

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 18671–18678

www.rsc.org/pccp

PAPER

# Carbohydrate hydration: heavy water complexes of $\alpha$ and $\beta$ anomers of glucose, galactose, fucose and xylose†

Nitzan Mayorkas,<sup>ab</sup> Svemir Rudić,<sup>b</sup> Emilio J. Cocinero,<sup>c</sup> Benjamin G. Davis<sup>d</sup> and John P. Simons\*<sup>b</sup>

Received 19th July 2011, Accepted 6th September 2011

DOI: 10.1039/c1cp22348h

The singly and doubly hydrated complexes of the  $\alpha$  and  $\beta$  anomers of a systematically varied set of monosaccharides, O-phenyl D-gluc-, D-galacto-, L-fuco- and D-xylopyranoside, have been generated in a cold molecular beam and probed through infrared-ultraviolet double resonance ion-dip (IRID) spectroscopy coupled with quantum mechanical calculations. A new ‘twist’ has been introduced by isotopic substitution, replacing H<sub>2</sub>O by D<sub>2</sub>O to separate the carbohydrate (OH) and hydrate (OD) vibrational signatures and also to relieve spectral congestion. The new spectroscopic and computational results have exposed subtle aspects of the intermolecular interactions which influence the finer details of their preferred structures, including the competing controls exerted by co-operative hydrogen bonding, bi-furcated and OH- $\pi$  hydrogen bonding, stereoelectronic changes associated with the anomeric effect, and dispersion interactions. They also reassert the operation of general ‘working rules’ governing conformational change and regioselectivity in both singly and doubly hydrated monosaccharides.

## Introduction

Non-covalent interactions and the conversations between biomolecules and their environment are part of the essential machinery of all living organisms and their interactions with small drug molecules lie at the heart of pharmaceutical science. Laser evaporation and interrogation of small biomolecules into seeded molecular beams provides a powerful strategy for probing their structures, preferred conformations and interactions, both inter- and intramolecular, ‘in vitro’.<sup>1</sup> Combined with mass-selected, double resonance IR-UV spectroscopy, and coupled with molecular mechanics, DFT and *ab initio* calculations, it enables the preparation, identification and structural assignment of their conformationally-selected molecular complexes.<sup>2–7</sup> When their partners are water molecules, this strategy also provides an entry into the ‘no-man’s land’

between isolated biomolecules, probed at low temperatures in the gas phase, and solvated molecules dissolved in aqueous solution at ambient temperatures.<sup>8,9</sup>

The influence of bound water molecules on the conformational landscapes of an expanding number of neutral amino acids,<sup>10–16</sup> small peptides<sup>17–19</sup> and especially carbohydrates<sup>20–26</sup> has been aided by the extreme sensitivity of their OH (and NH) vibrational spectra to hydrogen bonded solute-solvent interactions. In the latter case, it has shown how water binding can change the conformation of the substrate and helped to establish a set of ‘working rules’ governing the preferred binding sites in a representative series of biologically important monosaccharides.<sup>23,25</sup> Very recently, the differing OH vibrational signatures of doubly hydrated O-phenyl  $\alpha$ - and  $\beta$ -D-mannopyranoside<sup>26</sup> (and also carbohydrate-peptide complexes<sup>27</sup>) have exposed small structural differences arising from the stereoelectronic changes associated with the anomeric effect. The two water molecules were bound separately between OH4 and O6, and OH6 and the (axial) O2 sites, see Scheme 1, located on either side of the exocyclic hydroxymethyl group, crucially with the second one lying above the key endocyclic oxygen atom, O5. There it could act as a remarkably sensitive reporter, able to sense subtle stereoelectronic changes affecting the lone pair electron density on O5 and expose them through the resultant fine-tuning of the intermolecular hydrogen-bonds.

The sensitivity of this investigation<sup>26</sup> was enhanced further by substituting D<sub>2</sub>O for H<sub>2</sub>O, in order to isolate the

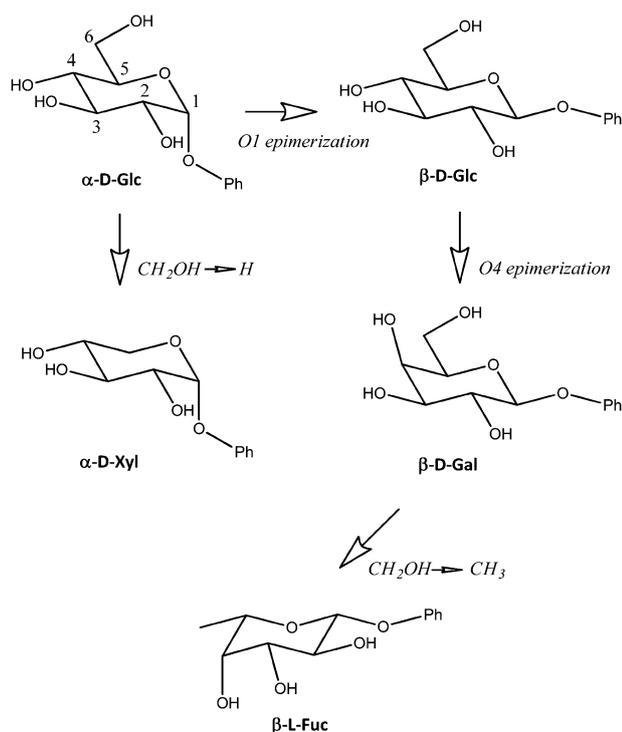
<sup>a</sup> Department of Physics, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel. E-mail: Mayorkas@bgu.ac.il

<sup>b</sup> Chemistry Department, University of Oxford, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK. E-mail: John.Simons@chem.ox.ac.uk

<sup>c</sup> Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco, (UPV-EHU), Apartado 644, E-48940, Bilbao, Spain. E-mail: Emiliojose.Cocinero@ehu.es

<sup>d</sup> Chemistry Department, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK. E-mail: Ben.Davis@chem.ox.ac.uk

† Electronic supplementary information (ESI) available: Experimental RZPI spectra and calculated relative energies, structures, atomic coordinates and vibrational frequencies. See DOI: 10.1039/c1cp22348h



**Scheme 1** Schematic representations and structural relationships of the selected set of monosaccharides.

carbohydrate (OH) bands from the water (OD) bands and also to relieve spectral congestion, a successful strategy which has prompted the new experimental and computational investigations presented here. They were designed initially, to seek further evidence of anomeric effects in the gas phase and were focused on two related systems: doubly (heavy water) hydrated O-phenyl  $\alpha$ - and  $\beta$ -D-gluco- and galactopyranoside [ $\alpha/\beta$ -PhGlc-(D<sub>2</sub>O)<sub>2</sub>], where again, the two water molecules should be located separately on each side of the hydroxymethyl group, and singly hydrated O-phenyl  $\alpha$ - and  $\beta$ -D-galactopyranoside [ $\alpha/\beta$ -PhGal-(D<sub>2</sub>O)], where binding at the preferred (4,6) site should be blocked by the (axial) OH<sub>4</sub> → O<sub>6</sub> hydrogen-bond. In each case, a bound water molecule would be expected to be located close to the anomeric site. The scope of the new investigations was subsequently extended by expanding the earlier studies of carbohydrate hydration (by H<sub>2</sub>O)<sup>20–25</sup> to include heavy water complexes of the  $\alpha$  and  $\beta$  anomers of the pentose carbohydrate, O-phenyl D-xylopyranoside ( $\alpha/\beta$ -PhXyl-(D<sub>2</sub>O)<sub>1,2</sub>) and the deoxy carbohydrate, O-phenyl  $\alpha$ -L-fucopyranoside ( $\alpha$ -PhFuc-(D<sub>2</sub>O)<sub>1,2</sub>): (the choice of L- rather than D-fucose reflects its biological relevance). Unlike glucose or galactose however, neither xylose (an analogue of glucose) nor fucose (6-deoxy galactose) incorporates an exocyclic hydroxymethyl group, see Scheme 1.

The new results have established the experimental advantages of substituting D<sub>2</sub>O for H<sub>2</sub>O: they have revealed subtle balancing controls exerted by 'fine-tuned' hydrogen bonding, their influence on the carbohydrate conformations, and the way in which they may mask the potential influence of stereo-electronic factors, associated with the anomeric effect, on hydrated structures. They have established the conformational structures of O-phenyl  $\alpha$ -D-gluco- and galactopyranoside and

confirmed many of the earlier conformational assignments based on investigations of (H<sub>2</sub>O) hydrated monosaccharides,<sup>21–25</sup> while reasserting the validity of the 'working rules' and the site-selective control exerted by the exocyclic hydroxymethyl group, when present. Finally, they suggest a possible rationale for the changes in the relative anomeric populations recorded in aqueous solution.<sup>28</sup>

## Experimental section

### Materials

The  $\alpha$ - and  $\beta$ -anomers of the O-phenyl pyranosides were obtained commercially (Aldrich) or in some cases, synthesized following methods described previously.<sup>21,23</sup> Their hydrated (D<sub>2</sub>O) complexes were generated in the gas phase using a combination of pulsed ns laser ablation (Nd:YAG, 1064 nm) and molecular beam procedures. Ground powdered samples of the carbohydrate were thoroughly mixed with graphite powder (80% sample:20% graphite, w/w), deposited as a thin uniform surface layer on a graphite substrate, and placed in a vacuum chamber close to and just below the exit of a pulsed, cylindrical nozzle expansion valve (0.8 mm diameter). The carbohydrates, desorbed by laser evaporation from the surface, were entrained and cooled in an expanding argon jet (~4 bar backing pressure) seeded with D<sub>2</sub>O, before passing into the detection chamber through a 2 mm diameter skimmer to create a collimated molecular beam. This was crossed by pulsed tunable IR and UV laser beams in the extraction region of a linear, Wiley-McLaren time-of-flight mass spectrometer (R. M. Jordan).

### Spectroscopy

Mass-selected resonant two photon ionization spectra of individual molecular complexes were recorded using a frequency-tripled pulsed Nd:YAG-pumped dye laser (Sirah). Their conformer-specific vibrational spectra were subsequently recorded in the OH and OD stretch regions, through infrared ion dip (IRID) spectroscopy, using IR radiation tuned over the range 2200–3800 cm<sup>-1</sup> and UV radiation tuned onto selected resonant two-photon ionization (R2PI) absorption bands. The first laser (IR) was pulsed at 5 Hz while the second one (UV), delayed by ~150 ns, ran at 10 Hz to allow subtraction of the background signals. The IR radiation (line-width 2–3 cm<sup>-1</sup>, ~5–10 mJ/pulse) was provided by the idler output of an OPO/OPA laser system (LaserVision), pumped by a pulsed Nd:YAG laser (Continuum, Surelite II). Several spectra, typically  $\geq 5$ , were recorded and averaged to achieve acceptable signal-to-noise levels.

### Computation

Analysis of the experimental data began with a series of unrestricted surveys of the large number of possible complex structures generated through molecular mechanics conformational search. This was done using the mixed Monte Carlo multiple minimization and Large Scale low mode method, as implemented in the MacroModel software (Schrödinger), until no additional new structures were obtained. All conformations with relative energies  $\leq 10$  kJ mol<sup>-1</sup>, as well as a number of

relevant higher energy structures, were subsequently submitted for geometry optimization using the Gaussian 03 suite of programs<sup>29</sup> and density functional theory (DFT) at the B3LYP/6-31+G(d) or in some cases where dispersion interactions were potentially influential, at the M05-2X and M06-2X/6-31+G(d,p) level. This led to a new set of relative energies (corrected for zero point energies), molecular structures and (harmonic) vibrational spectra which could then be compared with experiment. More accurate energies were calculated for the optimized DFT structures, at the MP2/6-311++G(d,p) level of theory. All energies were corrected for zero point energy using the unscaled DFT harmonic frequencies. Structural assignments were based primarily on the level of correspondence between the experimental and computed wavenumbers of the carbohydrate (OH) and water (OD) vibrational bands, scaled by the recommended<sup>30</sup> appropriate ‘anharmonicity’ factor, 0.9733 (B3LYP/6-31+G(d)), to bring them into better accord with experiment and secondly, on their calculated relative energies. The best agreement between experiment and theory for the most strongly populated complexes, was almost always obtained for the calculated minimum energy structures. The full list of relative energies, optimized geometries and harmonic vibrational frequencies for the low energy structures is given as Electronic Supplementary Information.†

### Notation

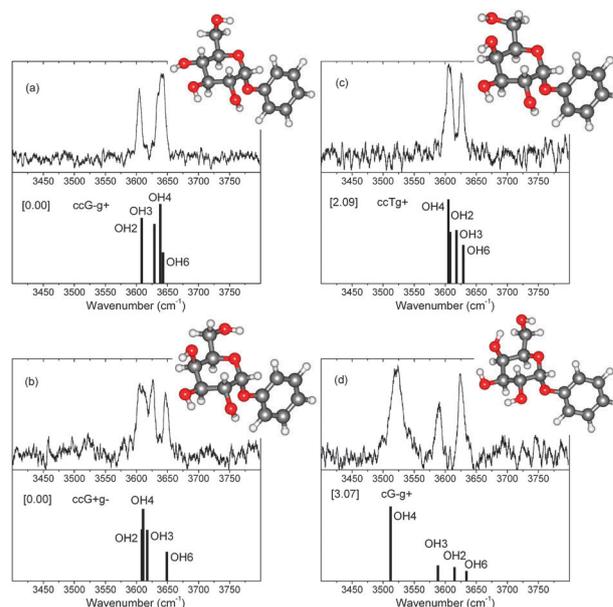
The designations, ccG + g–, cTg +, and so forth, indicate the conformation of the carbohydrate. Briefly, “cc” (“c”) indicates a counter-clockwise (clockwise) orientation of the peripheral OH groups, OH4 → OH3 → OH2 → O1 (O1 → OH2 → OH3 → OH4); the letters (G +, G–) or (T), indicate the *gauche* or *trans* orientation of the hydroxymethyl group with respect to the ring oxygen; and the final letters, g +, g– or t indicate the corresponding orientation of its hydroxyl group, OH6. In the case of the hydrated structures the insertion site of the water molecule is indicated by adding “ins(position)” to the nomenclature of the bare molecule; for example, ins(4,6) indicates a water molecule inserted between OH4 (acting as a hydrogen-bond donor) and OH6 (acting as the acceptor).

## Results and discussion

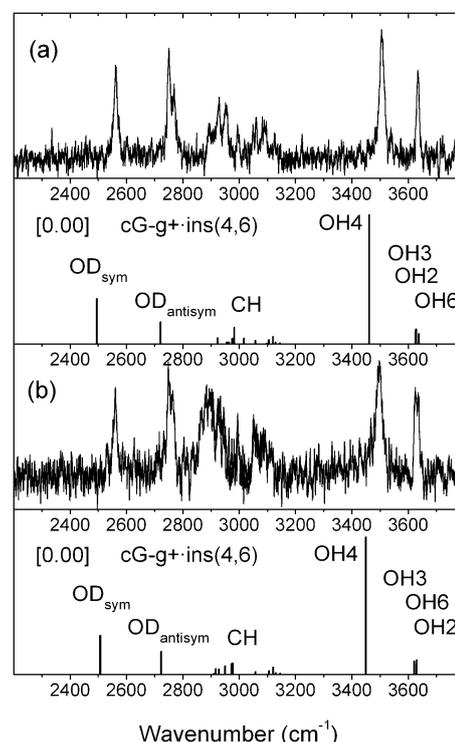
### O-phenyl $\alpha$ - and $\beta$ -D-glucopyranoside·(D<sub>2</sub>O)<sub>0,1,2</sub>

The experimental IRID and calculated vibrational spectra of  $\alpha$ -PhGlc and its singly and doubly hydrated complexes,  $\alpha$ -PhGlc·(D<sub>2</sub>O)<sub>1,2</sub>, are shown in Fig. 1(a) and Fig. 2 and 3, together with those of the corresponding hydrates of the  $\beta$  anomer,  $\beta$ -PhGlc·(D<sub>2</sub>O)<sub>1,2</sub> - their structures are shown later, in Fig. 7 and 8. The R2PI spectra, and the calculated relative energies, structures and vibrational spectra of each of their low lying conformers, are presented as Electronic Supplementary Information.

Unlike  $\beta$ -PhGlc, which can be detected in each of its three lowest energy conformations, ccG + g– (major), ccG–g + (minor) and ccTg + (trace),<sup>21</sup> the  $\alpha$  anomer displayed a somewhat different conformational landscape. Its IRID spectrum did not vary with the choice of R2PI probe wavenumber and it corresponded closely to the calculated vibrational spectrum of

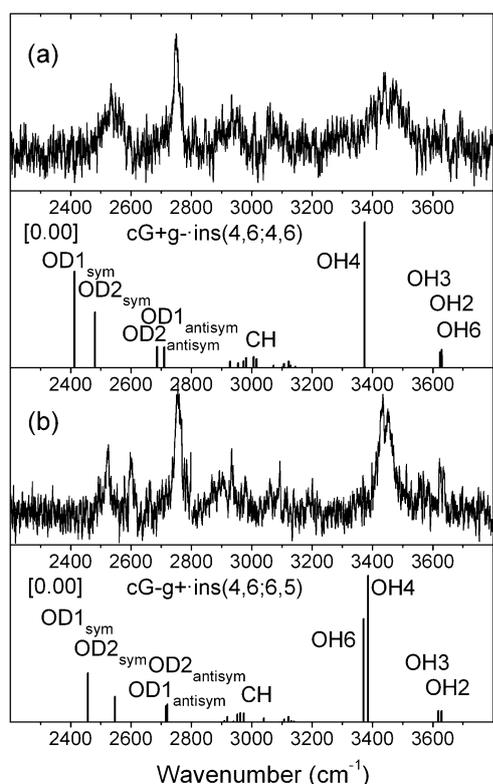


**Fig. 1** Experimental and calculated IRID and IR spectra, structures and relative energies (kJ mol<sup>-1</sup>) of the low lying conformers of (a)  $\alpha$ -PhGlc and (b,c,d)  $\alpha$ -PhGal.



**Fig. 2** Experimental and calculated IRID and IR spectra, structures and relative energies (kJ mol<sup>-1</sup>) of (a)  $\alpha$ -PhGlc·D<sub>2</sub>O and (b)  $\beta$ -PhGlc·D<sub>2</sub>O; their structures are shown in Fig. 7.

the ccG–g + conformer, also shown in Fig. 1. This was predicted to be the global minimum conformation (*cf.* the most strongly populated conformer in the  $\beta$  anomer, ccG + g–). The ccG + g– and ccTg + conformers, calculated to lie at energies respectively 1.7 and 2.8 kJ mol<sup>-1</sup> above the global minimum, could not be detected.



**Fig. 3** Experimental and calculated IRID and IR spectra, structures and relative energies ( $\text{kJ mol}^{-1}$ ) of (a)  $\alpha$ -PhGlc-( $\text{D}_2\text{O}$ )<sub>2</sub> and (b)  $\beta$ -PhGlc-( $\text{D}_2\text{O}$ )<sub>2</sub>; their structures are shown in Fig. 8.

The IRID spectra of the two singly hydrated anomers,  $\alpha/\beta$ -PhGlc- $\text{D}_2\text{O}$ , were identical, see Fig. 2, in contrast to those of the bare molecules, and they were in close correspondence with the calculated spectra associated with their (identical) global minimum structures,  $\text{cG-g+ins}(4,6)$ , shown later, in Fig. 7. In both cases, the spectra reflected the altered carbohydrate conformation, from  $\text{ccG-g+}$  ( $\alpha$  anomer) or  $\text{ccG+g-}$  ( $\beta$  anomer), to  $\text{cG-g+}$ , a change which creates a favourable water binding pocket located at the (4,6) site. The relative energy of the  $\text{cG-g+ins}(4,6)$  structure in the hydrated  $\alpha$  anomer, was predicted to be  $6.25 \text{ kJ mol}^{-1}$  below that of its nearest neighbour,  $\text{ccG-g+ins}(4,3)$ . The  $\text{cG-g+}$  conformer was not populated in either of the isolated anomers—their relative energies were predicted to be  $7.3 \text{ kJ mol}^{-1}$  ( $\alpha$  anomer) and  $9.7 \text{ kJ mol}^{-1}$  ( $\beta$  anomer) above their corresponding minima—but the binding energies associated with insertion of the water molecule into the (4,6) site, estimated to be  $\sim 30\text{--}40 \text{ kJ mol}^{-1}$ , should be more than enough to offset the barriers associated with the conformational change.

The good accord between the experimental and computed vibrational spectrum of  $\beta$ -PhGlc-( $\text{D}_2\text{O}$ )<sub>2</sub>, shown in Fig. 3, confirmed the earlier<sup>25</sup> structural assignment of the corresponding ( $\text{H}_2\text{O}$ )<sub>2</sub> complex to the global minimum conformation,  $\text{cG-g+ins}(4,6;6,5)$ . However, as well as the hydrogen bond linking the second water molecule to the endocyclic O5 atom, a closer examination of its structure, shown in Fig. 8, revealed an additional weak hydrogen bond linking it to the anomeric oxygen, O1, with a bond length,  $r(\text{O}(\text{H})\text{D}\cdots\text{O}1) \sim 2.81 \text{ \AA}$ .

Given the identical structures of the monohydrates, the  $\alpha$  and  $\beta$  di-hydrates might also have been expected to adopt similar structures but in contrast to the mono-hydrates, the IRID spectra of the doubly hydrated complexes were *not* identical, see Fig. 3. DFT calculations for the  $\alpha$  di-hydrate, initially conducted using the B3LYP functional alone, predicted a global minimum associated with a  $\text{cG-g+ins}(4,6;4,6)$  structure, with the two water molecules inserted as a dimer between OH4 and OH6, rather than separately,  $\text{ins}(4,6;6,5)$ . When the relative energy of the  $\alpha$  di-hydrate was recalculated, using single point MP2 theory to take account of dispersion interactions, a new global minimum energy structure,  $\text{cG+g-ins}(4,6;4,6)$ , was identified, shown later in Fig. 8(a). The two water molecules were still bound as a dimer but their spatial disposition was altered by the rotation of the hydroxymethyl group, from  $\text{G-g+}$  to  $\text{G+g-}$ , to bring OH6 and also one of the water molecules closer into the local environment of the phenyl group, offering the possibility of  $\text{OH} \rightarrow \pi$  hydrogen bonding. However, subsequent optimized DFT calculations employing the hybrid meta-GGA (Generalized Gradient Approximation) M05-2X or M06-2X functionals, specifically designed to include dispersion interactions,<sup>31</sup> led to the same result. Since the experimental and calculated vibrational spectra were also in reasonable accord, Fig. 3(a), it was identified as the preferred assignment. When the phenyl group was replaced by a methyl group, calculations for the  $\alpha$  di-hydrate (conducted at the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) level) predicted a quite different scenario; the water dimer was now inserted at the (3,2) site, remote from the methyl ‘tag’, to create the  $\text{ccG+g-ins}(3,2;3,2)$  structure shown in Fig. 8(c). In the phenylated  $\alpha$  di-hydrate,  $\beta$ -PhGlc-( $\text{D}_2\text{O}$ )<sub>2</sub>, the hydrated glucose appears to interact with its phenyl ‘tag’, most likely through a combination of hydrogen bonded and dispersion interactions, involving the phenyl group, OH6 and one of the neighbouring water molecule(s), so that the ‘tag’ no longer behaves as a simple ‘spectator’. In the  $\beta$  di-hydrate,  $\beta$ -PhGlc-( $\text{D}_2\text{O}$ )<sub>2</sub>, where the water molecules insert separately, the relative energy of the corresponding water dimer insertion structure,  $\text{cG+g-ins}(4,6;4,6)$ , was calculated to be  $9.5 \text{ kJ mol}^{-1}$  above the global minimum  $\text{cG-g+ins}(4,6;6,5)$ .

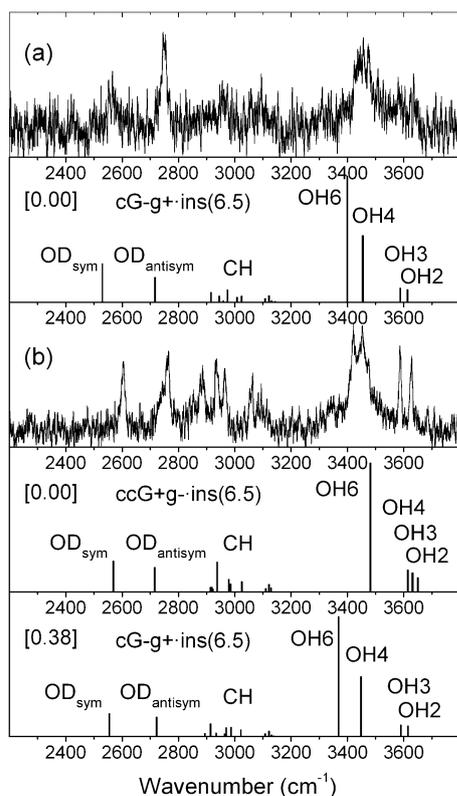
Why should the structures of the doubly hydrated  $\alpha$  and  $\beta$  anomers of PhGlc differ when those of their mono-hydrates are identical? The key lies in the location of the second bound water molecule. In the  $\text{ins}(4,6;6,5)$  di-hydrate structure, the additional hydrogen bond to O1 can only occur in the  $\beta$  complex; in the  $\alpha$  complex, the distance,  $r(\text{O}(\text{H})\text{D}\cdots\text{O}1)$  increases to  $> 4 \text{ \AA}$ . There is a fine balance between the influence of bifurcated H-bonding to O5 and O1 (not possible in the  $\alpha$  anomer); hydrogen bonding between the two water molecules themselves (only possible when they are bound as a dimer) or between the carbohydrate and the two separately bound water molecules; and in one particular case, a combination of hydrogen bonded and dispersion interactions involving the carbohydrate-water complex and the phenyl tag. This competition cannot influence the structures of the two mono-hydrates,  $\alpha/\beta$ -PhGlc- $\text{D}_2\text{O}$ , since the single bound water molecule occupies the favoured (4,6) site, remote from both O1 and O5.

### O-phenyl $\alpha$ - and $\beta$ -D-galactopyranoside-(D<sub>2</sub>O)<sub>0,1</sub>

The IRID spectrum of  $\alpha$ -PhGal is shown in Fig. 1b, c, and d where it can be compared with the corresponding spectrum of  $\alpha$ -PhGlc. The axial orientation of OH4 alters the conformational landscape and, unlike  $\alpha$ -PhGlc where only the minimum energy conformer (ccG-g+) was populated, three low lying conformers could be identified, ccG+g-, ccTg+ and cG-g+; their relative energies were predicted to be 0.00, 2.09 and 3.07 kJ mol<sup>-1</sup>, respectively. The IRID spectrum of its singly hydrated complex,  $\alpha$ -PhGal·D<sub>2</sub>O, shown in Fig. 4(a), is very similar to the calculated IR spectrum of its lowest energy structure, cG-g+·ins(6,5), shown in Fig. 7. The spectra associated with higher energy structures are presented as Electronic Supplementary Information.†

As anticipated earlier, the hydrogen bond linking the (axial) OH4 group to O6 blocks insertion into the (4,6) site and in the  $\alpha$  complex the water molecule opts for the most favoured alternative provided by the cG-g+ conformer. In the  $\alpha$  complex, the R2PI spectrum reflects the population of an additional conformer, ccG+g-ins(6,5),<sup>20</sup> but the more intense component is associated with the cG-g+·ins(6,5) structure. Its IRID spectrum is shown in Fig. 4(b) where it can be compared with the calculated spectra associated with the two alternatives; as expected, the cG-g+·ins(6,5) structure provides much better agreement. However, although the  $\alpha$  and  $\beta$  complexes are similar, neither their detailed structures nor their IRID spectra are identical.

Following the lead provided by the earlier investigations of  $\alpha/\beta$ -PhMan-ins(4,6;6,2),<sup>26</sup> the differences might have been



**Fig. 4** Experimental and calculated IRID and IR spectra, structures and relative energies (kJ mol<sup>-1</sup>) of (a)  $\alpha$ -PhGal·D<sub>2</sub>O and (b)  $\beta$ -PhGal·D<sub>2</sub>O; their structures are shown in Fig. 7.

attributed to an enhanced lone pair electron density on the endocyclic O atom, O5, in the  $\beta$  anomer—natural bond orbital calculations predict an increased electron density,  $\delta n(O5)_{\alpha \rightarrow \beta} \sim 0.012e$ , leading to a stronger OD  $\rightarrow$  O5 bond in the  $\beta$  complex. Unfortunately, the OD  $\rightarrow$  O5 bond is actually stronger in the  $\alpha$  complex:  $r(OD \cdots O5) = 1.89 \text{ \AA}$  ( $\alpha$  anomer) vs.  $1.96 \text{ \AA}$  ( $\beta$  anomer). This is reflected in the enhanced displacement of the corresponding OD vibrational band from  $\sim 2600 \text{ cm}^{-1}$  (in the  $\beta$  anomer) to  $\sim 2575 \text{ cm}^{-1}$  (in the  $\alpha$  anomer). The difference reflects the contribution of the bifurcated hydrogen bond which links the water molecule in the  $\beta$  complex to both O5 and O1,  $r(OD \cdots O1) = 2.82 \text{ \AA}$ . The additional bonding results in a slight tilting of the D<sub>2</sub>O plane towards O1 and away from O5. It also masks the potential influence of the stereoelectronic changes associated with the anomeric effect.

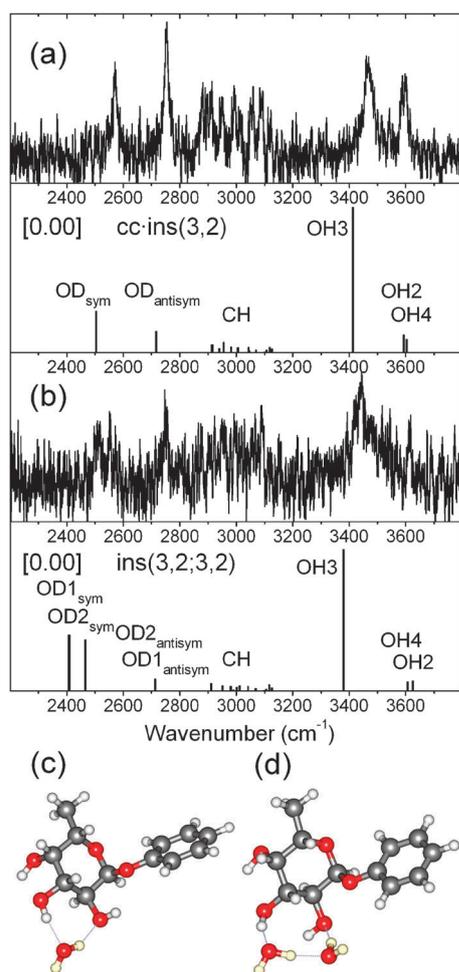
Although it was possible to detect weak R2PI signals in the  $\alpha/\beta$ -PhGal·(D<sub>2</sub>O)<sub>2</sub><sup>+</sup> ion channels their relative intensities were barely above the noise level and, unfortunately, they were too low to allow their IRID spectra to be recorded. With the (4,6) site blocked and the (6,5) site filled, binding of another water molecule appears to be less favoured than in  $\alpha/\beta$ -PhGlc, and also  $\alpha/\beta$ -PhMan, where both sites are accessible.

### O-phenyl $\alpha$ -L-fucopyranoside-(D<sub>2</sub>O)<sub>1,2</sub>

An earlier study of  $\alpha$ -PhFuc·H<sub>2</sub>O identified two structures,<sup>32</sup> one much more strongly populated than the other, but both based upon the preferred 'cc' conformation of the bare molecule. The IRID spectrum of the more strongly populated (heavy water) complex is shown in Fig. 5; it is in good agreement with the calculated spectrum associated with its minimum energy structure, cc-ins(3,2), confirming the earlier assignment. Fig. 5 also shows the IRID spectrum of  $\alpha$ -PhFuc·(D<sub>2</sub>O)<sub>2</sub>. It too, is in good accord with the spectrum associated with its calculated minimum energy structure which incorporates a water dimer, again inserted into the (3,2) site, the weakest link in the original counter-clockwise OH4  $\rightarrow$  OH3  $\rightarrow$  OH2  $\rightarrow$  O1 chain. While OH4 and OH3 are still oriented in a counter-clockwise 'cc' sense, the OH2 group is rotated to optimise its hydrogen bonded interaction with the second water molecule; this breaks the bond between OH2 and O1 and the 'cc' designation no longer applies.

### O-phenyl $\alpha$ - and $\beta$ -D-xylopyranoside-(D<sub>2</sub>O)<sub>1,2</sub>

Like  $\alpha$ -PhFuc, a previous R2PI/IRID investigation of  $\alpha$ -PhXyl·H<sub>2</sub>O identified the population of two distinct isomers;<sup>33</sup> Fig. 6 shows the corresponding IRID spectrum of the stronger isomer of the D<sub>2</sub>O complex. The calculated vibrational spectra associated with the two lowest energy structures, cc-ins(3,2) and cc-ins(4,3), are both very similar to the experimental spectrum. Although the minimum energy structure, cc-ins(3,2), might provide a slightly better fit—note the more widely spaced doublet at high wavenumber, assigned to the  $\sigma_2/\sigma_4$  vibrational modes—it is not possible to make a firm choice between the two. The  $\beta$  complexes follow a similar pattern though they differ in their choice of insertion site. The two lowest energy structures, separated by  $0.56 \text{ kJ mol}^{-1}$ , are both associated with insertion at the (2,1) site and differ only



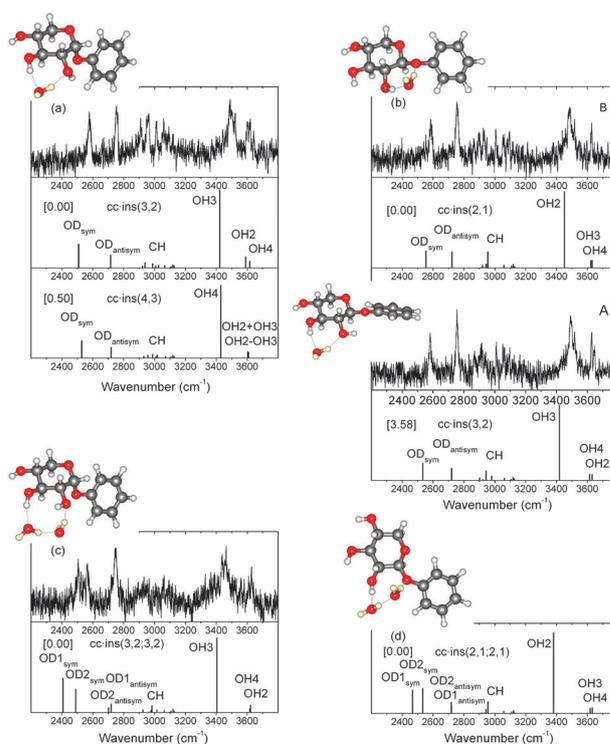
**Fig. 5** Experimental and calculated IRID and IR spectra and relative energies ( $\text{kJ mol}^{-1}$ ) of (a)  $\alpha$ -PhFuc- $\text{D}_2\text{O}$  and (b)  $\alpha$ -PhFuc- $(\text{D}_2\text{O})_2$ ; their structures are shown in (c) and (d).

in the relative orientations of the bound  $\text{D}_2\text{O}$  molecule. Their vibrational spectra are almost indistinguishable and the IRID spectra of the major and minor isomers of  $\beta$ -PhXyl- $\text{D}_2\text{O}$ , are in best correspondence with those associated with the cc-ins(2,1) and cc-ins(3,2) structures, as shown in Fig. 6(b). Like the two  $\alpha$ -complexes, the isomers are distinguished through the differing separations of the doublet at high wavenumber. Note, these preferred assignments differ a little from the earlier ones<sup>33</sup> based on the  $\text{H}_2\text{O}$  data, where both were identified with the cc-ins(2,1) structures.

The experimental ( $\alpha$  only) and calculated  $\alpha$  and  $\beta$  IRID spectra of the principal doubly hydrated complexes of the two anomers,  $\alpha/\beta$ -PhXyl- $(\text{D}_2\text{O})_2$ , both indicate insertion of a water dimer, once again at the (3,2) ( $\alpha$ ) or the (2,1) ( $\beta$ ), see Fig. 6c and d. (Unfortunately, the R2PI signal for the  $\beta$  complex was too weak to allow its IRID spectrum to be recorded at useful S/N levels).

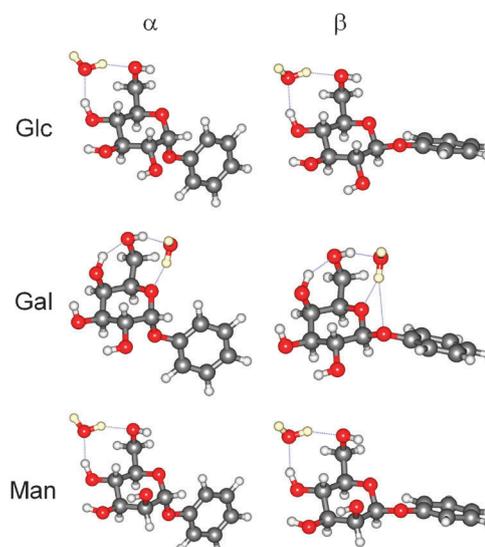
#### Anomeric differences, co-operative hydrogen bonding, conformational changes and regioselectivity

The preferred, experimentally-based structures of the singly hydrated anomers of PhGlc, PhGal and also PhMan, shown in



**Fig. 6** Experimental and/or calculated IRID and IR spectra, structures and relative energies ( $\text{kJ mol}^{-1}$ ) of (a,b)  $\alpha,\beta$ -PhXyl- $\text{D}_2\text{O}$  and (c,d)  $\alpha,\beta$ -PhXyl- $(\text{D}_2\text{O})_2$ .

Fig. 7, update an earlier comparison<sup>23</sup> where the data for  $\alpha$ -Glc/Gal were restricted to the calculated structures of the methyl derivatives,  $\alpha$ -MeGlc/Gal. Strikingly, in *all* six structures the carbohydrate hosts adopt the  $\text{cG-g+}$  conformation. The only carbohydrate to adopt this preferred conformation in the absence of hydration is  $\alpha$ -PhMan,<sup>20</sup> where it is favoured by the axial orientation of the OH2 group (which prevents OH2  $\rightarrow$  O1 bonding). The conformer is only weakly populated in  $\alpha$ -PhGal, and in all the other cases population of the



**Fig. 7** The preferred structures of the singly hydrated  $\alpha$  and  $\beta$  anomers of PhGlc, PhGal and PhMan.

cG–g+ conformers is prevented by their high relative energies. The situation is changed by hydration since relaxation into the cG–g+ conformation enables the bound water molecule (W) to access the favoured (4,6) site, or alternatively in  $\alpha/\beta$ -PhGal, the (6,5) site. This contributes to, and extends, the stabilizing chain of cooperative hydrogen bonds, either OH2 → OH3 → OH4 → W → OH6 → O5 (in  $\alpha/\beta$ -PhGlc) and  $\alpha/\beta$ -PhMan) or OH2 → OH3 → OH4 → OH6 → W → O5 (in  $\alpha/\beta$ -PhGal).

Fig. 8 compares the doubly hydrated structures of  $\alpha$ - and  $\beta$ -PhGlc, and also  $\alpha$ - and  $\beta$ -MeGlc, with those of  $\alpha$ - and  $\beta$ -PhMan, published earlier.<sup>26</sup> The carbohydrate unit in  $\beta$ -PhGlc·(D<sub>2</sub>O)<sub>2</sub> again adopts the cG–g+ conformation, to enable insertion of two separate water molecules on each side of the hydroxymethyl group. In this respect it resembles  $\alpha$ - and  $\beta$ -PhMan·(D<sub>2</sub>O)<sub>2</sub> but their hydrate structures are not the same. The second water molecule in both  $\alpha$ - and  $\beta$ -PhMan·(D<sub>2</sub>O)<sub>2</sub> inserts between OH6 and the axial OH2 group to create two near identical structures, allowing its interaction with the two carbohydrate anomers to be finely modulated by the differing lone pair electron repulsion between the neighbouring O(W) and O5(Man) atoms.<sup>26</sup> In contrast, in  $\alpha$ -PhGlc·(D<sub>2</sub>O)<sub>2</sub>, where OH2 is equatorial, the second water molecule binds to O5 and more weakly, to O1. In  $\alpha$ -PhGlc·(D<sub>2</sub>O)<sub>2</sub>, any structural comparisons which might have been based upon stereoelectronic effects alone, are compromised by the interaction with the phenyl ring, which alters both the hydrate structure and the carbohydrate conformation. The necessary conditions are only provided by  $\alpha/\beta$ -PhMan·(D<sub>2</sub>O)<sub>2</sub>.

In fucose (an analogue of galactose) and xylose (an analogue of glucose) the directing influence of the hydroxymethyl group is removed. The choice of preferred hydration sites is

now governed solely by the gain in cooperative hydrogen bonding achieved by selection of the weakest links in the original hydrogen bond chains.<sup>20,21,23,25</sup> The new data for the heavy water hydrated carbohydrate complexes confirm this ‘working rule’ in the mono-hydrates and reveal its continued operation when a second water molecule is introduced. In every case the two molecules insert as a dimer rather than separately, occupying the same preferred sites, (3,2) in fucose and (2,1) in xylose, as before.

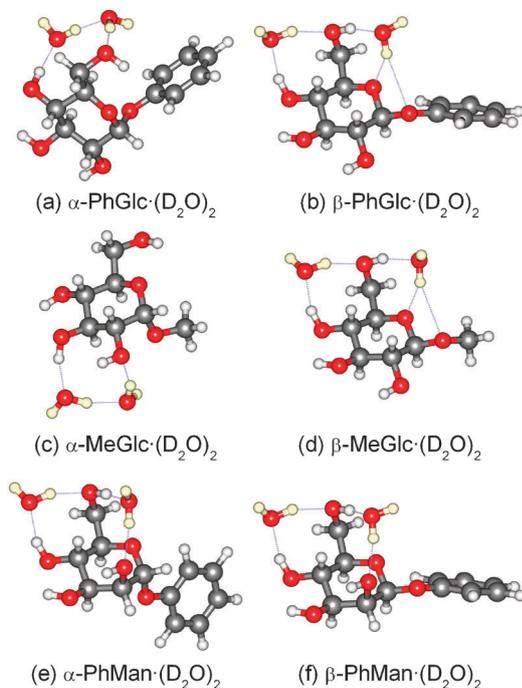
Although the anomeric effect favours population of the  $\alpha$  anomer in the isolated monosaccharides, in solution the relative populations of the  $\alpha$  and  $\beta$  anomers can change and in aqueous solutions of glucose, galactose and xylose, the  $\alpha$ : $\beta$  ratio changes from  $\sim 2:1$  to  $\sim 1:2$ . On the other hand, in mannose the  $\alpha$  anomer is still favoured and the ratio remains  $\sim 2:1$ .<sup>28</sup> This contrasting behavior bears a striking parallel with the variation in the pattern of hydrogen bonded interactions in the series,  $\alpha/\beta$ -PhGlc·(D<sub>2</sub>O)<sub>2</sub>,  $\alpha/\beta$ -PhGal·D<sub>2</sub>O and  $\alpha/\beta$ -PhXyl·(D<sub>2</sub>O)<sub>1,2</sub> vs.  $\alpha/\beta$ -PhMan·(D<sub>2</sub>O)<sub>2</sub>. The hydrates of  $\beta$ -PhGlc,  $\beta$ -PhGal and  $\beta$ -PhXyl each incorporate a hydrogen bonded interaction between one of the bound water molecules and the anomeric oxygen, O1. In hydrated mannose however, the corresponding hydrogen bond is directed towards the axial OH2 group, in both the  $\beta$  and the  $\alpha$  anomers: there is *no* interaction with O1 and the stereoelectronic influence is not submerged. A continued bias favouring these explicit local interactions in aqueous solution could provide a rationale for the  $\alpha$ : $\beta$  population ratios determined in aqueous solution.

## Conclusions

The advantage of substituting D<sub>2</sub>O for H<sub>2</sub>O in studying the influence of hydration on the conformational structures of isolated carbohydrates in the gas phase is now fully established. It allows the carbohydrate (OH) and hydrate (OD) vibrational signatures to be completely separated. There is no indication of isotope exchange in the gas phase and the new spectroscopic and computational results have exposed subtle aspects of the intermolecular interactions which govern the preferred structures of the hydrates. These include competing controls exerted by co-operative hydrogen bonding, bi-furcated and OH– $\pi$  hydrogen bonding, stereoelectronic changes associated with the anomeric effect, and dispersion interactions. The new data reassert the operation of the general ‘working rules’ that govern regioselectivity in both singly and doubly hydrated aldo-hexose and pentose carbohydrates. The specificity of the hydrogen bonded interactions also offers a possible rationale for the differing  $\alpha$ : $\beta$  population ratios found in aqueous solution.

## Acknowledgements

We thank the Leverhulme Trust for the award of an Emeritus Fellowship (J.P.S.), the Royal Society for a Wolfson Research Merit Award (B.G.D.), the Spanish Ministry (MICINN) for a ‘Ramón y Cajal’ Contract (E.J.C.), the support provided by STFC through the Laser Loan Pool (J.P.S.) and by EPSRC through an LSI Platform grant (B.G.D.). We also appreciate



**Fig. 8** Experimentally based structures of the doubly hydrated  $\alpha$  and  $\beta$  anomers of (a,b) PhGlc and (e,f) PhMan, and the corresponding calculated structures (c,d) of MeGlc.

discussions with Prof R. B. Gerber and the practical assistance provided by Dr Conor Barry.

## References

- J. W. Elam and D. H. Levy, *J. Phys. Chem. B*, 1998, **102**, 8113.
- T. S. Zwier, *J. Phys. Chem. A*, 2006, **110**, 4133.
- M. S. de Vries and P. Hobža, *Annu. Rev. Phys. Chem.*, 2007, **58**, 585.
- J.-P. Schermann, *Spectroscopy and Modelling of Biomolecular Building Blocks*, Elsevier, Amsterdam, 2008.
- J. P. Simons, *Mol. Phys.*, 2009, **107**, 2435.
- H. Fricke, A. Gerlach, C. Unterberg, M. Wehner, T. Schrader and M. Gerhards, *Angew. Chem., Int. Ed.*, 2009, **48**, 900.
- E. J. Cocinero, P. Çarçabal, T. D. Vaden, B. G. Davis and J. P. Simons, *J. Am. Chem. Soc.*, 2011, **133**, 4548.
- T. S. Zwier, *J. Phys. Chem. A*, 2001, **105**, 8827.
- J. P. Simons, *C. R. Chim.*, 2003, **6**, 17.
- L. C. Snoek, R. T. Kroemer and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2002, **4**, 2130.
- K. T. Lee, J. Sung, K. J. Lee, S. K. Kim and Y. D. Park, *J. Chem. Phys.*, 2002, **116**, 8251.
- P. Çarçabal, R. T. Kroemer, L. C. Snoek, J. P. Simons, J. M. Bakker, I. Compagnon, G. Meijer and G. von Helden, *Phys. Chem. Chem. Phys.*, 2004, **6**, 4546.
- J. L. Alonso, E. J. Cocinero, A. Lesarri, M. E. Sanz and J. C. Lopez, *Angew. Chem., Int. Ed.*, 2006, **45**, 3471.
- T. Ebata, T. Hashimoto, T. Ito, Y. Inokuchi, F. Altunsi, B. Brutschy and P. Tarakeshwar, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4783.
- M. N. Blom, I. Compagnon, N. C. Polfer, G. von Helden, G. Meijer, B. Paizs and J. Oomens, *J. Phys. Chem. A*, 2007, **111**, 7309.
- H. M. Kim, K. Y. Han, J. Park, G. S. Kim and S. K. Kim, *J. Chem. Phys.*, 2008, **128**, 041104.
- H. Zhu, M. Blom, I. Compagnon, A. M. Rijs, S. Roy, G. von Helden and B. Schmidt, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3415.
- H. Fricke, K. Schwing, A. Gerlach, C. Unterberg and M. Gerhards, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3511.
- H. S. Biswal, Y. Loquais, B. Tardivel, E. Gloaguen and M. Mons, *J. Am. Chem. Soc.*, 2011, **133**, 3931.
- P. Çarçabal, R. A. Jockusch, I. Hünig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin, I. Compagnon, J. Oomens and J. P. Simons, *J. Am. Chem. Soc.*, 2005, **127**, 11414.
- J. P. Simons, P. Çarçabal, B. G. Davis, D. P. Gamblin, I. Hünig, R. A. Jockusch, R. T. Kroemer, E. M. Marzluff and L. C. Snoek, *Int. Rev. Phys. Chem.*, 2005, **24**, 489.
- E. C. Stanca-Kaposta, D. P. Gamblin, E. J. Cocinero, J. Frey, R. T. Kroemer, B. G. Davis, A. J. Fairbanks and J. P. Simons, *J. Am. Chem. Soc.*, 2008, **130**, 10691.
- E. J. Cocinero, E. C. Stanca-Kaposta, E. M. Scanlan, D. P. Gamblin, B. G. Davis and J. P. Simons, *Chem.–Eur. J.*, 2008, **14**, 8947.
- E. J. Cocinero, D. P. Gamblin, B. G. Davis and J. P. Simons, *J. Am. Chem. Soc.*, 2009, **131**, 11117.
- E. J. Cocinero, E. C. Stanca-Kaposta, M. Dethlefsen, B. Liu, D. P. Gamblin, B. G. Davis and J. P. Simons, *Chem.–Eur. J.*, 2009, **15**, 13427.
- N. Mayorkas, S. Rudic, B. G. Davis and J. P. Simons, *Chem. Sci.*, 2011, **2**, 1128.
- E. J. Cocinero, P. Çarçabal, T. D. Vaden, J. P. Simons and B. G. Davis, *Nature*, 2011, **469**, 76.
- R. K. Schmidt, M. Karplus and J. W. Brady, *J. Am. Chem. Soc.*, 1996, **118**, 541.
- M. J. Frisch, *et al.*, *Gaussian 03, Revision C.02*, Gaussian, Inc, Pittsburgh, PA, 2003.
- Y. Bouteiller, J. C. Gillet, G. Grégoire and J.-P. Schermann, *J. Phys. Chem. A*, 2008, **112**, 11656.
- Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 289.
- P. Çarçabal, T. Patsias, I. Hünig, B. Liu, E. C. Kaposta, L. C. Snoek, D. P. Gamblin, B. G. Davis and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2006, **8**, 129.
- I. Hünig, A. J. Painter, R. A. Jockusch, P. Çarçabal, E. M. Marzluff, L. C. Snoek, D. P. Gamblin, B. G. Davis and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2474.