Quantum Chemical Studies of Carbohydrate Reactivity: Acid Catalyzed Ring Opening Reactions

Matthias Mohr,[†] Richard A. Bryce,[‡] and Ian H. Hillier*,[†]

Department of Chemistry, University of Manchester, Oxford Rd. M13 9PL, U.K., and School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Oxford Rd. M13 9PL, U.K.

Received: March 8, 2001; In Final Form: June 7, 2001

The appropriate quantum mechanical method for the investigation of enzymatic reactions, which involve oxocarbenium ions as reactive intermediates, is examined. 2-methoxy-tetrahydro-2H-pyran (**1**) was chosen as a model acetal for pyranose sugars, and its reactivity upon protonation of the glycosidic and the ring oxygen atom has been investigated using various density functional and post Hartree−Fock methods. Proton affinities calculated at the DFT levels of theory predict glycosidic protonation to be favorable by 2.5 kcal/mol. For ring-protonated **1**, among the density functionals, a strong dependence of the molecular structure on the density functional employed is found. Structures obtained with the BLYP and B3LYP functionals are at variance with those from the ab initio methods, MP2 and CCSD, as shown by differences in bond lengths of more than 0.4 Å for equivalent structures. By means of a valence bond analysis of the electron densities obtained at the DFT levels of theory, it is shown that this method dependence in this closed-shell species is caused by spurious self-interaction. This failure appears to be due to the subtle interplay between electron donating and accepting groups present in **1**. The BHLYP functional is found to perform best among the functionals under investigation, for describing the hypersurfaces for protonated pyranose sugars.

Introduction

The major significance of carbohydrates stems from their importance as biological building blocks, involved in for example the formation of bacterial peptidoglycan envelopes and the posttranslational modification of proteins. Carbohydrates are challenging systems to study theoretically, because of their dense polar functionality and inherent conformational flexibility. There are severe problems in predicting conformational equilibria in the gas phase[1] as well as condensed phase,[2] which are reflected in the large number of force field reparametrizations[3] and a number of extensive ab initio studies.[4,5] Most of these computational studies focus on predicting thermodynamic properties such as conformational equilibria, with little emphasis on the reactivity of sugars.[5] Sugar derived cationic oxocarbenium ions are of particular interest because they are believed to be key intermediates in enzymatic catalysis of a wide range of glycosyltransferases[6] and glycosidases.[7] The mechanism of the latter class of enzymes is known to proceed via protonation of the glycosidic oxygen atom, followed by transfer of the resulting oxocarbenium ion to a catalytic base located on the enzyme, or hydrolysis. However, the degree of concertedness of the bond cleaving and forming steps remains a matter of some debate.[8] A detailed picture, at the molecular level, of the reactivity of sugar derivatives in their native environments, such as aqueous solution or the binding pocket of an enzyme, can only be obtained by quantum chemical studies.

To choose the most appropriate quantum mechanical level, we here investigate the performance of several density functional theory (DFT), ab initio, and semiempirical methods to describe energetic and structural features of reaction intermediates arising from protonation of a model glycoside **1**. The primary objective of this study is to investigate whether these methods are suitable for use in subsequent hybrid quantum mechanical/molecular mechanical (QM/MM) investigations of enzyme mechanisms, which involve oxocarbenium intermediates.

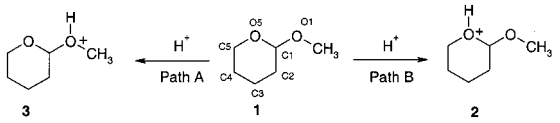

Figure 1. Schematic representation of the two distinct protonation sites of sugar model 2-methoxy-tetrahydro-2H-pyran **1**. Path A shows protonation of the glycosidic methoxy group, whereas path B shows the alternate ring protonation of **1**.

We report studies of the model acetal **1** which has the essential features of a pyranose sugar moiety with a glycosidic linkage (Figure 1). Hydroxyl groups have been omitted in order to restrict conformational space and to focus on the reactivity involving the glycosidic bond. Despite the well-known fact that hydroxyl substituents in the 2 position of pyranose sugars lead to approximately a 1000-fold increase in the rate of glycolysis, activation parameters of glycolysis of pyranoses and 2-deoxy-pyranoses indicate that both reactions proceed via the same mechanism.[9] Compound **1** was chosen as a model system to investigate the suitability of several theoretical methods in describing the potential energy surface of the basic protonation and accompanying bond-breaking steps. Special emphasis was put on several hybrid DFT methods, which differ in the fraction of exact exchange included in the functional. These results are compared to conventional ab initio methods which include correlation effects, such as MP2 or CCSD.

* To whom correspondence should be addressed.
[†] Department of Chemistry, University of Manchester.
[‡] School of Pharmacy and Pharmaceutical Sciences, University of Manchester.

10.1021/jp010901+ CCC: $20.00
Published on Web 08/09/2001

Theoretical Methods

Over the past decade, generalized gradient approximation (GGA) density functional theory as well as Hartree−Fock hybrid functionals have developed into a widely applied approach for the investigation of the electronic structure of systems of chemical interest. In particular, the B3LYP functional[10] developed by Becke is an alternative to the perturbation theory of Møller and Plesset, because DFT methods in general are computationally less demanding. Soon after development of GGA functionals, it became apparent that problems arising from the description of exchange contributions to the total energy could be, at least partly, overcome by including various amounts of exact Hartree−Fock exchange. Indeed, cases are known in which hybrid functionals give superior results to MP4 or even QCISD ab initio methods.[11] However, recently there are an increasing number of reports in the literature on the limitations of hybrid functionals.[12] In particular, the dissociation behavior of symmetric diradicals exemplifies this deficiency.[13]

All DFT calculations were performed using the Gaussian 94 revision E.1[14] and ab initio calculations at the MP2 and CCSD level of theory with the Gaussian 98 revision A.7[15] suites of programs. The hybrid functionals BHLYP and B3LYP were employed as implemented in Gaussian 94. All structures were optimized at the indicated level of theory and basis set and were characterized as stationary points by the calculation of harmonic frequencies. In DFT calculations, an SCF convergence criterion of 10^{-8} and the standard 75-radial, 302-angular point integration grid was used. Analysis of the electron density was based on Mulliken partial charges.

Results

To investigate the general acid-catalyzed reactions of 2-methoxy-tetrahydro-2H-pyran (**1**), two distinct protonation sites have been taken into account: the ring oxygen, according to sugar numbering O5, and the methoxy oxygen atom, referred to as O1. The two different reactions are shown in Figure 1 (Path A and B). Either of the two reactions can proceed through a variety of conformations of the ring system. Theoretical investigation[4] of the conformational space of **1** revealed the α-anomers in general to be more stable by approximately 1 kcal/mol in the gas phase. The most favorable conformation of the glycosidic bond is a trans orientation to the C1−C2 bond and gauche orientation to the O1−C1 bond. NMR-examinations of **1** confirm the trans/gauche conformer depicted in Figure 2 to be the most stable α conformer.[16]

Protonation of the Ring Oxygen. Investigation of the conformational space of **2** reveals structure **2a** depicted in Figure 2 to be the most stable α conformer[17] as observed for unprotonated species **1**. As expected, it is found that chairlike conformations in general are more stable compared to the corresponding boat conformations. For the α anomers, it turns out that axial protonation of **1** is slightly less favorable compared to equatorial protonation. Therefore, conformer **2a** was chosen to investigate the suitability of several methods to predict structural and energetic features of **2**.

Table 1 summarizes the results obtained at the indicated level of theory for the minimum energy conformer **2a**. The 6-31+G(d,p) basis was chosen according to the observation[18] that, for investigations of conformational equilibria of unprotonated pyranose sugars, reliable structural results are obtained with the 6-31G(d,p) basis set. A diffuse basis function was added because it was shown that the conformational preference at the anomeric center of **1** could not be predicted properly at the HF/6-31G(d) level of theory.[4] Usually, the conformational preference in **1** is biased toward the α anomer by approximately 1 kcal/mol when the diffuse function is omitted. It is expected that upon protonation hyperconjugative effects become important compared to the uncharged model **1**, as the bond between O5 and C1 is expected to weaken,[19] so that inclusion of a diffuse basis function becomes necessary. Furthermore inclusion of diffuse basis functions considerably reduces basis set superposition errors in comparable systems.[18]

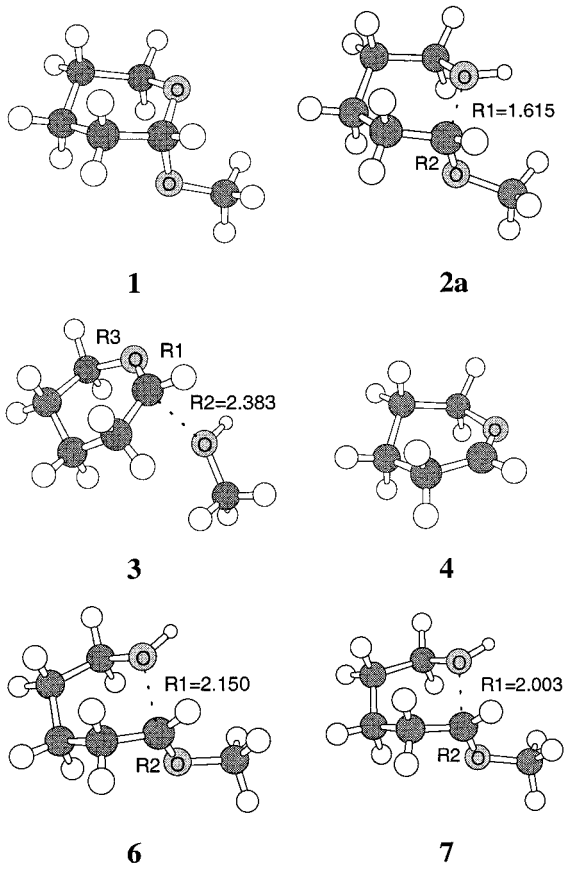

Figure 2. Schematic representation of the most stable conformer of 2-methoxytetrahydropyran **1**, the bond lengths R1, R2, and R3 of protonated 2-methoxy-tetrahydro-2H-pyran structures **2a**, **3**, **6**, **7**, and oxocarbenium ion **4**. Bond lengths reported here correspond to the BHLYP/6-31+G(d,p) level of theory and are given in angstroms.

TABLE 1: Structural Parameters of Structures 2a and 6 (Figure 2) and Gas Phase Proton Affinities (PA) for Protonation of O5

6-31+G(d,p)	R1 [Å]	R2 [Å]	PA [kcal/mol]	PA+ZPE [kcal/mol]	structure
BLYP	2.100	1.301	212.4	205.8	**2a**
B3LYP	1.954	1.293	209.5	203.0	**2a**
BHLYP	1.615	1.323	208.6	200.9	**2a**
BHLYP	2.150	1.261	-	-	**6**
HF	1.579	1.326	211.4	203.4	**2a**
HF	2.352	1.242	-	-	**6**
MP2	1.643	1.337	206.5	198.9	**2a**

The data given in Table 1 show the strong dependence of the bond distance R1 on the method chosen. For the pure GGA functional BLYP,[20] this distance is 2.100 Å, which significantly decreases upon inclusion of Hartree−Fock exchange. For the B3LYP functional, which includes 20% HF exchange, R1 reduces to 1.954 Å. Moreover, at the Hartree−Fock and BHLYP[21] levels of theory, two distinct minima **2a** and **6** can be located, which differ in the length R1. In **6** at the HF/6-31+G(d,p) and BHLYP/6-31+G(d,p) level, R1 increases to 2.352 and 2.150 Å, respectively. For **2a**, at the HF level, a

TABLE 2: Basis Set Dependence of Structure 2a (Figure 2, All Values Given in Å)

BHLYP	R1	R2
6-31G(d)	1.614	1.323
6-31G(d,p)	1.612	1.323
6-31+G(d,p)	1.615	1.323
6-311G(d,p)	1.623	1.317
6-311++G(3df,2p)	1.611	1.317

TABLE 3: Structural Parameters for Glycosidic Protonated Species 3 Given in Å (Figure 2) and Reaction Energies for Protonation and Methanol Loss Given in kcal/mol (Figure 3)

6-31+				protonation		methanol loss	
G(d,p)	R1	R2	R3	PA	PA+ZPE	ΔE	ΔE+ZPE
BLYP	1.286	2.361	1.529	214.1	208.1	12.3	10.6
B3LYP	1.267	2.372	1.502	211.3	205.4	13.1	11.5
BHLYP	1.249	2.383	1.481	209.4	203.5	14.5	12.9
HF	1.237	2.357	1.482	216.3	210.4	13.3	11.8
MP2	1.266	2.378	1.503	205.4	198.1	16.6	15.2
exp[a]					196.7		

[a] For tetrahydro-2H-pyran, taken from ref 22. See text for details.

slightly shorter bond length R1 of 1.579 Å is found compared to 1.615 Å at the BHLYP level. Inspection of other conformers[17] shows that the strong dependence of R1 upon the computational method is not limited to conformer **2a**, with the difference in R1 between the HF and BLYP method being at least 0.44 Å.

To exclude the possibility of artifacts due to an improper basis set, a variety of basis sets were employed to investigate the crucial bond lengths R1 and R2 of structure **2a** (Table 2). For the valence double-ζ basis sets, there is only minimal variation of the bond distance R1. The 6-31G(d) basis set yields an R1 of 1.614 Å, which lies between the value for the slightly more flexible basis sets 6-31G(d,p) and 6-31+G(d,p), which give R1 values of 1.612 and 1.615 Å, respectively. The largest value of R1 (1.623 Å) is found for the valence triple-ζ basis 6-311G-(d,p). The values for R2 found at these levels of theory are identical for all valence double-ζ basis sets. Increasing the basis set's flexibility to the 6-311G(d,p) and 6-311++G(3df,2p) basis sets, results in a slight shortening of R2 by 0.006 Å. The small deviations of R1 with respect to the choice of the basis show that the dependence of R1 upon the computational method is clearly not caused by an improper basis set. Therefore, for practical reasons, the 6-31+G(d,p) basis set which we use here is a compromise between accuracy and computational feasibility.

Protonation of the Glycosidic Oxygen. The second protonation site to be considered in model compound **1** is the O1 oxygen atom (Figure 1, path A). Table 3 summarises the geometrical and energetic results obtained at the indicated level of theory. In contrast to the effect observed for structure **2a**, the method dependence found for **3** is only moderate. Moreover, the trend found among the various density functionals in **1** is reversed.

Again, the shortest bond length is found for the Hartree−Fock method, which is expected for a method that does not include repulsive correlation contributions. In **3**, in contrast to **2a** among the density functional methods, the shortest distance for the exocyclic C−O bond R2 of 2.361 Å is found at the BLYP/6-31+G(d,p) level of theory. Upon gradual inclusion of Hartree−Fock exchange, this bond distance steadily increases, with the B3LYP functional yielding a distance of 2.372 Å and the BHLYP a value of 2.383 Å. The geometry calculated at the MP2 level yields an R2 value (2.378 Å) which lies between the corresponding B3LYP and BHLYP values.

Reaction Energies. Inspection of the proton affinities given in Tables 1 and 3 reveals only slight regioselectivity with regard

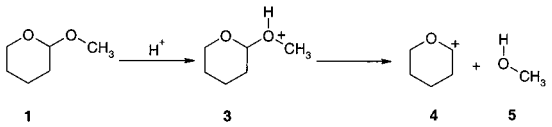

Figure 3. Schematic representation of the formation of oxocarbenium ion **4** upon protonation of the glycosidic oxygen O1 of **1** followed by elimination of methanol **5**.

to the protonation site. At the HF and DFT levels of theory, the ZPE-corrected gas-phase reaction energies for glycosidic oxygen protonation are predicted to be larger than those for ring protonation. For the BLYP, B3LYP, and BHLYP functionals, structure **3** is more stable than **2a** by 2.3, 2.4, and 2.6 kcal/mol, respectively. For the HF level, this difference is even larger and amounts to 7.0 kcal/mol. On the other hand, the relative stabilities of **2a** and **3** are reversed at the MP2 level of theory. Compound **3** is found to be less stable by 0.8 kcal/mol compared to **2a**. The closest related compounds for which experimental gas-phase proton affinities are available are tetrahydro-2H-pyran and 1,3-dioxane[22] and amount to 196.7 and 197.3 kcal/mol respectively. The small difference of only 0.6 kcal/mol suggests that the inductive effect of an α-oxygen substituent only slightly affects the proton affinity. Despite the quite dramatic changes in equilibrium structures of **2a** among the DFT methods, the differences in the proton affinity are only moderate. The largest value, 205.8 kcal/mol, is found at the BLYP/6-31+G(d,p) level of theory. Upon inclusion of HF exchange, the proton affinity is reduced to 203.0 and 200.9 kcal/mol, for the B3LYP and BHLYP functionals, respectively. The HF value (203.4 kcal/mol) is in close agreement with the B3LYP functional value. MP2 marks the lower end with a proton affinity of 198.9 kcal/mol. For structure **3**, the calculated proton affinities show a comparable trend to **2a**. Again, among the density functional methods, the largest proton affinity is found at the BLYP level of theory, which amounts to 208.1 kcal/mol and steadily decreases on going to the B3LYP and BHLYP functionals. Proton affinities of **3** at these levels are 205.4 and 203.5 kcal/mol respectively. In **3**, Hartree−Fock theory yields the largest value of 210.4 kcal/mol.

An additional reaction path has to be considered in system **3**. Upon formation of the protonated system **3**, in analogy to ring opening of **2a**, loss of methanol becomes a possible reaction (Figure 3). The reaction energies for dissociation into methanol **5** and oxocarbenium ion **4** are summarized in Table 3. The reaction is mildly endothermic, and the trend found for the reaction energies is reversed compared to that of proton affinities. Here, MP2 gives the largest interaction energy of 15.2 kcal/mol. In contrast to the trend found for the proton affinities, gradual inclusion of Hartree−Fock exchange in the density functional slightly increases the binding energy. The lowest value of 10.6 kcal/mol is found at the BLYP level of theory. The hybrid functionals B3LYP and BHLYP show slightly larger values of 11.5 and 12.9 kcal/mol, respectively. At the Hartree−Fock level, this energy is between the values from the hybrid functionals and amounts to 11.8 kcal/mol. The fact that for all levels of theory we have studied the bond distance R2 is much larger than 2 Å and the reaction energies are between 11 and 15 kcal/mol shows that the glycosidic protonated structure **3** can best be viewed as a loose complex between methanol **5** and oxocarbenium ion **4**, rather than a protonated 2-methoxy-tetrahydro-2H-pyran.

Because of the remarkable behavior of protonated ion **2a**, further investigations were performed by calculating the potential energy surface of the ring opening process of **2a**. The potential energy surface was obtained by restricting bond

TABLE 4: Relative Energies for Ring Opening in 2a[a]

R1 [Å]	HF	BHLYP	B3LYP	BLYP	MP2	AM1	CCSD//BHLYP	CCSD//BLYP	CCSD//MP2
1.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.70	0.44	0.20	−0.70	−1.58	0.03	0.64			
1.80	0.71	0.61	−0.93	−2.50	0.61	1.30	0.72		
1.90	0.44	0.87	−0.99	−3.06	1.25	1.61			
2.00	−0.27	0.95	−0.99	−3.36	1.75	1.64	1.23		
2.10	−1.09	0.92	−0.96	−3.46	2.10	1.58			
2.20	−1.76	0.94	−0.82	−3.39	2.39	1.62	1.15		
2.30	−2.12	1.12	−0.53	−3.12	2.74	1.86			
2.40	−2.12	1.51	−0.10	−2.70	3.25	2.35	1.60	2.06	1.67

[a] Calculated at the indicated level of theory with the 6-31+G(d,p) basis set. All values are given in kcal/mol. The energy at R1 = 1.6 Å was arbitrarily chosen to be zero. All structures were optimized with frozen bond distance R1, except for the CCSD single points. See text for details.

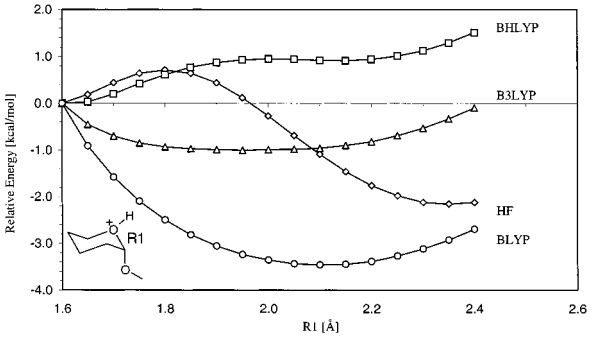

Figure 4. Plot of the relative energies at the indicated level of theory with the 6-31+G(d,p) basis set versus the reaction coordinate R1. For all methods, the geometry at R1 = 1.6 Å was arbitrarily chosen to be the reference geometry. All degrees of freedom except R1 are fully optimized.

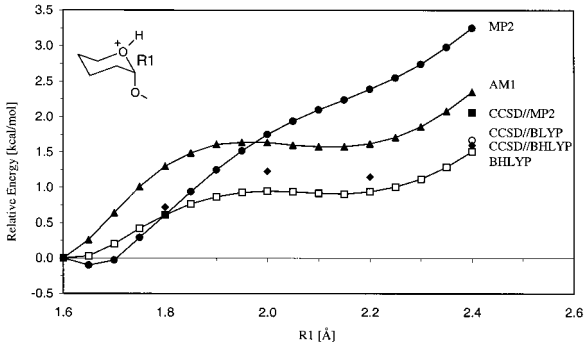

Figure 5. Plot of the relative energies at the indicated level of theory with the 6-31+G(d,p) basis set versus the reaction coordinate R1. For all methods, the geometry at R1 = 1.6 Å was arbitrarily chosen to be the reference geometry. All degrees of freedom except R1 are fully optimized.

distance R1 to values between 1.6 and 2.4 Å with a full optimization of the remaining degrees of freedom. Comparison of the dependence of the energy upon R1, for the HF method and for the three different density functionals employed in this study, confirms the dramatically different performance of the functionals (Figure 4). At the HF/6-31+G(d,p) level of theory, in addition to the minima **2a** and **6** discussed above, an energy maximum **7** at 1.801 Å is found. Optimization of that stationary point and analysis of the normal modes shows that maximum indeed to be the transition structure for ring opening. Nevertheless, this transition state is located only 0.74 kcal/mol above **2a**, and the barrier height reduces to only 0.20 kcal/mol upon correction for zero point vibration. The shapes of the hypersurfaces at the density functional levels of theory are entirely different. For the pure GGA functional BLYP/6-31+G(d,p), a comparatively deep minimum at 2.100 Å is found which is lower in energy by 3.5 kcal/mol, compared to the value at R1 = 1.600 Å. On the other hand, at the BHLYP/6-31+G(d,p) level of theory, there is a shallow minimum at R = 1.615 Å, and upon dissociation, the energy rises steadily. The most interesting feature of the BHLYP hypersurface is a tiny maximum, which was found to be a transition structure for ring opening **7**, with an associated energy barrier of 0.96 kcal/mol. An analysis of the imaginary normal mode reveals that this stationary point at R1 = 2.003 Å connects the R = 1.615 Å geometry of **2a** with the minimum **6** at R1 = 2.15 Å. However, this barrier vanishes when zero-point effects are included. The hypersurface at the BHLYP level exhibits a shape comparable to the one found at the HF level of theory, when R1 is smaller than 1.8 Å. At longer distances, the difference between the HF and BHLYP method increases significantly, with the long bonded minimum **6** at the HF level being far lower in energy. The third functional used here, the widely used B3LYP method, shows the same trend as that for the pure GGA functional BLYP. The shape is roughly parabolic, although less deep compared to the BLYP functional. At the equilibrium geometry, the energy is only 1.0 kcal/mol below the reference geometry at R1 = 1.6 Å (Table 4). Because of this surprising and somewhat disconcerting performance of the DFT methods, the bond breaking step in **2a** was further investigated, because it was not clear at this stage of the investigation which method describes the hypersurface correctly. Therefore, perturbation theory, the semiempirical method AM1, and higher correlated single-point calculations at the CCSD level were carried out (Figure 5).

At the MP2 level, a stationary point at R1 = 1.643 Å could be located (Table 1). The shape of the hypersurface is in good agreement with the one obtained at the BHLYP level of theory for R1 below 1.8 Å, but the energy rises more quickly upon further dissociation. To investigate if the MP2 or BHLYP hypersurface yields the better representation of the true shape of the surface, higher correlated single-point calculations at the CCSD/6-31+G(d,p)//BHLYP/6-31+G(d,p) level were carried out (CCSD//BHLYP in Figure 5). The BHLYP geometries were chosen as reference configurations, because, although energies and equilibrium geometries are quite different depending on the level of theory, the optimized geometries at a given oxygen−carbon distance R1 are very close as evidenced by their small RMS deviation (below 0.08 Å) when superimposed.[17] In general, the CCSD single-point calculations are in better agreement with the BHLYP results: The largest deviation between the two surfaces is found at R1 = 2.0 Å and amounts to only 0.28 kcal/mol, but at longer bond lengths the agreement improves again. At R1 = 2.4 Å the difference amounts to less than 0.1 kcal/mol. To exclude the possibility that this good agreement is caused by the choice of BHLYP as the reference geometry, additional CCSD single-point calculations were carried out. This time the geometries obtained at the BLYP and the MP2 level of theory for R1 = 2.4 Å were chosen for reference, because at

this distance the hypersurfaces differ most compared to that of BHLYP. The CCSD/6-31+G(d,p)//BLYP/6-31+G(d,p) energy is in good agreement with those obtained using the BHLYP geometry (CCSD//BLYP in Figure 5). This is surprising because the same geometry yields a BLYP energy, which is lower by 4 kcal/mol. The CCSD/6-31+G(d,p)//MP2/6-31+G(d,p) energy of 2.06 kcal/mol is slightly larger but still in better agreement with the BHLYP surface than the MP2 one. It is interesting to note that, at the semiempirical AM1 level, the shape of the hypersurface is in good agreement with the one obtained with the BHLYP functional (Figure 5). Over the whole range investigated, the difference between the two is below 1.0 kcal/mol and the previously mentioned CCSD single-point energies are found to be just between the ones obtained at the BHLYP and the AM1 level of theory. These results indicate, that the true energy surface looks rather like the one found at the BHLYP level, rather than those found with the other density functional methods.

We can conclude at this point that neither the BLYP nor the popular B3LYP functional are capable of describing the hypersurface of the system under investigation correctly. Results from the highly correlated CCSD single-point calculations suggest that MP2, despite its better performance at short bond distances, tends to rise in energy too quickly. Surprisingly good results are obtained at the semiempirical AM1 level. Results obtained at the AM1 and BHLYP levels of theory appear to be an upper and lower limit for the true reaction profile, as evidenced by the CCSD single-point calculations.

Discussion

A similar method dependence was found by Yang et al. in homonuclear open shell dimers[13] and is explained by the occurrence of fractional occupation numbers at the dissociation limit. It has been shown that all known GGA density functionals do not fulfill the scaling relation

$$q^2E_x[\rho] = E_x[q\rho], \quad q \in [0,1] \tag{1}$$

which is a necessary condition for any q electron density to be free of self-interaction. Moreover, for all current density functionals, the right-hand side of eq 1 is larger, which leads to an artificial stabilization of noninteger electron densities.[13] The occurrence of fractional occupation numbers during the dissociation is strongly linked to the difference between the ionization energy of the electron donor and the electron affinity of the electron acceptor of the dissociating fragments. Fractional electron numbers are most likely to occur when this splitting is small. For the radical cation of the hydrogen molecule, in which this difference equals zero, the spurious self-interaction is due to the resonance of the bonding and antibonding linear combination of the following configurations in the dissociation limit:

$$\phi_1[H^+ - H^\bullet], \quad \phi_2[H^\bullet - H^+] \tag{2}$$

At short intermolecular distances, the energy separation between bonding and antibonding linear combination of ϕ_1 and ϕ_2 is large, which explains why most of the density functionals perform well at equilibrium geometries. During separation of the atoms, the energetic difference of these states gradually becomes smaller, and self-interaction increases significantly. This is in line with the observation that GGAs usually underestimate reaction barriers, especially in open shell systems.[23]

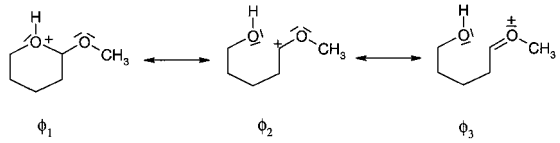

Figure 6. Schematic representation of the valence bond reference determinants under consideration.

There is no possibility to examine the properties of excited states by DFT if the states do not correspond to different irreducible representations. To circumvent this methodological limitation, we analyzed the electron density, evidenced by its Mulliken partial charges, in the following way: first, we assumed the total wave function can be expressed as a linear combination of only the reference valence bond (VB) configurations depicted in Figure 6. If the total charge distribution arises from these configurations only, the unnormalized wave function can be expressed by eq 3

$$\psi_T = \phi_1 + c_2\phi_2 + c_3\phi_3 \tag{3}$$

and the partial charges on the corresponding molecular fragments based on Ψ_T equal the following expressions:

$$q_{O5} = \frac{1}{1 + c_2^2 + c_3^2}, \quad q_{C1} = \frac{c_2^2}{1 + c_2^2 + c_3^2}, \quad q_{O1} = \frac{c_3^2}{1 + c_2^2 + c_3^2} \tag{4}$$

After basic transformation and normalization of the wave function, one obtains equations expressing the VB reference configuration's mixing coefficients c_1, c_2, and c_3 in terms of the partial charges of the molecular fragments, which were obtained from a Mulliken population analysis.[24]

$$c_1 = \sqrt{\frac{1}{Nq_{O5}}}, \quad c_2 = \sqrt{\frac{q_{C1}}{Nq_{O5}}}, \quad c_3 = \sqrt{\frac{q_{O1}}{Nq_{O5}}} \tag{5}$$

If such an analysis is performed for the potential energy surface shown in Figure 4, a qualitative picture of the relative importance of the states ϕ_1, ϕ_2, and ϕ_3 can be obtained. The molecular fragments were defined as follows: O5, C1, and O1 are the atoms on which the positive charge is formally located at the VB-reference configurations ϕ_1, ϕ_2, and ϕ_3. To account for inductive effects between oxygen and carbon atoms, the nearest heavy atom is included. Thus, q_{O5}, q_{C1}, and q_{O1} denote the sum of the Mulliken charges at centers O5/C5, C1/C2, and O1/CH_3, respectively. Table 5 summarizes the mixing coefficients for reference states ϕ_1, ϕ_2, and ϕ_3 obtained at the various levels of theory.

Inspecting the mixing coefficients at R1 = 1.6 Å reveals a significant difference in the relative importance of reference states, as shown by Figure 7. At the HF level, reference configuration ϕ_2 is the most important, whereas the pure GGA functional BLYP puts emphasis on configuration ϕ_1. This observation is quite remarkable, because chemical intuition is more in line with the results from the BLYP functional, rather than the results from Hartree−Fock theory. Again, as observed for the geometric differences, the hybrid functionals are between these two cases. At R1 = 1.6 Å, the B3LYP coefficients are close to the ones found with BLYP. On the other hand, BHLYP,

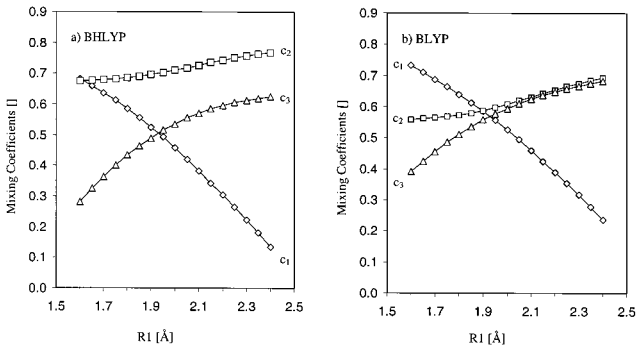

Figure 7. Plot of the mixing coefficients c_1, c_2, and c_3 versus the bond distance R1. Calculated (a) at the BHLYP and (b) at the BLYP level of theory with the 6-31+G(d,p) basis set. See text for details.

TABLE 5: Mixing Coefficients (c_1−c_3) for the VB Reference Configurations ϕ_1−ϕ_3 for the Ring Opening in 2A (See Text for Details)

R1 [Å]		1.60	1.70	1.80	1.90	2.00	2.10	2.20	2.30	2.40
HF	c_1	0.65	0.60	0.55	0.48	0.41	0.33	0.25	0.16	0.03
	c_2	0.75	0.75	0.76	0.77	0.78	0.79	0.80	0.82	0.82
	c_3	0.12	0.26	0.35	0.42	0.48	0.51	0.54	0.56	0.57
BHLYP	c_1	0.68	0.64	0.59	0.53	0.46	0.38	0.30	0.22	0.13
	c_2	0.68	0.68	0.69	0.70	0.71	0.73	0.74	0.76	0.77
	c_3	0.28	0.36	0.43	0.49	0.54	0.57	0.60	0.61	0.63
B3LYP	c_1	0.71	0.67	0.62	0.56	0.51	0.43	0.35	0.27	0.18
	c_2	0.61	0.62	0.62	0.64	0.65	0.67	0.70	0.71	0.73
	c_3	0.35	0.42	0.48	0.53	0.57	0.61	0.63	0.65	0.66
BLYP	c_1	0.73	0.69	0.64	0.59	0.53	0.46	0.39	0.32	0.26
	c_2	0.56	0.57	0.57	0.59	0.61	0.63	0.65	0.68	0.69
	c_3	0.39	0.46	0.51	0.56	0.60	0.63	0.65	0.67	0.68

where the reference configurations ϕ_1 and ϕ_2 are of equal importance, appears to be closer to the HF results. The trends found for coefficient c_1 are similar at all levels of theory. Upon dissociation of the bond, formal localization of the positive charge on the ring oxygen atom O5 should become increasingly unfavorable. A similar argument holds true for coefficient c_3. As long as there is donation of electron density from the lone pair of O5 into the cationic center at C1, ϕ_3 is high in energy. When dissociated, stabilization of the cationic center by O5 no longer takes place and a lone pair of the glycosidic oxygen atom O1 can stabilize the positive charge at C1. This is reflected in the steady increase of c_3, with increase in R1. The major difference between the wave functions at the different levels of theory is the behavior of c_2. At the HF and BHLYP levels of theory, ϕ_2 is the most important for the whole reaction coordinate. Especially for large values of R1, there is a strong bias toward ϕ_2. For the BLYP functional, and to a limited extent for B3LYP, the behavior is different. At short distances, in contrast to the BHLYP functional, configuration ϕ_1 is most important. At a bond distance R1 of 1.90 Å, the curves showing the variation of the various coefficients with R1 intersect (Figure 7), indicating a complete delocalization of the positive charge on all three centers. Moreover, upon further dissociation the difference between c_2 and c_3 remains very small. This means the wave function for the structure at large R1 (**6**) is a superposition of reference states ϕ_2 and ϕ_3. In terms of the explanation presented by Yang et al.,[13] the artificial stabilization found for the structures with long R1 (**6**) at the BLYP and B3LYP levels stems from self-interaction errors because of an artificial delocalization of the positive charge on carbon atom C1 and the glycosidic oxygen atom O1. This failure appears to be due to the subtle interplay between the electron donating oxygen and electron accepting carbon atoms.

Conclusions

An important conclusion of our calculations is that the occurrence of spurious self-interaction errors previously observed for a variety of open shell species is not limited to such radicals and is here found for closed shell carbohydrates. Such sugar derivatives are systems in which hyperconjugation can affect structure, such as the well-known anomeric effect. From a detailed analysis of the electron densities obtained at various levels of theory we conclude the following:

(1) The case reported in this study suggests that whenever hyperconjugation competes with weak chemical bonding the effect of self-interaction can become quite large and is *not* limited to open shell systems.

(2) Among the methods under investigation, the semilocal hybrid functional BHLYP yields results in good agreement with high level *ab initio* CCSD results. Surprisingly, the semiempirical method AM1 also shows good agreement with the BHLYP and CCSD results. Decreasing the amount of HF exchange in the hybrid functional steadily worsens the functional's performance. The BLYP and B3LYP are not capable of reproducing the hypersurface of the ring opening reaction of model system **2a**. To investigate enzymatic reactions, which are believed to involve protonated sugar moieties, by means of QM/MM calculations, the BHLYP functional is the best choice among the DFT methods.

(3) Analysis in terms of the electron density in terms of VB reference configurations shows this problem is caused by an artificial delocalization of the electron density between carbon atom C1 and oxygen O1 at long distances. As a result, this resonance between VB reference configurations ϕ_2 and ϕ_3 yields an physically unrealistic lowering of the energy which becomes more evident when the electronic interaction between the ring oxygen atom's lone pair and the cationic center decreases.

(4) Protonation of the glycosidic oxygen in **1** leads to a loosely bound ion dipole complex. The different levels of theory predict binding energies of the complex in the range of 11−15 kcal/mol.

Acknowledgment. This research has been supported by a Marie Curie Fellowship of the European Community program Improving Human Potential under Contract No. HPMF-CT-2000-00550. We thank EPSRC and the Ramsay Trust (RAB), for financial support.

Supporting Information Available: Full tables of total energies and partial charges. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) (a) Ma, B.; Schaeffer, H. F., III; Allinger, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 3411. (b) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *Chem. Phys Lett.* **1996**, *257*, 49.

(2) (a) Maw, S. A.; Bryce, R. A.; Hall, R. J.; Masters, A. J.; Hillier, I. H. *J. Phys. Chem. B* **1998**, *102*, 4089. (b) Tvarovska, I.; André, I.; Carver, J. P. *J. Phys. Chem. B* **1999**, *103*, 2560.

(3) (a) Glennon, T. M.; Zheng, Y.-J.; LeGrand, S. M.; Shutzberg, B. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15*, 1019−1040. (b) Homans, S. W. *Biochemistry* **1990**, *29*, 9110−9118. (c) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832−3846. (d) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078−2086. (e) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jørgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955−1970.

(4) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138.

(5) (a) Morpurgo, S.; Brahimi, M.; Bossa, M.; Morpurgo, G. O. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2707. (b) Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1999**, *121*, 9843.

(6) (a) Beyer, T.a.; Sadler, J. E.; Rearick, J. I.; Paulson; Hill, R. L.; *Adv. Enzymol.* **1981**, *52*, 23. (b) Kleene, R.; Berger, E. G. *Biochim. Biophys. Acta* **1993**, *1154*, 283. (c) Morera, S.; Imberty, A.; Schake-Sonnenborn,

U.; Ruger, W.; Freemont, P. *J. Mol. Biol.* **1999**, *292*, 717. (d) Kakinuma, H.; Yuasa, H.; Hashimoto, H. *Carbohydr. Res.* **1998**, *312*, 103.

(7) (a) Zechel, D. L.; Withers, S. G *Acc. Chem. Res.* **2000**, *33*, 11. (b) Heigthman, T. D.; Vasella, A. T. *Angew. Chem.* **1999**, *38*, 750. (c) Sinnott, M. L. *Chem. Rev.* **1990**, *90*, 1171. (d) Legler, G. *Carbohydr. Res.* **1993**, *250*, vii. (e) Withers, S. G.; Aersbold, R. *Protein Sci.* **1995**, *4*, 361. (e) Hays, W. S.; Van der Jagt, D. J.; Bose, B.; Serianni, A. S.; Glew, R. H. *J. Biol. Chem.* **1998**, *273*, 3494.

(8) (a) Andrews, S. R.; Charnock, S. J.; Lakey, J. H.; Davies, G. J.; Claeyssens, M.; Nerinckx, W.; Underwood, M.; Sinnott, M. L.; Warrens, R. A. J.; Gilbert, H. J. *J. Biol. Chem.* **2000**, *30*, 23027. (b) Post, C. B.; Brooks, B. R.; Karplus, M. *J. Mol. Biol.* **1986**, *190*, 455. (c) Post, C. B.; Karplus, M. *J. Am. Chem. Soc.* **1986**, *108*, 1317.

(9) (a) Dyer, E.; Glaudermans, C. P. J.; Koch, M. J.; Marckesvault, R. H. *J. Chem. Soc., Faraday Trans.* **1962**, 3361. (b) Armour, C.; Bunton, C. A.; Patai, S.; Selman, L. H.; Vernon, C. A. *J. Chem. Soc., Faraday Trans.* **1961**, 412.

(10) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(11) Ventura, O. N.; Kieninger, M. *Chem. Phys. Lett.* **1993**, *245*, 488.

(12) (a) Mohr, M.; Zipse, H.; Marx, D.; Parrinelllo, M. *J. Phys. Chem. A* **1997**, *101*, 8942. (b) Bally, T.; Sastry, G. N. *J. Phys. Chem. A* **1997**, *101*, 7923. (c) Hrounda, V.; Roeselova, M.; Bally, T. *J. Phys. Chem. A* **1997**, *101*, 3925. (d) Gauld, J. W.; Radom, L. *Chem. Phys. Lett.* **1997**, *275*, 28. (e) Parkinson, C. J.; Mayer, P. M.; Radom, L. *Theor. Chem. Acc.* **1999**, *102*, 92. (f) Bally, T.; Hrovat, D. A.; Thatcher Borden, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3363.

(13) Zhang, Y.; Yang, W. J. *J. Chem. Phys.* **1998**, *109*, 2604.

(14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision E.1; Gaussian, Inc.: Pittsburgh, PA, 1995.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(16) Booth, H.; Dixon, J. M.; Khedhair, K. A.; Readshaw, S. *Tetrahedron* **1990**, *46*, 1625.

(17) See the Supporting Information for details.

(18) Lii, J.-H.; Ma, B.; Allinger, N. L. *J. Comput. Chem.* **1999**, *20*, 1593.

(19) (a) Glukovtsev, M. N.; Pross, A.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 3498. (b) Chandra, A. K.; Goursot, A. *J. Phys. Chem.* **1996**, *100*, 11596.

(20) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 2098. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1992**, *199*, 557. (c) Slater, J. C. *Phys. Rev.* **1951**, *81*, 385. (d) Vosko, S. J.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(22) Lias, S. G.; Levin, R. D.; Kafafi, S. A. Ion Energetics Data. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2000; http://webbook.nist.gov.

(23) Gonzales, C. A.; Gill, P. M. W.; Pople, A. J. *Chem. Phys. Lett.* **1994**, *221*, 100.

(24) See the Supporting Information for the full set of charges used in the analysis.