained from a supersonic expansion of propofol and water in He. Each species appears in a different mass channel, enabling exploration of their spectroscopy without interferences Adapted from Ref.[16].

a different mass channel (Fig. 1). Nevertheless, use of mass selection is not that relevant in the high-resolution techniques, such as high-resolution LIF[18] or microwave (MW) spectroscopy.[19] Especially in the latter, the high specificity of the technique enables the efficient identification of the spectra of multiple species in the beam. A beautiful example of the power of MW spectroscopy may be found in the study of water aggregates,[20] phenol[21] and aniline homotrimers[22] or difluoromethane–water.[23]

Characterization of the molecular aggregates usually demands the use of several very elaborated spectroscopic techniques. Even very small aggregates very often present a collection of conformational isomer (abbreviated as conformers or isomers), consequence of the different ways in which the molecules can interact, leading to formation of aggregates of similar stability (Fig. 2). The most popular technique to identify the number of conformers of a given system is the so-called UV/UV double resonance or "hole burning" (Fig. 3). This technique isolates the contribution from each species to the excitation spectrum of a given aggregate. Once isolated, the IR/UV technique enables extraction of the IR spectrum of each conformer, yielding important structural information.[25] Certainly, the position of the IR bands is very sensitive to the environment of the chemical moiety. Thus, formation of a hydrogen bond usually results in a shift in the position of the corresponding stretching vibration. This shift is usually proportional to the strength of the hydrogen bond formed. Additional techniques, such as determination of the ionization energy thresholds[26] or dispersed fluorescence spectroscopy,[27] enable extraction of physical observables that will conclude with an accurate identification of the structure of the isomers of a given aggregate, by comparison with the in silico simulations and predictions.

With these experimental and computational tools, an impressive number of systems have been tackled in the last decades (see, for example, references[29–57]). We will give in the following several examples, highlighting the influence of the hydrogen bond in the structure of biological molecules.

1.1 DNA and DNA Bases

One of the most intriguing questions in biology is why nature chose CGAT (cytosine, guanine, adenine, and thymine) as the alphabet of life. Some authors speculate on the idea that in the primal earth where life first appeared, the molecules were exposed to strong VUV radiation, and therefore, only those species resistant to solar radiation were able to survive and form the first molecules with auto-replication ability.[58–60] Certainly, all DNA bases present very short excited state lifetimes, which allow the molecules to dissipate efficiently the electronic excitation and transform it into vibrational (thermal) energy. However, DNA also contains sugar units and phosphate groups, which can also be ionized by VUV radiation. Several recent studies deal with the photodamage induced in the deoxyribose and other sugars, and demonstrate that UV radiation easily induces their dissociation.[60] In addition, other works demonstrate that new deactivation channels open when two nucleobases establish stacking interactions[61, 62] or that interaction with water also modifies the excited state dynamics, increasing the lifetimes.[63] Therefore, it is not clear that VUV radiation alone may have played such a determinant role.

Another possibility is that CGAT were the best candidates to build a molecule to store information, because their ability to interact with many molecules at the same time. The special combination of functional groups in DNA bases gives them the ability to interact preferentially with their complementary base, but at the same time, to form stacking interactions to build the biopolymer. Still, they present additional functional groups to form hydrogen bonds with other molecules. Thus, DNA bases contain a combination of CO/and NH groups that confer them a marked preference to pair with their complementary base (C-G/A-T). However, in addition, they also allow the nucleobases to be easily recognized through non-covalent interactions, enabling the DNA/

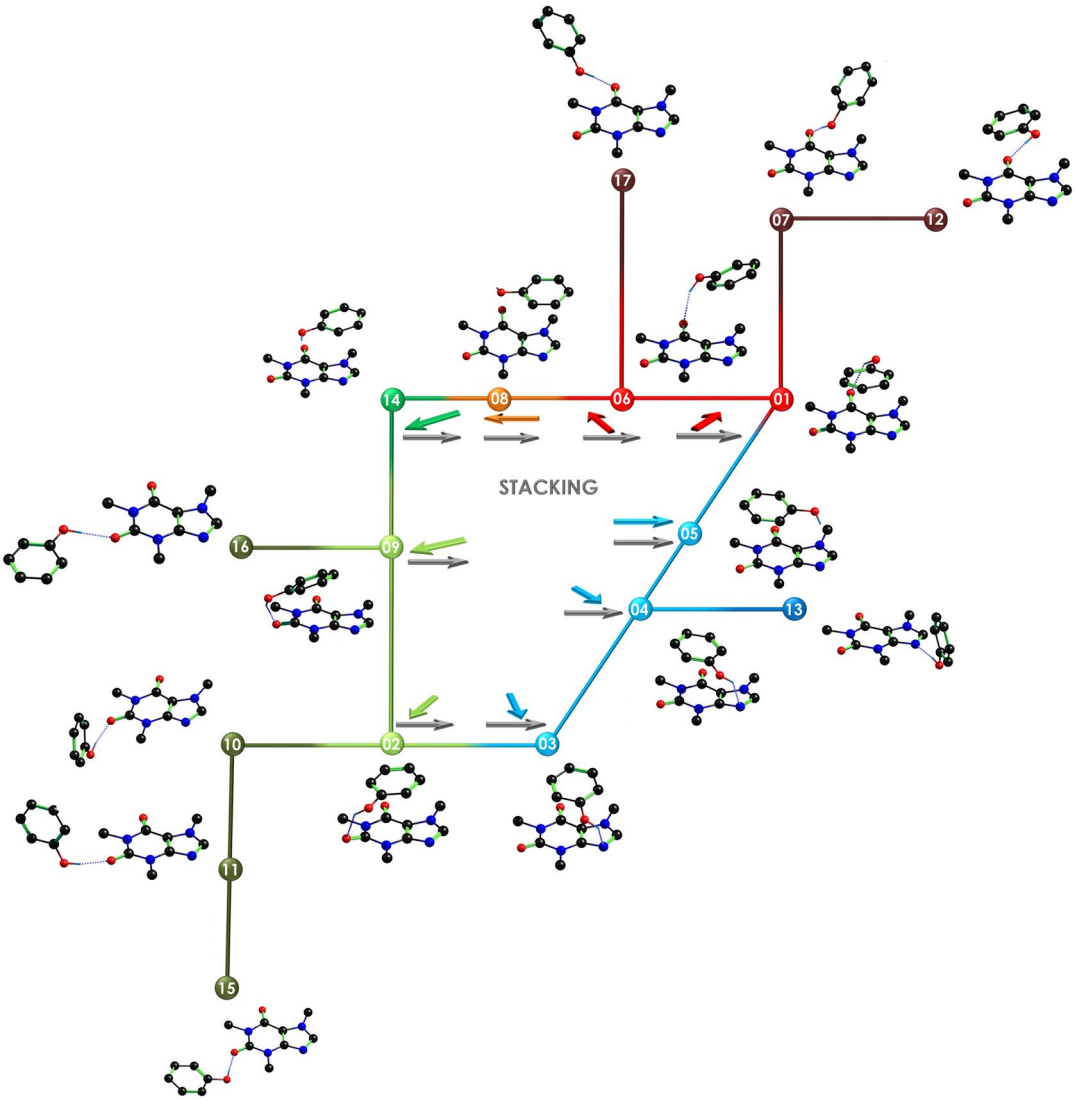

Figure 2: Schematic representation of the conformational landscape for the interaction between phenol and caffeine. The conformers are organized depending on the relative orientation of the two molecules and the lines represent connections between the structures. Hydrogen atoms were omitted except that of the hydroxyl group of phenol for the shake of clarity Adapted from Ref.[24].

RNA strand to be "read" and the information that it contains interpreted. Furthermore, such combination permits their simultaneous interaction with several proteins. This is essential, as DNA is usually stored wrapped around disk-shaped protein ensembles called histones.[64] Unfolding DNA is a complex process that involves interaction with several highly specialized proteins that bind to the DNA strand with a high affinity and pull from it, liberating and leaving it ready to couple with the transcriptional machinery. Methylation of NH sites in the nucleobases blocks the grips that such proteins use, silencing the gene. These so-called "epigenetic marks" give the cell a dynamic mechanism to choose when and what genes to express.[65]

All the above highlights the importance of having a good knowledge of the aggregation preferences of nucleobases and has motivated publication of a large number of experimental and computational works.[31, 59, 61, 63, 66–74]

Gas-phase spectroscopic studies on DNA bases pairing are not easy. First, their vapor pressure is low, and therefore, they are not transferred efficiently into the gas phase by simple warming. Therefore, more sophisticated desorption systems have to be used. Among them, the most popular is laser desorption. Second, DNA bases present several tautomers. In the biological environment,

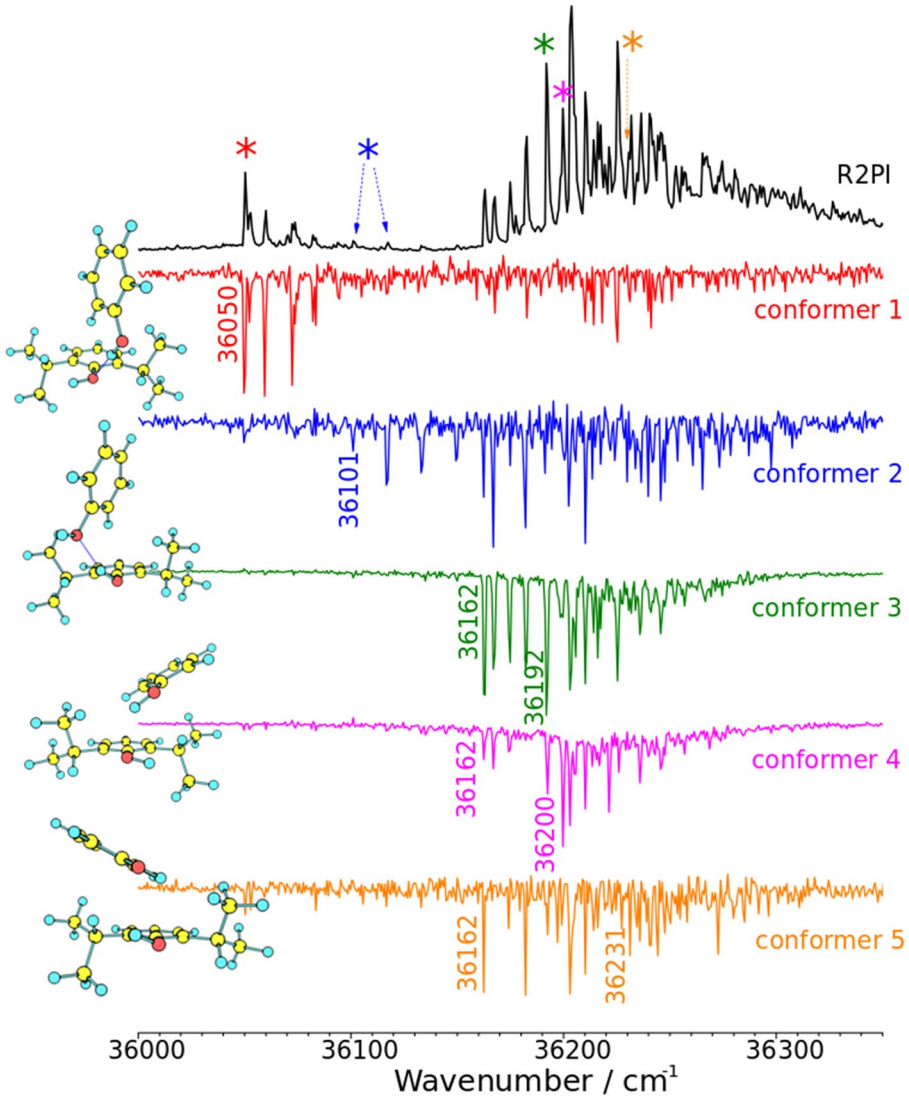

Figure 3: Comparison between the two-color REMPI spectrum of the propofol···phenol aggregate and the hole burning traces obtained probing different transitions of the REMPI spectrum. The tentative assignments included with the hole burning traces demonstrate the sensitivity of the technique that enables discrimination between conformers differing in the relative position of the two molecules Adapted from Ref. [28].

water preferentially stabilizes the ones that usually appear in the text books, which are not necessarily those detected in jets.[31, 61, 71, 73, 75–80] Especially, when laser desorption sources are used, the large amounts of energy deposited in the sample by the desorption laser opens isomerization paths, leading to tautomeric species that are further stabilized by the collisions during the expansion. As an example, Fig. 4 shows the tautomers of guanine, with the most stable forms in the center of the figure: the 9-keto and 7-keto tautomers. The small structural difference between them, the position of the hydrogen atom shaded in orange, is enough to produce a measurable difference in the IR spectrum. From these keto forms, the enol tautomers, the next most stable structures, are obtained by moving a single proton. This modification is also clearly reflected in the spectrum. From a spectroscopic point of view, the two possible orientations of the hydroxyl hydrogen atom, highlighted with a double arrow in Fig. 4, also produce two different and relevant species, as they present a different landscape of inter-molecular interactions.

The same happens with the imine-keto tautomers, formed by transferring a proton from the NH_2 moiety into the N4. Although these species are relatively high in energy, they are also detected

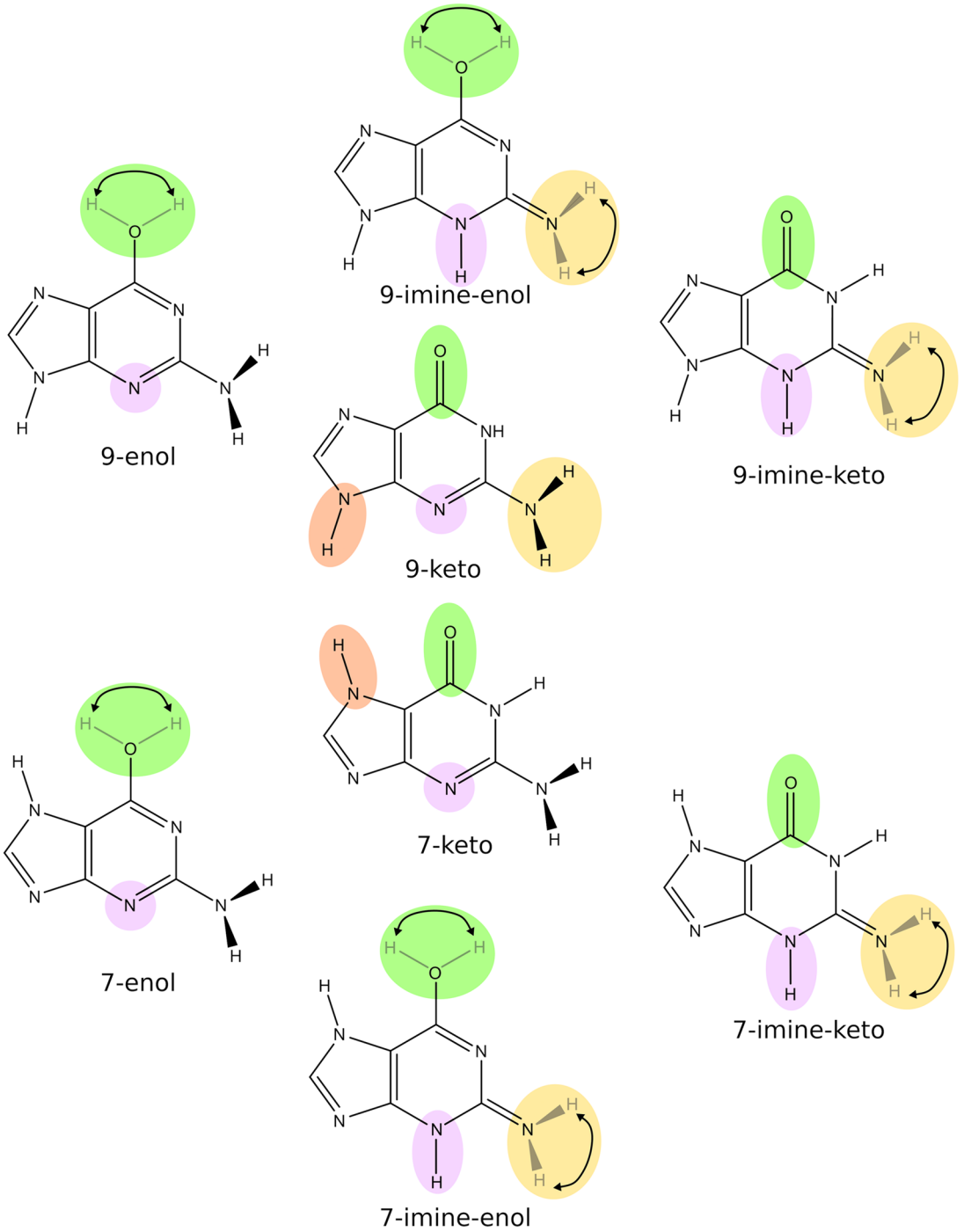

Figure 4: Tautomers of guanine. The two structures in the center of the picture correspond to the two most stable tautomers. The colors highlight the protons that must be moved between chemical groups to form a different tautomer. The double arrow highlights that some of the tautomers have two different isomers, which differ in the orientation of a single hydrogen atom. Some tautomers, such as 9-imine-enol, present four conformational isomers.

in supersonic expansions. Finally, each of the two imine-enol tautomers has four conformational isomers, depending on the relative orientation of the NH and OH groups. The presence of so many species of a given nucleobase strongly complicates the spectroscopic studies, but it has the advantage of providing structural information on a larger portion of the potential energy surface. This is particularly interesting for the search of life in exoplanets. The 9-keto is the most stable form of

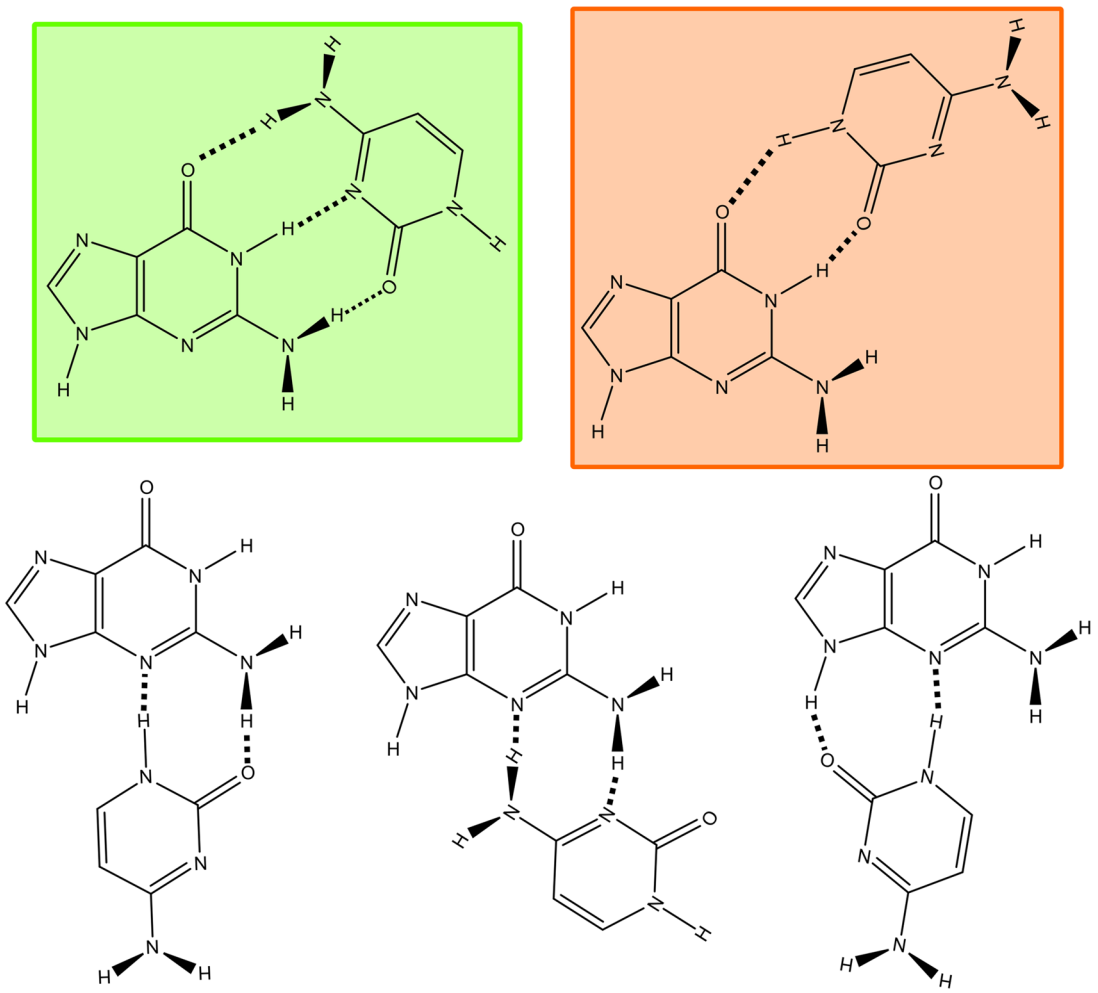

Figure 5: The Watson–Crick pairing (green rectangle) of guanine–cytosine is neither the only possible interaction structure nor the most stable in gas phase (red rectangle). Other three possible pairings are also shown.

guanine under biological conditions, as we know them in this planet. However, life may appear (or even has appeared) in other worlds under very different conditions. For example, in colder environments, oceans of liquid ammonia may exist. Such extremely alkaline conditions will favor different tautomers of the same nucleobases, in case that life still uses such molecules as biological building blocks.

Aggregation under jet conditions is also different from real biological environments, but still, important information may be extracted using laser spectroscopy and computational chemistry. Thus, guanine–cytosine do not spontaneously form the Watson–Crick (WC) pair (Fig. 5), characterized by the formation of three strong hydrogen bonds. Conversely, they adopt a different conformation in which only two symmetric hydrogen bonds are formed.[78, 81] Probably, such conformation gains stability by the resonance between keto and enol tautomers, which very likely results in a strong delocalization of the shared protons. Only if the hydrogen atom on N1 of cytosine is replaced by a methyl group resembling the attachment of a sugar unit, the two bases choose a WC configuration.[81] Other tree possible G–C aggregates formed by the two most stable tautomers of the two bases are also presented in Fig. 5. All these aggregates are not very different in energy, and therefore, any alteration of the environment may tip the balance towards adopting a different conformation. For example, addition of water has been demonstrated to strongly perturb the conformational landscape.[82]

Formation of DNA not only requires a strong propensity of nucleobases to form hydrogen bonds, but also their ability to establish $\pi \cdots \pi$ interactions. Actually, some studies point to a

competition between stacking and hydrogen bond in the aggregates of nucleosides, starting from the dimer.[83] Furthermore, inclusion of water molecules in the simulations favors formation of stacked structures, as a way to "hide" the hydrophobic sides of the nucleobases (the aromatic ring),[47] while the NH, OH, and CO groups remain available for the formation of hydrogen bonds with water.[71, 84] Our own studies on the subject point to formation of stacking structures, even in the absence of water, once certain size is reached.[85] Figure 6 shows the conformers of cytosine dimer, trimer, and tetramer detected in supersonic expansions. Two conformers were found for the dimer, based on different combinations of two very strong hydrogen bonds. According to calculations carried out at M06-2X/6-311++G(d,p), the binding energy of the system is close to 80 kJ/mol, which is comparable with the binding energy of a covalent bond, although one must take into account that two hydrogen bonds are contributing to such binding energy. In this situation, the two shared protons are delocalized between the two molecules, which do not have a well-defined tautomerism. When a third molecule is added, it attaches to the pre-existing ones in either of the two ways observed for the dimer, forming a planar structure. However, according to the calculations, there is a competition on the tetramer between stacked and linear conformers. Unfortunately, it is not possible to discern from the congested spectrum of such a large aggregate which of the two structures, either planar or stacked, is formed. The computational experiment demonstrates that methylation of N1 is enough to favor stacking over linear cluster growth.[85] Such tendency is reinforced in the nucleosides, due to the extra stabilization energy of the sugar stacking.[86] This propensity of the nucleobases towards forming stacked dimers may have also favored the appearance of a double strand of proto-DNA. The tendency of the nucleosides to form stacked dimers may have facilitated their fusion in a double strand, by addition of phosphate groups. However, more experiments are required to probe this hypothesis.

Confirmation of the importance of the ability of DNA bases to form hydrogen bonds also requires additional experiments on related metabolites already present in living beings, which exhibit similar combination of NH/CO groups. One of the most evident molecules is xanthine and its derivatives.[24] They are somehow a mixture between the backbone of adenine and the CO/N/CO/N motif in thymine, giving them certain affinity for some nucleobases receptors. This is probably the reason of the well-known stimulant properties of caffeine, one of the molecules of the family. Why nature chose CGAT over, for example, the molecules of the xanthine family are still a mystery, but characterization of their respective aggregation properties may help shedding some light on the subject. For example, exploration of their interactions with proteins or peptides and comparing the results from those of nucleobase–peptide interaction[87] may help understanding if a biopolymer built using xanthine derivatives would be as versatile as DNA, from an interaction point of view.

1.2 Amino Acids and Peptides

Proteins are the nano-machines of the cell. They are complex molecules, carefully designed, and optimized to carry out all kinds of tasks: from purely structural to complex catalytic processes. The number of proteins that compose the human proteome is still undetermined, but some authors report existence of > 100.000 different proteins,[88] all of them built using a combination of only 20 amino acids. The design of the amino acids is somehow similar to the concept behind the Lego blocks: they have a carboxylic acid group at one end and an amino group at the other (Fig. 7), enabling fast and energetically efficient assembly in the so-called peptide bond. In addition, they present a side chain that may contain pure aliphatic groups, NH, CO, aromatic or SH groups, and even charged groups, such as quaternary amines or carboxylic acids. The sequence of amino acids in a protein is optimized to induce its folding into a well-defined spatial conformation. Misfolding of a protein does not only involve loss of function, but it may also have pathological consequences, as the mad-cow disease demonstrated several years ago.[89]

Control over the final shape of a protein is carefully exerted by two mechanisms: formation of hydrogen bonds between the peptide bonds and by the correct combination of amino acids in the sequence of the protein.[90] The former is responsible for the formation of α-helices, β-sheets, and several types of turns (Fig. 8),[91] which are the most common structural motifs in proteins, although not the only ones. The α-helix is a structure in the shape of a right-handed helix, held together by C=O···H–N hydrogen bonds between non-consecutive amino acids. Each helix contains six AAs. This is the most predictable and prevalent structure in the proteins. However, not all the amino acids present the same propensity towards formation of α-helices and some

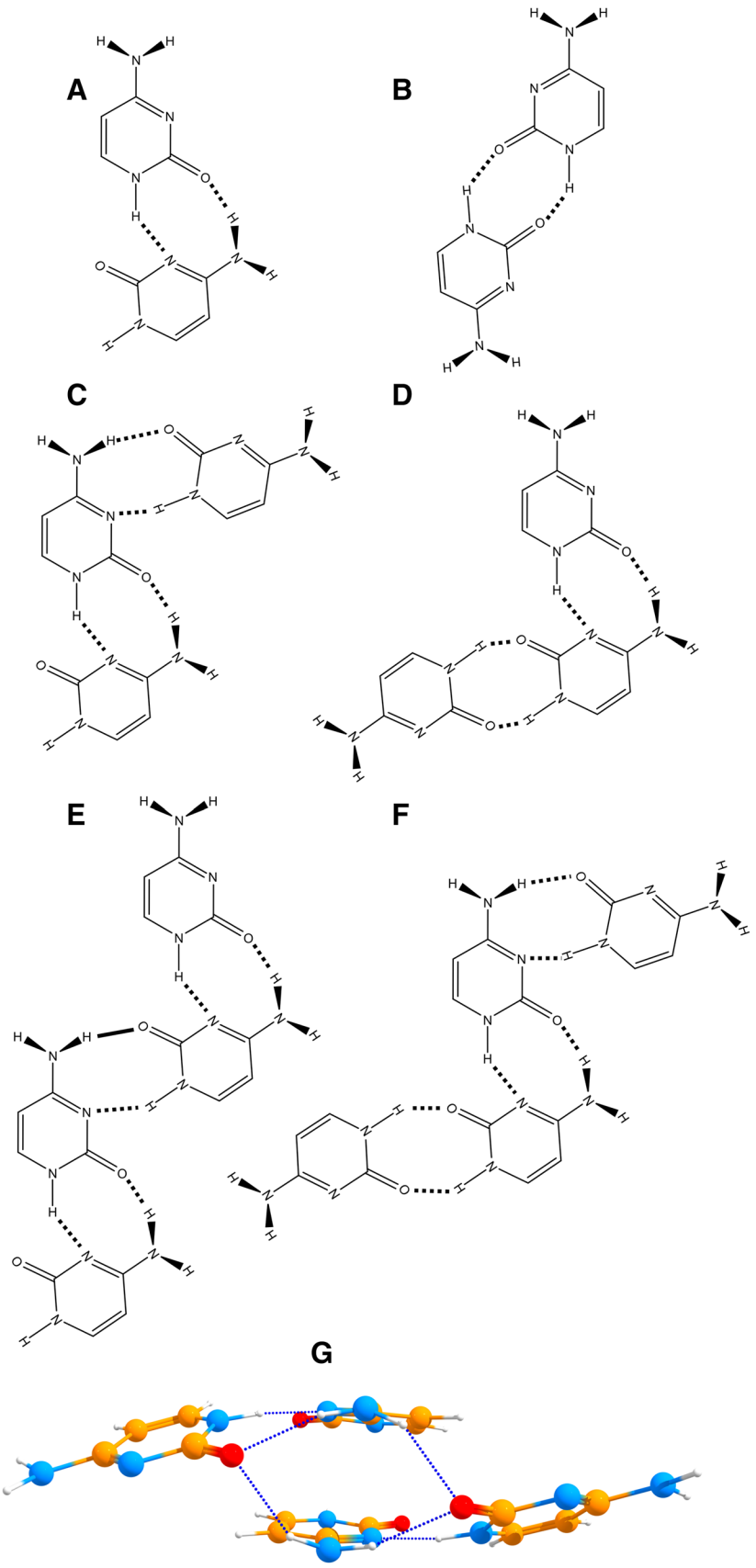

Figure 6: Structure of the detected aggregates of cytosine dimer (**a**, **b**), trimer (**c**, **d**) and tetramer (**e**–**g**) in supersonic expansions. The experimental data could not determine if the tetramer adopts a planar or stacked structure.

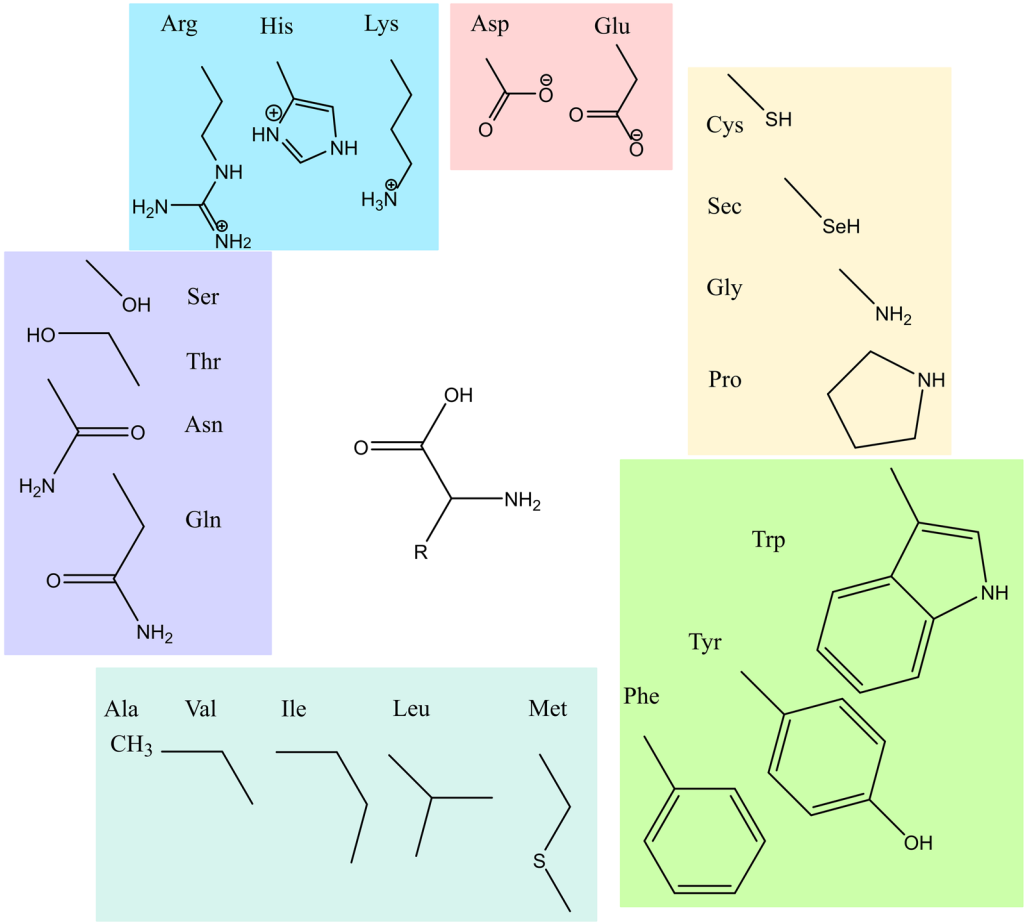

Figure 7: Structure of the 20 natural amino acids, plus de seleniated version of cysteine. The lateral chains are grouped according to their nature. Starting from the blue square and moving to the right: positively charged, negatively charged, special cases, aromatic, aliphatic, and polar side chains.

amino acid sequences are more prone to form β-sheets, structures composed by two sequences of AAs that run parallel and are held together by C=O···H–N interactions between the peptide bond of the AAs. The two sequences are usually separated by a variable number of AAs forming one (or several) of the many possible turns.[92]

One of the most frequent turns is the so-called γ-turn:[94] ordered, but not periodic, structures that may be classified by the torsional angles of the participant carbon atoms. All kinds of turns are usually stabilized by formation of hydrogen bonds between the oxygen and NH atoms in the protein's back bone. Putting all these elements together, the structure of a protein may be envisioned as α-helices and β-sheets connected by turns. It is a priori difficult to predict the folding of a protein, because it depends on a subtle balance between all the forces at play: formation of intramolecular hydrogen bonds, conformational preferences of each amino acid, and interaction of the AAs with the environment (water or lipids in the case of membrane proteins).[90] In principle, the preference of a given AA sequence to form and helix or a sheet may depend on the propensity of a section of the protein to fold in a certain way, to allow interactions to take place between closer or more distant AAs. Thus, the shape of the whole protein may be determined by the sequence of small, key sections. In this sense, several groups have conducted spectroscopic studies in jets to determine the conformational preferences of different combinations of AAs, yielding invaluable information to depurate the in silico models.[32, 33, 48, 49, 95–104]

The most striking observation is that all the interactions constituent of the above-mentioned structural elements were observed in jets, despite the size limitations of these types of experiments, and the requirement of introducing a chromophore (an aromatic ring) in the structure to carry out the spectroscopy.[46] Studies on peptides of

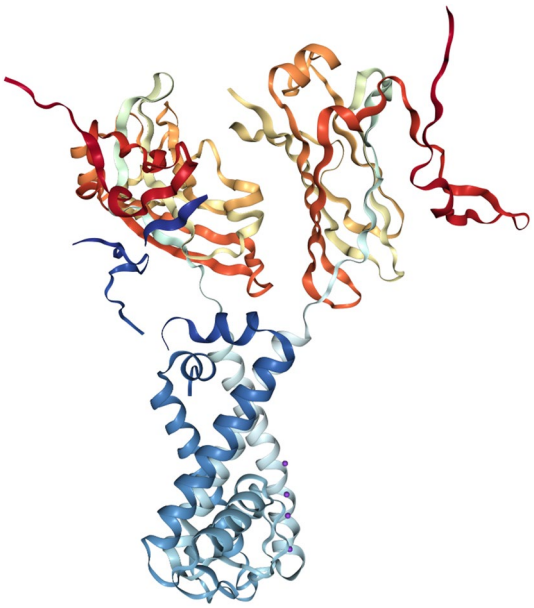

Figure 8: Model of a potassium channel (2WLI in the Protein Data Bank),[93] highlighting the tertiary structure. α-helix are depicted in blue and β-sheets in red–yellow–orange. Several turns connecting such structures are also visible.

increasing size determined that $Ac(Ala)_2$–O–Bn preferentially adopts conformations that resemble those in β-strands: the extended sequences that in the end result in formation of the β-sheets.[105] Replacing an alanine residue by a phenylalanine to form Ac-Phe-Ala-NH_2 produces a strong modification in the structure of the peptide, which now shows a marked preference towards γ-turn like conformations.[106, 107] If the AA sequence is reversed, the resulting dipeptide prefers adopting a β-turn like conformation.[108] If alanine is replaced by glycine, the resulting Ac–Gly–Phe–NH_2 peptide prefers formation of 2_7-ribbon like structures.[109] This is not a common structure, but its stability has been demonstrated for peptides containing Leu AAs.[110] Further elongation of the peptide introducing more alanine residues results in formation of helix-like structures (Ac–Ala–Phe–Ala–NH_2) and β-hairpin conformations.[111] Finally, β-sheet-like interactions were observed for some peptide dimers.[105, 112]

One must take into account that all these structures were formed in vacuum, i.e., in absence of any solvent, and therefore, they demonstrate that the basic structural motifs are already "coded" in the structure of the peptide backbone. There is still a long path ahead until complete understanding of the protein folding mechanism is achieved, in part due to the modulations introduced in the peptide structure by each side chain. Exploratory studies revealed that, for example, introduction of aromatic AAs may force the folding of the backbone in such a way that a direct $\pi \cdots \pi$ interaction may take place.[113] This effect may be reinforced in solution, due to the hydrophobic character of the aromatic rings. Gas-phase studies may also yield important information on the influence of a limited number of water molecules in the final structure of a peptide. However, these are technically challenging studies, and therefore, a very limited number of works have been published till the date.[114–118]

1.3 Sugars and Glycans

In addition to nucleobases and proteins, the cell contains an undetermined number of molecular species. Most of them are grouped into the so-called metabolome and their study is essential to connect genotype with phenotype. Among such a large collection of molecules, sugars and lipids are probably the most abundant families, and therefore, they deserve a special treatment. The study of lipids in jets presents technical difficulties that are almost impossible to solve with current techniques: they adopt flexible and dynamic structures formed by aliphatic chains that produce broad absorptions in the IR. Very limited structural information can be extracted from those spectral signatures.

On the other hand, sugars are small molecules, essential for many processes.[9] Apart from their well-known role in metabolism, they are key molecules for the immune system. The lipid membrane is decorated with glycans that serve as cellular ID:[119, 120] the cells of the immune system patrol the tissues probing such polysaccharides, and if the sugar combination is not recognized, the cell is marked as foe and an immune response is elicited. It is not surprising that sugars were chosen for such specialized task, due to their structural characteristics. As polyhydroxyaldehydes, they contain a combination of hydroxyl groups that are coupled through formation of intramolecular hydrogen bonds. Figure 9 compares the structure of the hexoses in their pyranose form. The substituent in C1, the anomeric carbon, can be in either axial or equatorial position, resulting in α or β anomers. Starting from β-glucose, which is the sugar that has all the OH substituents in equatorial position, the rest of the sugars are obtained moving one or several hydroxyl groups to axial conformation.[121] As all the hydroxyl groups are involved in formation of cooperative hydrogen-bond networks,

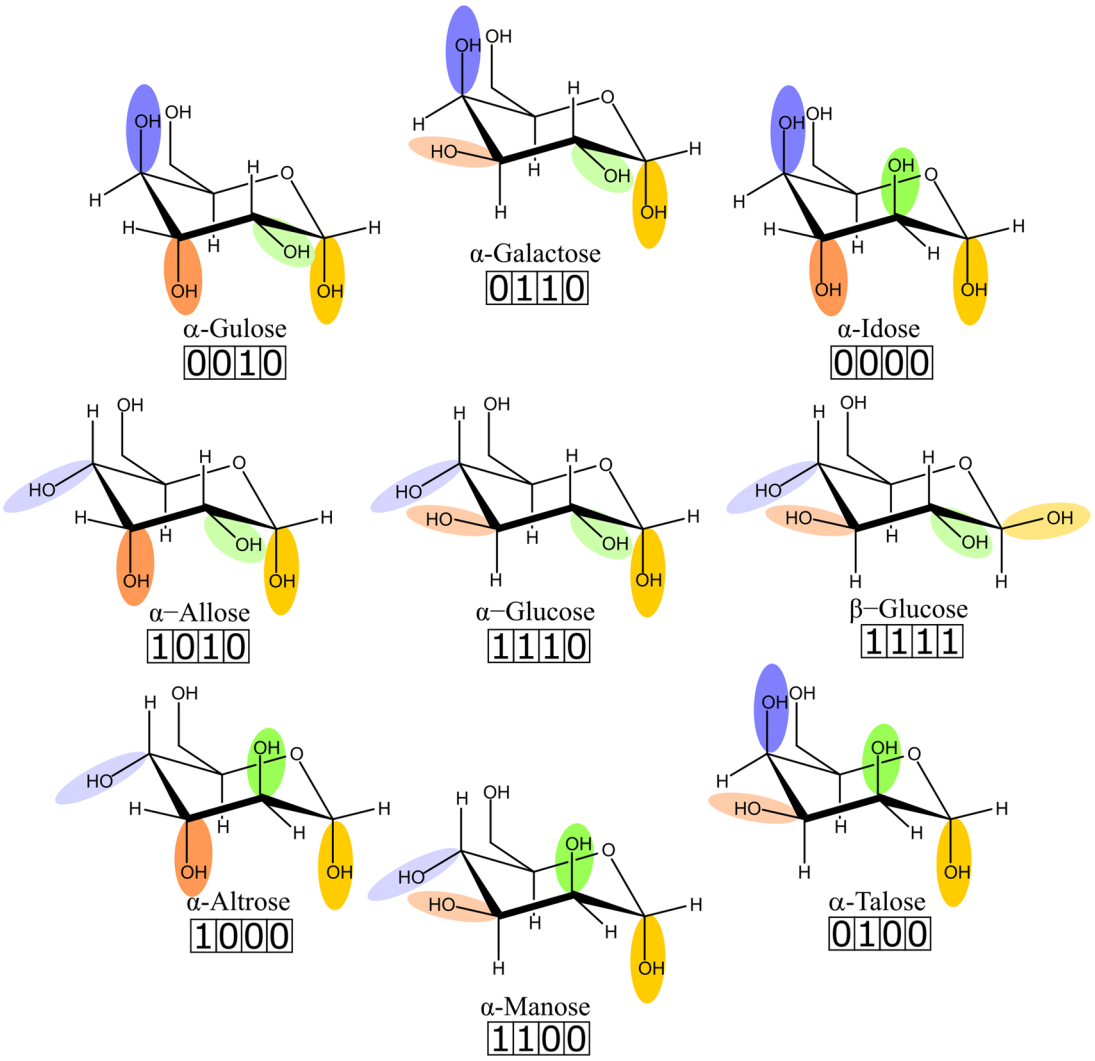

Figure 9: Structure of the hexoses in pyranose form. Starting from β-glucose, which has all the hydroxyl groups in equatorial position, the rest of the sugars are obtained by moving one or several (all in the case of α-idose) OH substituents into axial position. Darker colors highlight OHs in axial position. In a sense, sugar molecules resemble a binary register of a computer and the axial/equatorial positions of the OHs represent the 1 s and 0 s.

modification of the position of a single OH has a noticeable impact in the shape of the network, which somehow amplifies the structural changes. In a sense, the axial/equatorial disposition of the OH groups in the sugar molecules may be envisioned as the 1 s and 0 s in the binary register of a computer. Adding several sugar units to form a polysaccharide, a significantly long register can be built. Only the correct combination of axial/equatorial OHs, the equivalent to the correct value in the binary register, is recognized as valid by the receptors in the cells of the immune system.[120, 122–124]

Exploration of the conformational preferences of sugars in jets was started by Simon's group using laser spectroscopy.[40, 41, 123, 125–133] Such studies were later complemented with others using MW spectroscopy to map the structure of all main sugar molecules.[134–140] Several conclusions were derived from such studies. First, sugars have strong preference for the pyranose form in jets. Certainly, monosaccharides can adopt in solution an extended form or they can cyclize through intramolecular nucleophilic attack. For example, hexoses can be found in pyranose (6-membered ring, Fig. 10) or furanose (5-membered ring) forms. As cycle formation may take place with two possible orientations, monosaccharides can interconvert between α and β anomeric forms in solution, although they are not isoenergetic.

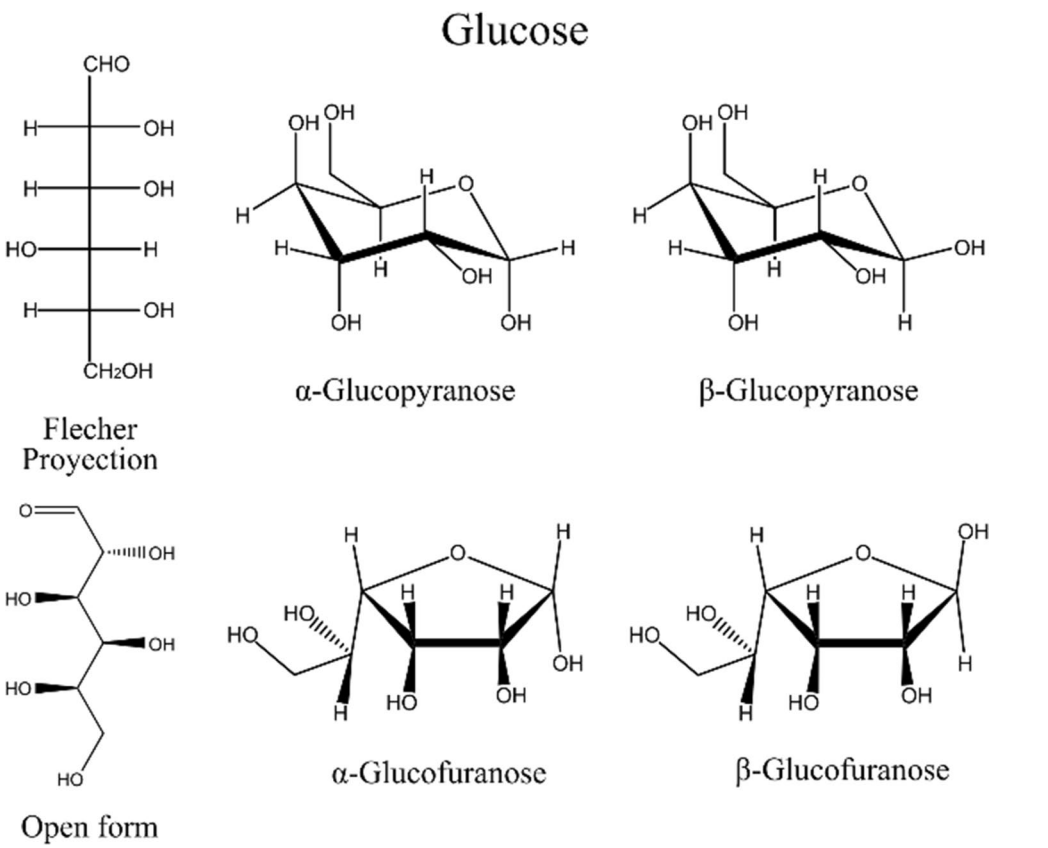

Figure 10: Sugars may adopt linear or cyclic structures. Depending on the cyclization mechanism, hexoses may be found as six-membered (pyranosides) or five-membered (furanosides) rings.

To block the mutarotation and to enable the use of electronic spectroscopy to characterize the sugar molecules, a substituent is introduced in the anomeric carbon, usually an aromatic ring. Such modification has a non-negligible impact in the conformational preferences, but the benefits derived from such modification compensate the small alteration introduced in the molecular conformational landscape.

Despite the presence of several OH groups in the monosaccharides, they interact with other molecules in a limited number of ways, with the hydroxymethyl group as the preferred interaction site.[41, 141–145] This is a consequence of the formation of cooperative hydrogen bonds between the OH groups. Inclusion of, for example, a water molecule between two hydroxyl groups requires of a first step that involves H-bond breaking, and therefore, the process has a substantial potential energy barrier.[146, 147] The second reason that favors interaction with the hydroxymethyl group is its flexibility, which enables optimizing its position to maximize the interaction energy. This is clearly seen in Fig. 11, where the interaction between β-phenyl-D-glucopyranose (β-PhGlc) and α-/β-glucopyranose (α-/β-Glc), *N*-methylacetamide, paracetamol, phenol, and α-/β-methyl-D-glucopyranose (α-/β-MeGlc) is analyzed.[141] All the molecules are attracted towards the hydroxymethyl group of β-PhGlc and are trapped between it and the aromatic ring. Conversely, interaction with α-/β-Glc and α-/β-MeGlc takes place through β-PhGlc O3 and O4. Interestingly, β-PhGlc resembles the primer used by the enzyme glycogenin to start glycogen synthesis from glucose. On the light of these results, one is tempted to speculate that perhaps tyrosine was chosen by nature as the docking amino acid for the first glucose molecule, because the combination of the sugar's hydroxymethyl group and the tyrosine's aromatic ring constitutes a kind of trap for the wandering molecules, keeping them away from the interaction site where the next glucose molecule has to bind. In this way, blocking the polymerization site would be prevented. In favor of this hypothesis, replacing the aromatic ring of β-PhGlc by a methyl group results in the shift of the interaction site towards O3/O4.

An interesting property of sugars is the differences in the inter-molecular interactions that the

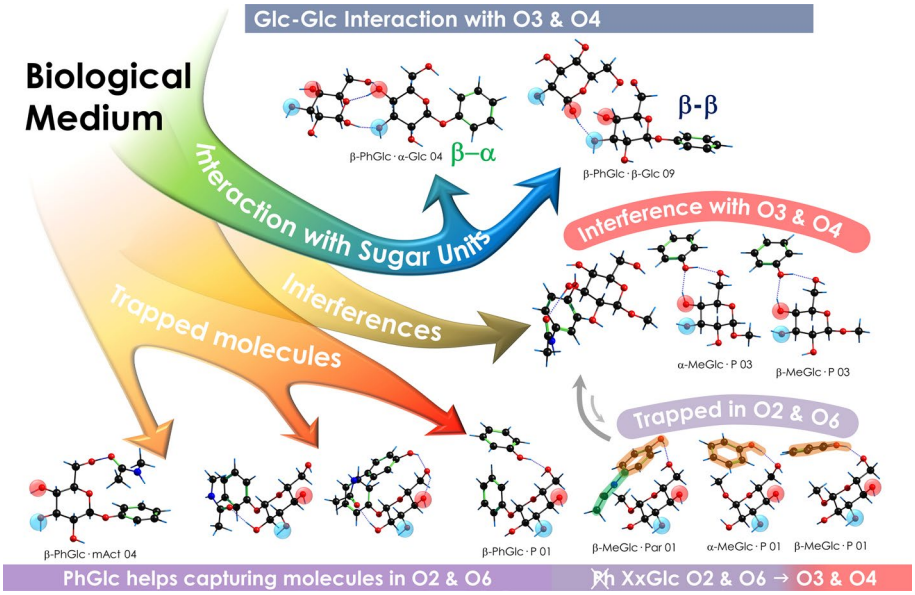

Figure 11: Interaction of b-phenyl-D-glucopyranose (β-PhGlc) and α-/β-methyl-D-glucopyranose (α-/β-MeGlc) with α-/β-glucopyranose (α-/β-Glc), *N*-methylacetamide, paracetamol, phenol, and α-/β-MeGlc. The combination of the hydroxymethyl group with the aromatic ring in β-PhGlc create a kind of trap that captures wandering molecules, keeping them away from O3 and O4, which are the sites for addition of the next glucose unit during glycogen production. Only sugar units are able to avoid such trap, and interestingly, they interact preferentially with O3/O4 of β-PhGlc Adapted from Ref. [141].

position of the hydroxyl groups introduces. For example, as demonstrated in previous works,[142] interaction between β-PhGlc and MeGlc is substantially stronger with the β anomer than with the α anomer. The axial orientation of the anomeric substituent enables formation of a very symmetric structure, in which the two molecules fusion their respective H-bond networks in a single one that runs along both structures. Conversely, interaction with the α anomer is substantially less favorable. These conclusions were extended to other combinations of sugar derivatives using DFT calculations. In all cases, it was demonstrated that interaction between β anomers was more stable.[142]

The same effect was observed when the structure of β-PhGlc dimer was compared with that of β-PhGal despite that the difference between the two molecules is the position of a single OH.[148] While formation of extended H-bond networks is possible in $(\beta\text{-PhGal})_2$, such superstructures are not favored in $(\beta\text{-PhGlc})_2$. All these observations may be extrapolated to the structure of large glycans and their interaction with the corresponding receptor: even in such large molecules, detection of alterations in the position of a single hydroxyl group may be possible thanks to the amplification mechanism that the formation of H-bond networks has in the final structure of the polysaccharide.

2 Summary

We have revised in this mini-review the influence of the hydrogen bond in the structure and function of three families of biomolecules: DNA, proteins, and saccharides. The examples presented highlight the importance of hydrogen bond in the final structure of those molecules and how nature makes extensive use of H bond to produce complex 3D structures in proteins. The structure of saccharides is also largely conditioned by the formation of intra- and inter-molecular hydrogen bonds. We speculate here with the idea of how small structural changes in the position of the hydroxyl groups are amplified by the intramolecular network of hydrogen bonds, enabling easy recognition by other biomolecules. Perhaps, this is one of the reasons why nature chose carbohydrates to code the cellular ID.

We also present the importance of stacking but specially hydrogen-bond interactions in DNA and how it may have influenced the election of CGAT as the molecules to build the so-called "alphabet of life". Most of the information presented in this short review comes from

experiments in supersonic expansions using a collection of spectroscopic techniques, whose results were interpreted in the light of calculations using computational chemistry. Despite the exotic environment that the molecules encounter in the expansion, the results presented demonstrate that biologically relevant information can be extracted from such experiments. Especially, taking into account the increase in size of the systems tackled in the last 20 years. Still, there is a long road ahead until hydrogen bond is fully understood in biological environments. Evaluation of the influence of the biological medium is one of the variables that present the greatest challenges. The studies published are not able to introduce more than a dozen of solvent molecules before the spectra became so congested that it is no longer possible to extract structural information.

Another interesting aspect for future research is the interaction of the biomolecules reported here with molecules that may act as solvents in other planets. The different environmental conditions found in other words may lead to existence of seas of methane or ammonia, or to very acidic conditions. Understanding of how amino acids, nucleobases, and sugars behave and interact under such (for us) extreme conditions may help to guide the search for life in exoplanets, and may demonstrate if life can survive very different conditions such as extreme pH or temperature. Exploration of H bond in such systems is an exciting perspective.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

This work has received funds from MINECO/FEDER (CTQ2015-68148-C2-1-P), MICIU(PGC2018-098561-B-C21) and Basque Government (IT1162-19). Most of the results presented in this work have been possible thanks to the support of the computing infrastructure of the i2BASQUE academic network and the SGI/IZO-SGIker network. I would like to thank the technical support from the personnel from the UPV/EHU laser facility.

Received: 24 October 2019 Accepted: 2 December 2019
Published online: 27 December 2019

References

1. Crowe SA, Dossing LN, Beukes NJ, Bau M, Kruger SJ, Frei R, Canfield DE (2013) Atmospheric oxygenation three billion years ago. Nature 501:535
2. Hobza P, Muller-Dethlefs K (2010) Non-covalent interactions. RCS Publishing, Milan
3. Johnson ER, Keinan S, Mori-Sánchez P, Contreras-García J, Cohen AJ, Yang W (2011) Revealing Non-Covalent Interactions. J Am Chem Soc 132:6498–6506
4. Reed AE, Curtiss LA, Weinhold F (1988) Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. Chem Rev 88:899–926
5. Stone AJ (2002) The theory of intermolecular forces. Oxford University Press, Oxford
6. Müller-Dethlefs K, Hobza P (2000) Noncovalent interactions: a challenge for experiment and theory. Chem Rev 100:143–168
7. Biochemistry of Lipids (2002) Lipoproteins and membranes. Elsevier Science, New York
8. Brown DA, London E (1998) Functions of lipid rafts in biological membranes. Annu Rev Cell Dev Biol 14:111–136
9. Lehninger A, Nelson D, Cox M (2008) Lehninger principles of biochemistry. W. H. Freeman, New York
10. Poblotzki A, Gottschalk HC, Suhm MA (2017) Tipping the scales: spectroscopic tools for intermolecular energy balances. J Phys Chem Lett 8:5656–5665
11. Puzzarini C, Bloino J, Tasinato N, Barone V (2019) Accuracy and interpretability: the devil and the holy grail. New routes across old boundaries in computational spectroscopy. Chem Rev 119:8131–8191
12. Dopfer O, Fujii M (2016) Probing solvation dynamics around aromatic and biological molecules at the single-molecular level. Chem Rev 116:5432–5463
13. Smalley RE, Levy DH, Wharton L (1975) Molecular-jet spectroscopy. Laser Focus 11:40–43
14. Smalley RE, Blazy JA, Fitch PSH, Kim MS, Wharton L, Levy DH (1976) Laser spectroscopy in supersonic molecular-beams. Opt Commun 18:59–60
15. Zwier T (2001) Laser spectroscopy of jet-cooled biomolecules and their water-containing clusters: Water bridges and molecular conformation. J Phys Chem A 105:8827–8839
16. León I, Millán J, Cocinero EJ, Lesarri A, Fernández JA (2013) Magic numbers in the solvation of the propofol dimer. Chem Phys Chem 14:1558–1562
17. Ashfold MNR, Howe JD (1994) Multiphoton spectroscopy of molecular species. Annu Rev Phys Chem 45:57–82
18. Held A, Pratt DW (1992) Hydrogen bonding in the symmetry-equivalent C_{2h} dimer of 2-pyridone in its S_0 and S_2 electronic states. Effect of deuterium substitution. J Chem Phys 96:4869–4876
19. Becucci M, Melandri S (2016) Resolution spectroscopic studies of complexes formed by medium-size organic molecules. Chem Rev 116:5014–5037

20. Pérez C, Muckle MT, Zaleski DP, Seifert NA, Temelso B, Shields GC, Kisiel Z, Pate BH (2012) Structures of cage, prism, and book isomers of water hexamer from broadband rotational spectroscopy. Science 336:897–901
21. Seifert NA, Steber AL, Neill JL, Pérez C, Zaleski DP, Pate BH, Lesarri A (2013) The interplay of hydrogen bonding and dispersion in phenol dimer and trimer: structures from broadband rotational spectroscopy. Phys Chem Chem Phys 15:11468–11477
22. Pérez C, León I, Lesarri A, Pate BH, Martínez R, Millán J, Fernández JA (2018) Isomerism of the aniline trimer. Angew Chem Int Ed 57:15112–15116
23. Calabrese C, Li W, Prampolini G, Evangelisti L, Uriarte I, Cacelli I, Melandri S, Cocinero EJ (2019) A general treatment to study molecular complexes stabilized by Hydrogen-, Halogen-, and Carbon-Bond networks: experiment and theory of $(CH_2F_2)_n\cdots(H_2O)_m$. Angew Chem Int Ed 58:8437–8442
24. Usabiaga I, Camiruaga A, Calabrese C, Maris A, Fernández JA (2019) Exploring caffeine-phenol interactions by the inseparable duet of experimental and theoretical data. Chem Eur J 25:14230–14236
25. Zwier TS (1996) The spectroscopy of solvation in hydrogen-bonded aromatic clusters. Annu Rev Phys Chem 47:205–241
26. Fernández JA, Unamuno I, Longarte A, Castaño F (2001) S-0, S-1, and ion I-0 binding energies of the p-methoxyphenethylamine(H_2O)(1-4) complexes. J Phys Chem A 105:961–968
27. Longarte A, Fernández JA, Unamuno I, Castaño F (2000) Isomer structures and vibrational assignment of the methyl-p-aminobenzoate(H_2O)(1) complex. J Chem Phys 112:3170–3180
28. León I, Millan J, Cocinero E, Lesarri A, Castaño F, Fernández JA (2012) Mimicking anaesthetic-receptor interaction: a combined spectroscopic and computational study of propofol···phenol. Phys Chem Chem Phys 14:8956–8963
29. Siffert L, Blaser S, Ottiger P, Leutwyler S (2018) Transition from water wires to bifurcated H-bond networks in 2-Pyridone·$(H_2O)_n$, n = 1-4 Clusters. J Phys Chem A 122:9285–9297
30. Knochenmuss R, Sinha RK, Leutwyler S (2019) Face, Notch, or Edge? Intermolecular dissociation energies of 1-naphthol complexes with linear molecules. J Chem Phys 150:234303
31. Asami H, Tokugawa M, Masaki Y, Ishiuchi S, Gloaguen E, Seio K, Saigusa H, Fujii M, Sekine M, Mons M (2016) Effective strategy for conformer-selective detection of short-lived excited state species: application to the IR spectroscopy of the N1H Keto tautomer of guanine. J Phys Chem A 120:2179–2184
32. Habka S, Sohn WY, Vaquero-Vara V, Geleoc M, Tardivel B, Brenner V, Gloaguen E, Mons M (2018) On the turn-inducing properties of asparagine: the structuring role of the amide side chain, from isolated model peptides to crystallized proteins. Phys. Chem. Chem. Phys. 20:3411–3423
33. Kumar S, Mishra KK, Singh SK, Borish K, Dey S, Sarkar B, Das A (2019) Observation of a weak intra-residue C5 hydrogen-bond in a dipeptide containing Gly-Pro sequence. J Chem Phys 151:104309
34. Bernhard D, Fatima M, Poblotzki A, Steber AL, Pérez C, Suhm MA, Schnell M, Gerhards M (2019) Dispersion-controlled docking preference: multi-spectroscopic study on complexes of dibenzofuran with alcohols and water. Phys Chem Chem Phys 21:16032–16046
35. Mori H, Kugisaki H, Inokuchi Y, Nishi N, Miyoshi E, Sakota K, Ohashi K, Sekiya H (2002) LIF and IR dip spectra of jet-cooled p-aminophenol-M (M = CO, N_2): hydrogen-bonded or van der waals-bonded structure?. J Phys Chem A 106:4886–4890
36. Dietrich F, Bernhard D, Fatima M, Pérez C, Schnell M, Gerhards M (2018) The effect of dispersion on the structure of diphenyl ether aggregates. Angew Chem Int Ed Engl 57:9534–9537
37. Weiler M, Bartl K, Gerhards M (2012) Infrared/ultraviolet quadruple resonance spectroscopy to investigate structures of electronically excited states. J Chem Phys 136:114202–114206
38. Bartl K, Funk A, Gerhards M (2008) IR/UV spectroscopy on jet cooled 3-hydroxyflavone $(H_2O)_n$ (n = 1,2) clusters along proton transfer coordinates in the electronic ground and excited states. J Chem Phys 129:234306–234311
39. Cocinero EJ, Çarçabal P (2015) Rijs AM, Oomens J (eds) In gas-phase IR spectroscopy and structure of biological molecules. Springer, New York, pp 299–334
40. Barry CS, Cocinero EJ, Çarçabal P, Gamblin DP, Stanca-Kaposta E, Remmert SM, Fernández-Alonso MC, Rudic S, Simons JP, Davis BG (2013) 'Naked' and hydrated conformers of the conserved core pentasaccharide of N-linked glycoproteins and its building blocks. J Am Chem Soc 135:16895–16903
41. Stanca-Kaposta E, Çarçabal P, Cocinero EJ, Hurtado P, Simons JP (2013) Carbohydrate-aromatic interactions: vibrational spectroscopy and structural assignment of isolated monosaccharide complexes with p-hydroxy toluene and N-acetyl l-tyrosine methylamide. J Phys Chem B 117:8135–8142
42. Aguado E, León I, Cocinero EJ, Lesarri A, Fernández JA, Castaño F (2009) Molecular recognition in the gas phase: benzocaine-phenol as a model qof anaesthetic-receptor interaction. Phys Chem Chem Phys 11:11608–11616
43. Bhattacherjee A, Wategaonkar S (2016) Water bridges anchored by a C-HO hydrogen bond: the role of weak interactions in molecular solvation. Phys Chem Chem Phys 18:27745–27749
44. Bhattacherjee A, Wategaonkar S (2015) Conformational preferences of monohydrated clusters of

imidazole derivatives revisited. Phys Chem Chem Phys 17:20080–20092
45. Biswal HS, Gloaguen E, Mons M, Bhattacharyya S, Shirhatti PR, Wategaonkar S (2011) Structure of the indole-benzene dimer revisited. J Phys Chem A 115:9485–9492
46. Rijs AM and Oomens J (eds) (2015) Gas-phase IR spectroscopy and structure of biological molecules. Springer International Publishing, Heidelberg, New York, Dordrecht, London
47. de Vries MS (2015) In Rijs MA, Oomens J (eds) Gas-phase ir spectroscopy and structure of biological molecules. Springer International Publishing, Cham, pp 271–297
48. Jaeqx S, Du WN, Meijer EJ, Oomens J, Rijs AM (2013) Conformational study of Z-Glu-OH and Z-Arg-OH: dispersion interactions versus conventional hydrogen bonding. J Phys Chem A 117:1216–1227
49. Rijs AM, Kabelac M, Abo-Riziq A, Hobza P, de Vries MS (2011) Isolated gramicidin peptides probed by IR spectroscopy. Chem Phys Chem 12:1816–1821
50. Rijs AM, Kay ER, Leigh DA, Buma WJ (2011) IR spectroscopy on jet-cooled isolated two-station rotaxanes. J Phys Chem A 115:9669–9675
51. León I, Cocinero EJ, Rijs AM, Millán J, Alonso E, Lesarri A, Fernández JA (2013) Formation of water polyhedrons in propofol-water clusters. Phys Chem Chem Phys 15:568–575
52. León I, Cocinero E, Millán J, Jaeqx S, Rijs A, Lesarri A, Castaño F, Fernández JA (2012) Exploring microsolvation of the anesthetics propofol. Phys Chem Chem Phys 14:4398
53. León I, Cocinero EJ, Millán J, Rijs AM, Usabiaga I, Lesarri A, Castaño F, Fernández JA (2012) A combined spectroscopic and theoretical study of propofol·$(H^2O)_3$. J Chem Phys 137:074303
54. Nakamura T, Schmies M, Patzer A, Miyazaki M, Ishiuchi S, Weiler M, Dopfer O, Fujii M (2014) Solvent migration in microhydrated aromatic aggregates: ionization-induced site switching in the 4-aminobenzonitrile-water cluster. Chemistry 20:2031–2039
55. Schmies M, Patzer A, Schuetz M, Miyazaki M, Fujii M, Dopfer O (2014) Microsolvation of the acetanilide cation (AA(+)) in a nonpolar solvent: IR spectra of AA(+)-L-n clusters (L = He, Ar, N-2; n <= 10). Phys Chem Chem Phys 16:7980–7995
56. Nakamura T, Miyazaki M, Ishiuchi S, Weiler M, Schmies M, Dopfer O, Fujii M (2013) IR Spectroscopy of the 4-AminobenzonitrileAr Cluster in the S0, S1 Neutral and D0 Cationic States. Chem Phys Chem 14:741–745
57. Beckstead AA, Zhang Y, de Vries MS, Kohler B (2016) Life in the light: nucleic acid photoproperties as a legacy of chemical evolution. Phys Chem Chem Phys 18:24228–24238
58. Kleinermanns K, Nachtigallová D, de Vries MS (2013) Excited state dynamics of DNA bases. Int Rev Phys Chem 32:308–342
59. Abo-Riziq A, Grace L, Nir E, Kabelac M, Hobza P, de Vries MS (2005) Photochemical selectivity in guanine–cytosine base-pair structures. Proc Natl Acad Sci USA 102:20–23
60. Shin J, Bernstein ER (2014) Vacuum ultraviolet photoionization of carbohydrates and nucleotides. J Chem Phys 140:044330
61. Sobolewski AL, Domcke W, Hättig C (2005) Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: the role of electron-driven proton-transfer processes. Proc Natl Acad Sci USA 102:17903–17906
62. Zhang W, Yuan S, Wang Z, Qi Z, Zhao J, Dou Y, Lo GV (2011) A semiclassical dynamics simulation for a long-lived excimer state of π-stacked adenines. Chem Phys Lett 506:303–308
63. Urashima S, Asami H, Ohba M, Saigusa H (2010) Microhydration of the guanine-guanine and guanine-cytosine base pairs. J Phys Chem A 114:11231–11237
64. McGinty RK, Tan S (2015) Nucleosome structure and function. Chem Rev 115:2255–2273
65. Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. Cell 2007(128):635–638
66. Mondal S, Puranik M (2017) Ultrafast structural dynamics of photoexcited adenine. Phys Chem Chem Phys 19:20224–20240
67. Schwell M, Hochlaf M (2015) In: Barbatti M, Borin AC, Ullrichin S (eds) Photoinduced phenomena in nucleic acids I: nucleobases in the gas phase and in solvents. Springer International Publishing, Cham, pp 155–208
68. Sun C, Ding F, Ding Y, Wang C (2015) The nonadditivity of stacking interactions in adenine–thymine and guanine–cytosine stacked structures: Study by MP2 and SCS-MP2 calculations. J Theor Comput Chem 14:1550037
69. Hunger K, Buschhaus L, Biemann L, Braun M, Kovalenko S, Improta R, Kleinermanns K (2013) UV-light-induced hydrogen transfer in guanosine–guanosine aggregates. Chem A Eur J 19:5425–5431
70. Mishra D, Pal S (2009) Ionization potential and structure relaxation of adenine, thymine, guanine and cytosine bases and their base pairs: A quantification of reactive sites. J Mol Struct (Thoechem) 902:96–102
71. Zeleny T, Hobza P, Kabelac M (2009) Microhydration of guanine···cytosine base pairs, a theoretical Study on the role of water in stability, structure and tautomeric equilibrium. Phys Chem Chem Phys 11:3430–3435
72. Shukla MK, Kuramshina GM, Leszczynski J (2007) Evidence of structural non-planarity in excited state: New findings provided by vibrational analysis of the guanine-cytosine base pair. Chem Phys Lett 447:330–334
73. Crews B, Abo-Riziq A, Grace L, Callahan M, Kabeláč M, Hobza P, de Vries MS (2005) IR-UV double resonance

spectroscopy of guanine-H_2O clusters. Phys Chem Chem Phys 7:3015–3020

74. Jurecka P, Hobza P (2003) True stabilization energies for the optimal planar hydrogen-bonded and stacked structures of guanine···cytosine, adenine···thymine, and their 9- and 1-methyl derivatives: complete basis set calculations at the MP2 and CCSD(T) levels and comparison with experiment. J Am Chem Soc 125:15608–15613
75. Triandafillou CG, Matsika S (2013) Excited-state tautomerization of gas-phase cytosine. J Phys Chem A 117:12165–12174
76. Villani G (2005) Theoretical investigation of hydrogen transfer mechanism in the adenine–thymine base pair. Chem Phys 316:1–8
77. Nir E, Hunig I, Kleinermanns K, de Vries MS (2003) The nucleobase cytosine and the cytosine dimer investigated by double resonance laser spectroscopy and ab initio calculations. Phys Chem Chem Phys 5:4780–4785
78. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2002) Pairing of the nucleobases guanine and cytosine in the gas phase studied by IR-UV double-resonance spectroscopy and ab initio calculations. Phys Chem Chem Phys 4:732–739
79. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2001) Guanine tautomerism revealed by UV-UV and IR-UV hole burning spectroscopy. J Chem Phys 115:4604–4611
80. Plutzer C, Nir E, de Vries M, Kleinermanns K (2001) IR-UV double-resonance spectroscopy of the nucleobase adenine. Phys Chem Chem Phys 3:5466–5469
81. Nir E, Kleinermanns K, de Vries MS (2000) Pairing of isolated nucleic-acid bases in the absence of the DNA backbone. Nature 408:949–951
82. Florián J, Leszczynski J, Scheiner S (1995) Ab initio study of the structure of guanine-cytosine base pair conformers in gas phase and polar solvents. Mol Phys 84:469–480
83. Asami H, Yagi K, Ohba M, Urashima S, Saigusa H (2013) Stacked base-pair structures of adenine nucleosides stabilized by the formation of hydrogen-bonding network involving the two sugar groups. Chem Phys 419:84–89
84. Kabelac M, Ryjacek F, Hobza P (2000) Already two water molecules change planar H-bonded structures of the adenine···thymine base pair to the stacked ones: a molecular dynamics simulations study. Phys Chem Chem Phys 2:4906–4909
85. González J, Usabiaga I, Arnaiz PF, León I, Martínez R, Millán J, Fernández JA (2017) Competition between stacked and hydrogen bonded structures of cytosine aggregates. Phys Chem Chem Phys 19:8826–8834
86. Ligare MR, Rijs AM, Berden G, Kabeláč M, Nachtigallova D, Oomens J, de Vries MS (2015) Resonant infrared multiple photon dissociation spectroscopy of anionic nucleotide Monophosphate clusters. J Phys Chem B 119:7894–7901
87. González J, Baños I, León I, Contreras-García J, Cocinero EJ, Lesarri A, Fernández JA, Millán J (2016) Unravelling protein-DNA interactions at molecular level: a DFT and NCI study. J Chem Theory Comput 12:523–534
88. Ponomarenko EA, Poverennaya EV, Ilgisonis EV, Pyatnitskiy MA, Kopylov AT, Zgoda VG, Lisitsa AV, Archakov AI (2016) The size of the human proteome: the width and depth. Int J Anal Chem 2016:7436849
89. Ng SBL, Doig AJ (1997) Molecular and chemical basis of prion-related diseases. Chem Soc Rev 26:425
90. Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. Annual Rev Biophys 37:289–316
91. Eisenberg D (2003) The discovery of the alpha -helix and beta -sheet, the principal structural features of proteins. Proc Natl Acad Sci 100:11207–11210
92. Feldmann RJ (1989) In Wittmann-Liebold B (eds) Methods in protein sequence analysis. Springer, Berlin, Heidelberg, pp 379–384
93. Clarke OB, Caputo AT, Hill AP, Vandenberg JI, Smith BJ, Gulbis JM (2010) Domain reorientation and rotation of an intracellular assembly regulate conduction in kir potassium channels. Cell 141:1018–1029
94. Matthews BW (1972) The gamma turn. evidence for a new folded conformation in proteins. Macromolecules 5:818–819
95. Forsting T, Gottschalk HC, Hartwig B, Mons M, Suhm MA (2017) Correcting the record: the dimers and trimers of trans-N-methylacetamide. Phys Chem Chem Phys 19:10727–10737
96. Ishiuchi S, Yamada K, Chakraborty S, Yagi K, Fujii M (2013) Gas-phase spectroscopy and anharmonic vibrational analysis of the 3-residue peptide Z-Pro-Leu-Gly-NH2 by the laser desorption supersonic jet technique. Chem Phys 419:145–152
97. Chakraborty S, Yamada K, Ishiuchi S, Fujii M (2012) Gas phase IR spectra of tri-peptide Z-Pro-Leu-Gly: effect of C-terminal amide capping on secondary structure. Chem Phys Lett 531:41–45
98. Schwing K, Fricke H, Bartl K, Polkowska J, Schrader T, Gerhards M (2012) Isolated beta-turn model systems investigated by combined IR/UV spectroscopy. Chem Phys Chem 13:1576–1582
99. Cocinero EJ, Stanca-Kaposta EC, Gamblin DP, Davis BG, Simons JP (2009) Peptide secondary structures in the gas phase: consensus motif of N-linked glycoproteins. J Am Chem Soc 131:1282–1287
100. Baquero EE, James WH, Choi SH, Gellman SH, Zwier TS (2008) Single-conformation ultraviolet and infrared spectroscopy of model synthetic foldamers: beta-peptides Ac-beta(3)-hPhe-NHMe and Ac-beta(3)-hTyr-NHMe. J Am Chem Soc 130:4784–4794

101. Haber T, Seefeld K, Kleinermanns K (2007) Mid- and near-infrared spectra of conformers of H-Pro-Trp-OH. J Phys Chem A 111:3038–3046
102. Toroz D, Van Mourik T (2006) The structure of the gas-phase tyrosine-glycine dipeptide. Mol Phys 104:559–570
103. Gregurick SK, Fredj E, Elber R, Gerber RB (1997) Vibrational spectroscopy of peptides and peptide-water complexes: Anharmonic coupled-mode calculations. J Phys Chem B 101:8595–8606
104. Cable JR, Tubergen MJ, Levy DH (1989) Fluorescence spectroscopy of jet-cooled tryptophan peptides. J Am Chem Soc 111:9032–9039
105. Gloaguen E, De Courcy B, Piquemal J, Pilmei J, Parisel O, Pollet R, Biswal HS, Piuzzi F, Tardivel B, Broquier M, Mons M (2010) Gas-phase folding of a two-residue model peptide chain: on the importance of an interplay between experiment and theory. J Am Chem Soc 132:11860–11863
106. Chin W, Piuzzi F, Dognon J, Dimicoli I, Mons M (2005) Gas-phase models of gamma turns: effect of side-chain/backbone interactions investigated by IR/UV spectroscopy and quantum chemistry. J Chem Phys 123:084301
107. Gloaguen E, Pagliarulo F, Brenner V, Chin W, Piuzzi F, Tardivel B, Mons M (2007) Intramolecular recognition in a jet-cooled short peptide chain: Î³-turn helicity probed by a neighbouring residue. Phys Chem Chem Phys 9:4491
108. Chin W, Dognon J, Piuzzi F, Tardivel B, Dimicoli I, Mons M (2005) Intrinsic folding of small peptide chains: spectroscopic evidence for the formation of beta-turns in the gas phase. J Am Chem Soc 127:707–712
109. Chin W, Dognon J, Canuel C, Piuzzi F, Dimicoli I, Mons M, Compagnon I, Von Helden G, Meijer G (2005) Secondary structures of short peptide chains in the gas phase: double resonance spectroscopy of protected dipeptides. J Chem Phys 122:054317
110. Nandel FS, Jaswal R (2007) New type of helix and 2_7 ribbon structure formation in poly (delta)leu peptides: construction of a single-handed template. Biomacromolecules 8:3093–3101
111. Chin W, Piuzzi F, Dognon J, Dimicoli I, Tardivel B, Mons M (2005) Gas phase formation of a 310-helix in a three-residue peptide chain: role of side chain-backbone interactions as evidenced by IR-UV double resonance experiments. J Am Chem Soc 127:11900–11901
112. Fricke H, Funk A, Schrader T, Gerhards M (2008) Investigation of secondary structure elements by IR/UV double resonance spectroscopy: analysis of an isolated (beta)-sheet model system. J Am Chem Soc 130:4692–4698
113. Gloaguen E, Loquais Y, Thomas JA, Pratt DW, Mons M (2013) Spontaneous formation of hydrophobic domains in isolated peptides. J Phys Chem B 117:4945–4955
114. Yoon I, Seo K, Lee S, Lee Y, Kim B (2007) Conformational study of tyramine and its water clusters by laser spectroscopy. J Phys Chem A 111:1800–1807
115. Blom MN, Compagnon I, Polfer NC, von Helden G, Meijer G, Suhai S, Paizs B, Oomens J (2007) Stepwise solvation of an amino acid: the appearance of zwitterionic structures. J Phys Chem A 111:7309–7316
116. Sakota K, Kouno Y, Harada S, Miyazaki M, Fujii M, Sekiya H (2012) IR spectroscopy of monohydrated tryptamine cation: rearrangement of the intermolecular hydrogen bond induced by photoionization. J Chem Phys 137:224311
117. Biswal HS, Loquais Y, Tardivel B, Gloaguen E, Mons M (2011) Isolated monohydrates of a model peptide chain: effect of a first water molecule on the secondary structure of a capped phenylalanine. J Am Chem Soc 133:3931–3942
118. Fricke H, Schwing K, Gerlach A, Unterberg C, Gerhards M (2010) Investigations of the water clusters of the protected amino acid Ac-Phe-OMe by applying IR/UV double resonance spectroscopy: microsolvation of the backbone. Phys Chem Chem Phys 12:3511–3521
119. Solís D, Bovin NV, Davis AP, Jiménez-Barbero J, Romero A, Roy R, SmetanaJr K, Gabius H (2015) A guide into glycosciences: how chemistry, biochemistry and biology cooperate to crack the sugar code. Biochim Biophys Acta (BBA) General Subj 1850:186–235
120. Gabius H, André S, Jiménez-Barbero J, Romero A, Solís D (2011) From lectin structure to functional glycomics: principles of the sugar code. Trends Biochem Sci 36:298–313
121. Grindley TB (2008) In: Fraser-Reid B, Tatsuta K, Thiem J (eds) Springer, Berlin Heidelberg, pp 3–55
122. Ghazarian H, Idoni B, Oppenheimer SB (2011) A glycobiology review: Carbohydrates, lectins and implications in cancer therapeutics. Acta Histochem 113:236–247
123. Stanca-Kaposta EC, Gamblin DP, Cocinero EJ, Frey J, Kroemer RT, Fairbanks AJ, Davis BG, Simons JP (2008) Solvent interactions and conformational choice in a core N-glycan segment: gas phase conformation of the central, branching trimannose unit and its singly hydrated complex. J Am Chem Soc 130:10691–10696
124. Gabius HJ (2000) Biological information transfer beyond the genetic code: the sugar code. Naturwissenschaften 87:108–121
125. Cocinero EJ, Carcabal P, Vaden TD, Simons JP, Davis BG (2011) Sensing the anomeric effect in a solvent-free environment. Nature 469:76-U1400
126. Cocinero EJ, Çarçabal P, Vaden TD, Davis BG, Simons JP (2011) Exploring carbohydrate–peptide interactions in the gas phase: structure and selectivity in complexes of pyranosides with N-acetylphenylalanine methylamide. J Am Chem Soc 133:4548–4557
127. Cocinero EJ, Gamblin DP, Davis BG, Simons JP (2009) The building blocks of cellulose: the intrinsic conformational structures of cellobiose, its epimer, lactose,

and their singly hydrated complexes. J Am Chem Soc 131:11117–11123

128. Cocinero EJ, Stanca-Kaposta EC, Scanlan EM, Gamblin DP, Davis BG, Simons JP (2008) Conformational choice and selectivity in singly and multiply hydrated monosaccharides in the gas phase. Chem A Eur J 14:8947–8955
129. Simons JP (2009) CARB 24-conformational landscapes of polysaccharides and glycopeptides: spectroscopy and modeling of the building blocks. Mol Phys 107:2435–2458
130. Jockusch RA, Kroemer RT, Talbot FO, Simons JP (2003) Hydrated sugars in the gas phase: spectroscopy and conformation of singly hydrated phenyl b-d-glucopyranoside. J Phys Chem A 107:10725–10732
131. Jockusch RA, Talbot FO, Simons JP (2003) Sugars in the gas phase Part 2: the spectroscopy and structure of jet-cooled phenyl small beta]-D-galactopyranoside. Phys Chem Chem Phys 5:1502–1507
132. Talbot FO, Simons JP (2002) Sugars in the gas phase: the spectroscopy and structure of jet-cooled phenyl small beta]-D-glucopyranoside. Phys Chem Chem Phys 4:3562–3565
133. Robertson EG, Simons JP (2001) Getting into shape: conformational and supramolecular landscapes in small biomolecules and their hydrated clusters. Phys Chem Chem Phys 3:1–18
134. Peña I, Cabezas C, Alonso JL (2015) Unveiling epimerization effects: a rotational study of Î±-d-galactose. Chem Commun 51:10115–10118
135. Bermúdez C, Peña I, Mata S, Alonso JL (2016) Sweet structural signatures unveiled in ketohexoses. Chem Eur J 22:16829–16837
136. Alonso ER, Peña I, Cabezas C, Alonso JL (2016) Structural expression of exo-anomeric effect. J Phys Chem Lett 7:845–850
137. Alonso JL, Lozoya MA, Pena I, Lopez JC, Cabezas C, Mata S, Blanco S (2014) The conformational behaviour of free d-glucose-at last. Chem Sci 5:515–522
138. Cocinero EJ, Lesarri A, Écija P, Cimas A, Davis BG, Basterretxea FJ, Fernández JA, Castaño F (2013) Free fructose is conformationally locked. J Am Chem Soc 135:2845–2852
139. Cocinero EJ, Lesarri A, Écija P, Basterretxea FJ, Grabow J, Fernández JA, Castaño F (2012) Ribose found in the gas phase. Angew Chem Int Ed 51:3119–3124
140. Peña I, Kolesnikova L, Cabezas C, Bermúdez C, Berdakin M, Simao A, Alonso JL (2014) The shape of d -glucosamine. Phys Chem Chem Phys 16:23244–23250
141. Camiruaga A, Usabiaga I, Insausti A, Cocinero EJ, León I, Fernández JA (2017) Understanding the role of tyrosine in glycogenin. Mol BioSyst 2017(13):1709–1712
142. Usabiaga I, González J, León I, Arnaiz PF, Cocinero EJ, Fernández JA (2017) Influence of the anomeric effect in the intermolecular interactions of sugars. J Phys Chem Lett 8:1147–1151
143. Usabiaga I, González J, Arnaiz PF, León I, Cocinero EJ, Fernández JA (2016) Modeling the tyrosine-sugar interactions in supersonic expansions: the glucopyranose-phenol clusters. Phys Chem Chem Phys 18:12457–12465
144. Su Z, Cocinero EJ, Stanca-Kaposta EC, Davis BG, Simons JP (2009) Carbohydrate-aromatic interactions: a computational and IR spectroscopic investigation of the complex, methyl alpha-L-fucopyranoside center dot toluene, isolated in the gas phase. Chem Phys Lett 471:17–21
145. Fernández-Alonso del Carmen M, Cañada FJ, Jiménez-Barbero J, Cuevas G (2005) Molecular recognition of saccharides by proteins. insights on the origin of the carbohydrate-aromatic interactions. J Am Chem Soc 127:7379–7386
146. Simons JP, Davis BG, Cocinero EJ, Gamblin DP, Stanca-Kaposta EC (2009) Conformational change and selectivity in explicitly hydrated carbohydrates. Tetrahedron-Asymmetry 20:718–722
147. Cocinero EJ, Stanca-Kaposta EC, Dethlefsen M, Liu B, Gamblin DP, Davis BG, Simons JP (2009) Hydration of sugars in the gas phase: regioselectivity and conformational in N-acetyl glucosamine and glucose. Chem A Eur J 15:13427–13434
148. Usabiaga I, Camiruaga A, Insausti A, Çarçabal P, Cocinero EJ, León I, Fernández JA (2018) Phenyl-(beta)-D-glucopyranoside and phenyl-(beta)-D-galactopyranoside dimers: small structural differences but very different interactions. Front Phys 6:3

José A. Fernández is Professor in the Department of Physical Chemistry of the University of the Basque Country (UPV/EHU). His main research activity is in laser spectroscopy in supersonic expansions of systems of biological interest, from nucleobases to drugs and from isolated molecules to large clusters. He has also a research line in molecular imaging of lipids in tissue sections in collaboration with several hospitals and international institutions.