Aspects of Glycosidic Bond Formation in Aqueous Solution: Chemical Bonding and the Role of Water

John M. Stubbs*[a] and Dominik Marx[b]

Abstract: A model of the specific acid-catalyzed glycosidic bond formation in liquid water at ambient conditions is studied based on constrained Car–Parrinello ab initio molecular dynamics. Specifically the reaction of α-D-glucopyranose and methanol is found to proceed by a D_NA_N mechanism. The D_N step consists of a concerted protonation of the O^1 hydroxyl leaving group; this process results in the breaking of the C^1–O^1 bond, and oxocarbenium ion formation involving C^1=O^5. The second step, A_N, is the formation of the C^1–O^m glycosidic bond, deprotonation of the methanol hydroxyl group O^mH^m, and re-formation of the C^1–O^5 single bond. A focus of this study is the analysis of the electronic structure during this condensed phase reaction relying on both Boys/Wannier localized orbitals and the electron localization function ELF. This analysis allows the clear elucidation of the chemical bonding features of the intermediate bracketed by the D_N and A_N steps, which is a non-solvent equilibrated oxocarbenium cation. Most interestingly, it is found that the oxygen in the pyranose ring becomes "desolvated" upon double bond/oxocarbenium formation, whereas it is engaged in the hydrogen-bonded water network before and after this period. This demonstrates that hydrogen bonding and thus the aqueous solvent play an active role in this reaction implying that microsolvation studies in the gas phase, both theoretical and experimental, might lead to qualitatively different reaction mechanisms compared to solution.

Keywords: glycosides · reaction mechanisms · reactive intermediates · solvent effects · transition states

Introduction

There are several archetypal "bonding patterns" found in nature connecting simple molecular building blocks via covalent bonds to biopolymers possessing a vast spectrum of functionalities. Among them are the peptide bond, the disulfide bond, and the glycosidic bond thus highlighting the importance of these motifs within (bio)chemistry. In particular, monosaccharides such as α-D-glucopyranose are the building blocks of carbohydrate chemistry, linked together by glycosidic bonds. Simple sugars make up a wide variety of substances found in nature, from DNA to cellulose to biological energy storage mechanisms. Consequently, the glycosidic bond has been studied experimentally for a wide variety of compounds,[1–13] predominantly through hydrolysis, under a myriad of conditions, and is important in the consideration of enzymatic glycosyl-transfer reactions. One of the reasons it has been studied so extensively is because of the rich and often subtle variety of the reactions that take place. Widely varying amounts of inversion and retention of the anomeric center are readily observed depending upon anomeric starting configuration, leaving group, additional ring substituents, and solvent. Adding to the complexity are competing hydrolysis reactions when multiple products are possible.[1] It is accepted that the reaction proceeds through an oxocarbenium cation intermediate; the mechanism involved in hydrolysis of glucopyranosides has been concluded to be dissociative (D_N+A_N) or partially-dissociative ($A_N^*D_N$)[1,2,13] which lies between S_N2 and S_N1 behavior, although a concerted mechanism (A_ND_N)[5,6] has been observed for the case of aqueous hydrolysis of fluoro-α-D-glucopyranoside (see ref. [14,15] for notation details). Upon reviewing previous work it has been concluded[8] that the hydrolysis of methyl

[a] Dr. J. M. Stubbs
Lehrstuhl für Theoretische Chemie
Ruhr-Universität Bochum, 44780 Bochum (Germany)
on leave from:
Department of Chemistry, University of Minnesota
Minneapolis, MN 55455-0431 (USA)
Current address: Department of Chemistry, Grinnell College
Grinnell, IA 50112-1690 (USA)
Fax: (+1) 641-269-4285
E-mail: stubbsj@grinnell.edu

[b] Prof. Dr. D. Marx
Lehrstuhl für Theoretische Chemie
Ruhr-Universität Bochum, 44780 Bochum (Germany)

 DOI: 10.1002/chem.200400773

α- and methyl β-glucopyranosides is specific acid-catalyzed[2,16] and that the mechanism has been explained by the pre-equilibrium formation of a protonated leaving group intermediate which provides insufficient time for solvent equilibration. The transition state has an estimated life time[2,10] of ~1–3 ps.

Although the hydrolysis of glucopyranosides has been thoroughly investigated, the forward reaction to form polysaccharides has not received much attention for several reasons, including energetic unfavorability and lack of appearance in carbohydrate digestion (where precisely the opposite occurs). Due to the subtlety in the effects of leaving group[5,6] and temperature[12] on the hydrolysis reaction the mechanism of the forward reaction is not necessarily the same as its hydrolysis, though reaction energetics should be comparable. In particular, the questions of the reaction being late or early, the transition state being solvent equilibrated or not, and the degree of C^1–O^1 bond cleavage at the transition state can be answered and compared with the hydrolysis reaction.

Previous theoretical treatments have investigated aspects of such hydrolysis reactions by using quantum chemical gas phase or continuum solvation methods[17–21] where the conformation of exocyclic substituents such as the C^6–O^6 hydroxymethyl orientation was found to be particularly important for the hydrolysis of methyl β-D-glucopyranoside. These studies focused mainly on aspects of reactivity and conformational stability of mono- or disaccharides in the ground state and transition states, including relative anomeric stability and substituent effects. A notable exception to these studies includes the investigation of uracil-DNA glycosylase via a QM/MM approach.[22] Others have focused on the solvation and hydrogen-bonding behavior without studying chemical reactions using molecular dynamics methods.[23–25] Combining these findings it can be anticipated that the presence of water will be critical for the mechanism so that condensed phase results are expected to differ substantially from gas-phase.

Stimulated by the lack of theoretical studies at typical "wet chemistry conditions" in conjunction with the general importance of the glycosidic bond we launched a Car–Parrinello ab initio molecular dynamics[26–28] study; preliminary results have been published very recently in a short note.[29] The goal of this investigation was to determine the mechanistic details including the transition state structure, extent of oxocarbenium ion formation and free energy of reaction at conditions relevant to experiment, that is, in acidic liquid water at room temperature and atmospheric pressure. In order to obtain a computationally tractable model for dynamical ab initio simulations one of the monosaccharide partners was replaced by a methanol molecule. Thus, we investigate how methyl β-D-glucopyranose is formed via the specific acid-catalyzed reaction of methanol with α-D-glucopyranose. The method of Car–Parrinello molecular dynamics[26–28] was chosen because of its ability to study bond formation and breaking, its incorporation of dynamical constraints, the ability to explicitly include a solvent at the same footing as the reactants, and the ability to do all this at a biologically relevant temperature and pressure.

Computational Methods and Modeling Approach

Car–Parrinello molecular dynamics[26,27] as implemented in the CPMD code[27,28] was used for both the dynamic ab initio simulations and the static optimizations in the liquid and gas phase, respectively. For these electronic structure calculations Troullier–Martins norm-conserving pseudopotentials[30] were employed together with the BLYP functional[31,32] by using a plane wave cutoff of 70 Ry at the Γ point of the simulation cell; periodic boundary conditions were applied for the condensed phase simulations, whereas gas phase calculations where carried out within isolating "cluster" boundary conditions.

Initially, a study of the proton affinity of the respective oxygens of α-D-glucopyranose was carried out with static ab initio calculations in the gas phase to find out which of the six oxygen atoms is a preferred site for protonation, see Figure 1. Starting with the crystal structure[33] of α-D-glucopyranose gas phase optimizations were performed, followed by the generation of six different protonated structures by placing a proton in one of the two "tetrahedral lone pair sites" of each of the six oxygen atoms and optimizing the resulting protonated molecules.

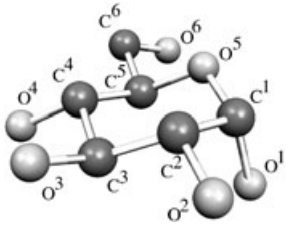

Figure 1. Numbering scheme of α-D-glucopyranose molecule's heavy atoms.

In order to prepare the system in solution in a chemically relevant state an initial pre-equilibration was performed with classical molecular dynamics, which satisfactorily describes the non-reactive interactions involved in this process. In particular, a (neutral) starting configuration was prepared by very slowly pulling together the O^m atom of methanol and the C^1 site on the sugar ring (see Figure 1 for labeling) in a large simulation cell; again the crystal structure[33] of α-D-glucopyranose was used as an initial structural guess. For this purpose the OPLS-AA all-atom force field for carbohydrates[34,35] was used in conjunction with the TIP3P[36] water model for the classical simulations of α-D-glucopyranose, methanol, and water at a density of 0.997 g cm^{-3}. These classical simulations were carried out with the TINKER simulation package.[37] The electrostatic interactions were computed with a distance-based cutoff at 9 Å and periodic boundary conditions were applied. A time step of 1 fs was used with the velocity–Verlet integration scheme, and a temperature of 300 K was maintained by a Berendsen thermostat with coupling time of 1 ps.

This force field based setup was used to pull the reactants slowly together by applying a harmonic distance constraint (or "restraint") on the C^1–O^m distance, starting with an interatomic distance of 10 Å and decrementing by 0.5 Å from 10 to 3.5 Å followed by 0.25 Å from 3.5 to 2.5 Å by using a force constant of 100 kcal mol^{-1}. In addition, it was necessary to apply such a constraint on the O^1–O^m distance as to keep O^m close to the C^1–O^1 vector in order to avoid hydrogen bonding between O^1 and methanol at short distances; this would be an unfavorable starting configuration for a substitution reaction on the C^1 carbon. Three α-D-glucopyranose rotamers were used, all sharing the $^4C_1$ conformation of the pyranose ring but differing in both the conformation of the hydroxymethyl group (rotation around the C^5–C^6 bond) and direction of the hydrogen bonds of the hydroxyl groups. Finally, one configuration was chosen at a constraint distance of about 2.7 Å such that the group was located opposite the O^1 atom with respect to the sugar ring based on "chemical intuition" of a displacement reaction. Subsequently, the number of water molecules was decreased (by using NVT and NpT simulations at 300 K and 1 atm) from 483 to 89, 73, 57, and 44 solvating molecules. No significant difference was found between the radial distribution function of water oxygen atoms around both C^1 and O^1 as obtained by 89, 73, 57 molecules, whereas qualitative deviations were found with 44 molecules. In order to prepare a starting configuration for the follow-up ab initio simulations a configuration was sampled from a NpT run (within a cubic supercell) with 57 water molecules where O^m was close to the C^1–O^1 vector subject to having system parameters such as volume and total potential energy close to their respective average values. This particular system had a box length of 12.255 Å which yielded a density of 1.118 g cm^{-3}.

This configuration of the neutral system was then equilibrated for about 2.5 ps by using Car–Parrinello ab initio molecular dynamics. After this simulation phase, a proton was added at one of the two tetrahedral oxygen lone-pairs of a water molecule that was hydrogen bonding with the O^1 oxygen; the latter was found to be the energetically most favorable protonation site in the gas phase. This system composed of 57 water molecules to solvate the protonated reactive complex in a cubic box of 12.255 Å was propagated by Car–Parrinello ab initio molecular dynamics; the overall positive net charge was compensated as usual by applying a uniform negative background charge. A temperature of 300 K was established via Nosé–Hoover thermostatting, thus properly sampling the canonical ensemble. Hydrogen was substituted by deuterium allowing for a time step of 0.12 fs in conjunction with a fictitious electron mass of 1000 au and separate thermostatting of the orbitals.

Since the reaction barrier is certainly too high (experimental estimates[38] for the hydrolysis of methyl β-D-glucopyranoside, which is at least as large of a barrier as for the forward reaction studied here, are in the range of about 31–34 kcal mol^{-1}) to be overcome by thermal fluctuations on the picosecond time scale of ab initio molecular dynamics runs, constraint techniques[39,40] have to be used. Here, we continue to use the distance between the O^m atom of methanol and the C^1 site on the sugar ring as a possible reaction coordinate to drive the reactive complex to the product state. The free energy along this particular variable is obtained via thermodynamic integration from the integral of the average Lagrange multiplier in the constrained ensemble which defines the mean force for the simple interatomic distance constraint used.[40] Starting from 2.69 Å the constraint distance was decreased in a stepwise fashion to 1.46 Å by using increments from 0.15 to 0.075 Å. Overall, a total ab initio simulation time of 13.3 ps (after ab initio equilibration) was necessary in order to simulate the entire process of glycosidic bond formation including a meaningful sampling of the average Lagrange multiplier at each increment.

Results and Discussion

Gas phase—Structure and proton affinity: The results from the structural optimizations of α-D-glucopyranose in isolation are collected in Table 1 in terms of dihedral angles; other structural parameters such as bond lengths and bond angles are within the expected ranges. The structure of the neutral sugar molecule agrees well with the experimentally determined structure in the crystal[33] and previous density functional and MP2 calculations of the isolated molecule.[19,41,42] As seen in Table 2, the O^1 site is the most energetically favorable site for protonation in the gas phase. For comparison, the proton affinity of water obtained with the very same approach was determined to be 169 kcal mol^{-1}; a recent G2 calculation[43] yielded 163 kcal mol^{-1} whereas the experimental value[44] is 165 kcal mol^{-1}. This suggests that O^1 might also be the thermodynamically preferred site for protonation in solution. An interesting and unexpected result of these gas phase calculations was the "product" of the O^4 protonation. Having an energy value of only 0.18 kcal mol^{-1} above that of the protonated O^1 site, which is essentially a degenerate situation in view of the accuracy of the electronic structure method, this optimization resulted in a complete rearrangement of the six-membered pyranose to a five-membered furanose ring with the H_2O leaving group, involving O^4, stabilizing the carbocation at C^3 formed as shown in Figure 2. This rearrangement is probably the cause for our finding the proton affinity of O^4 to be higher than

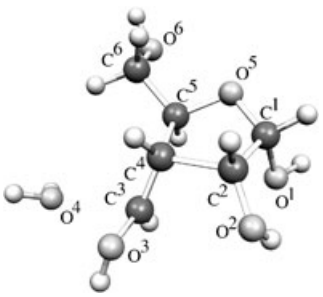

Figure 2. Optimized gas-phase structure for α-D-glucopyranose protonated at the O^4 site. Note the qualitative rearrangement from the pyranose ring from Figure 1 into a five-membered furanose ring.

O^6 in contrast to previous Hartree–Fock calculations,[45] where no such complete restructuring of the sugar ring was reported. Additionally, although the proton affinities are (essentially) the same for O^1 and O^4, the required rearrangement in the latter case would not lead to the desired glycosidic bonding pattern and would in any case almost certainly not occur in aqueous solution where interactions with neighboring water molecules would not allow such a rearrangement. However, this result might have interesting implications for gas phase studies. Thus, protonation of O^1 is most promising and will be considered in detail in aqueous solution.

Table 1. Comparison of α-D-glucopyranose torsional angles [°]; note that the experimental structure[33] is obtained from crystallographic data.

Atoms	Experimental[33]	This work	BLYP[41]	MP2[19]	B3LYP[42]
exocyclic angles					
C^2-C^1-O^1-H	−163.1	−170.2	−170		
C^3-C^2-O^2-H	24.4	75.3	75		
C^4-C^3-O^3-H	73.8	56.7	60		
C^5-C^4-O^4-H	−103.6	−81.9	−90		
C^5-C^6-O^6-H	−143.8	−161.1	180		−58.0
O^5-C^5-C^6-O^6	70.2	75	74.9		60.8
C^4-C^5-C^6-O^6	−170.3	−165.4	−165		
O^5-C^1-O^1-H	74.9	66.4			67.6
C^1-O^5-C^5-C^6	−176.5	−176.6			−178.7
endocyclic angles					
C^1-C^2-C^3-C^4	−51.3	−49.3		−54.0	
C^2-C^3-C^4-C^5	53.3	52.9		56.5	55.9
C^3-C^4-C^5-O^5	−57.5	−56.3		−57.9	−56.6
C^4-C^5-O^5-C^1	62.2	60.3		60.1	58.7
C^5-O^5-C^1-C^2	−60.9	−58.8		−57.9	
O^5-C^1-C^2-C^3	54.1	51.7		53.8	

Table 2. Gas phase α-D-glucopyranose proton affinities in kcal mol^{-1}; see text for a discussion of the comparison with the Hartree–Fock data.[45]

Site	This work	HF/6-31G[45]
O^1	214	
O^2	204	
O^3	205	
O^4	214	190
O^5	208	
O^6	207	196

Condensed phase—Mechanism: Upon starting from a situation where α-D-glucopyranose is close to the methanol molecule, the distance between the O^m atom of methanol and the C^1 site on the sugar being 2.7 Å and a protonated water molecule donating a hydrogen bond to O^1, the approach of the methanol molecule was enforced by decreasing the C^1–O^m distance in a stepwise fashion; see Section on Computational Methods above for a detailed description of the protocol. The reaction mechanism can be inferred from the evolution of interatomic separations plotted in Figure 3a, hydrogen bond coordinates in b) and sugar ring torsional angles in c). The hydrogen-bond coordinates depicted in Figure 3b are asymmetric stretch coordinates defined such that a value of zero implies that a given proton resides midway between the respective donor and acceptor atoms. Note that only for convenience, the entire reaction sequence is plotted on one time axis and events occurring during the simulation will be referred to according to at what "time" they occurred; however, it must be emphasized that due to the effect of the constrained dynamics this should not be interpreted as real time.

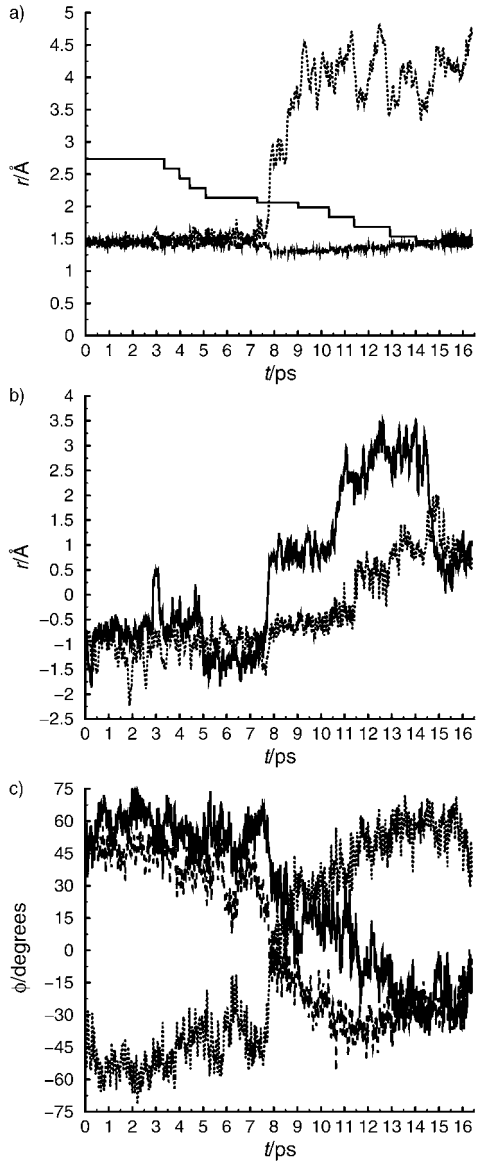

Figure 3. Evolution of distances and torsional angles as the C^1–O^m constraint is decreased from 2.689 Å at $t=0$ ps to 1.462 Å at $t=15$ ps where unconstrained dynamics sets in. a) Solid, dotted and dashed lines represent C^1–O^m, C^1–O^1 and C^1–O^5 distances, respectively. b) Solid and dotted lines represent the hydrogen bond coordinates $r_{O^wH^+}-r_{O^1H^+}$ and $r_{O^mH^m}-r_{O^wH^m}$, respectively. c) Solid, dotted and dashed lines represent C^4-C^5-O^5-C^1, C^5-O^5-C^1-C^2 and O^5-C^1-C^2-C^3 angles, respectively. Labeling scheme according to Figure 1.

The formation of methyl β-D-glucopyranoside is found to occur in two sequential steps along the chosen reaction coordinate, the C^1–O^m distance (Figure 3a, ——), which is decreased throughout the reaction. The first step, I, consists of concerted protonation of O^1 in the leaving hydroxyl group (Figure 3b, ——), breaking of the C^1–O^1 bond (a, dotted line), and shortening of the endocyclic C^1–O^5 bond (a, graph at $r = 1.5$ Å) close to 7.76 ps at which point the C^5-O^5-C^1-C^2 torsion moves from negative to positive values (c, dotted line). These events are initiated by irreversible protonation of O^1. This also brings about a change in the orientation of the attacking methanol as evidenced by the O^5-C^1-O^m angle moving from 80 to 100°, see also the schematic representation in Figure 4a,b. The second step, II, involves the formation of the new C^1–O^m bond (Figure 3a, ——), deprotonation of O^m (Figure 3b, dotted line crossing zero), an increase in the C^1–O^5 bond length (Figure 3a), dashed line), and cessation of changes to the ring conformation (Figure 3c, see also Figure 4c,d. Prior to this successful step two temporary protonation events of the leaving hydroxyl group, that is, of O^1, were observed at about 3.0 and 4.7 ps (Figure 3b, ——), which were, however, unsuccessful in breaking the C^1–O^1 covalent bond. These events imply that it was necessary to have the methanol O^m sufficiently close to C^1 before irreversible protonation and thus dehydration could occur. This scenario does not support mechanisms where the intermediate cation is assumed to have sufficient time to become solvent equilibrated before the C^1–O^m bond is formed, similar to glycosyltransferase mechanisms[46] with close proximity of the nucleophile before ionization.

Interestingly, each reversible protonation of O^1 was followed by changes in bond lengths between C^1–O^1 which increases and C^1–O^5 which decreases, most noticeable in Figure 3 at 3.0 ps. This observation is characteristic for both specific acid catalysis[2,16] and the near synchronicity of step I. The coupling of C^1–O^5 bond shortening and C^1–O^1 bond breaking, which occur nearly simultaneously at 7.76 ps, is at odds with kinetic isotope effect data for the hydrolysis of methyl α- and methyl β-D-glucopyranoside[2] which indicates a delay between the two events. However, the same hydrolysis reaction for methyl xylopyranosides[47] is known to be simultaneous. The three endocyclic ring torsional angles, which can be used to monitor the conformation during the simulation, are depicted in Figure 3c. The change from the initial $^4C_1$ to the final predominantly $^1S_3$ twist boat

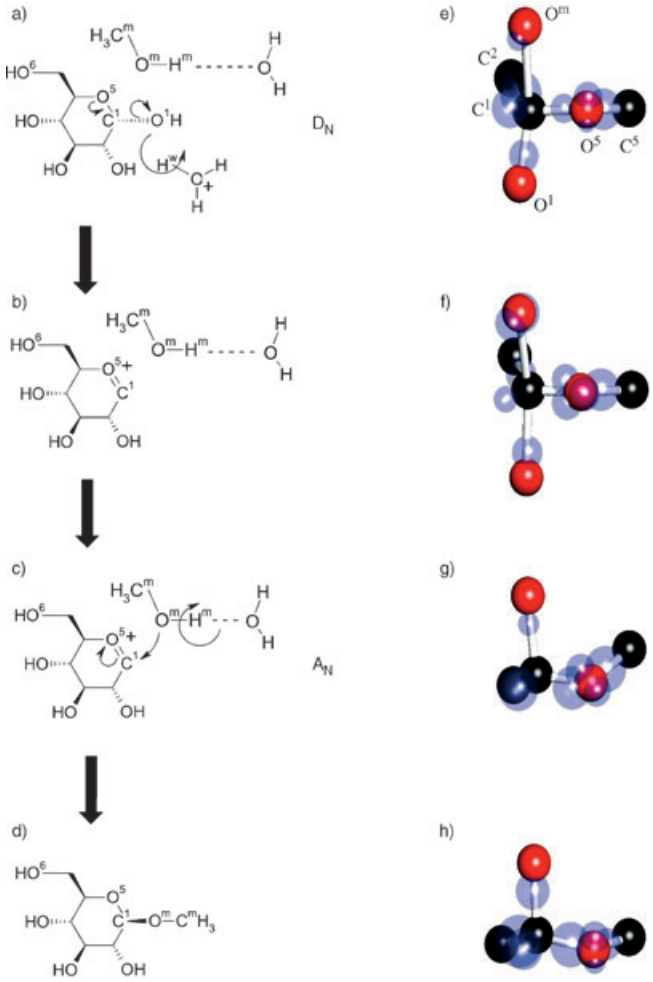

Figure 4. Scheme of the reaction mechanism a)–d) and ball-and-stick representation of the most important atoms (O: red, C: black) in representative trajectory snapshots taken at 7.27 ps (e), 7.74 ps (f), 12.5 ps (g), and 13.5 ps (h). The Wannier centers are represented by transparent spheres (blue) with radii being proportional to the spread of the localized orbitals.

conformation, which occurs continuously starting at 7.76 ps, is clearly seen. Panels a) and c) in Figure 5 depict these conformations as sampled from the trajectory. Once the C^1–O^m distance has reached 2.06 Å at 7.27 ps, it takes about 500 fs until the C^1–O^5 bond length contracts abruptly by as much as ≈0.15 Å after which it relaxes only slowly back to the

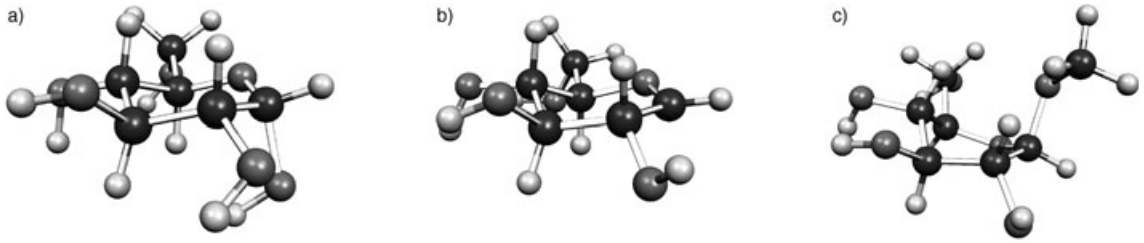

Figure 5. Conformations of the glucopyranose molecule throughout the simulation. a) $^4C_1$ conformation at 3.07 ps; b) $^1S_5$ conformation at 7.75 ps; c) $^1S_3$ conformation at 13.90 ps.

original value, see the dashed line in Figure 3a. The reason for this phenomenon, which is the transient formation of an oxocarbenium ion, can be understood at the molecular level upon analyzing electronic structure changes during the course of the simulation, see next section.

Shortly after the oxocarbenium formation it is observed by analyzing coordination numbers that O^5 expels its hydrogen-bonded solvating water molecule at about 8 ps. Furthermore, it is found that this site remains unsolvated during the lifetime of the oxocarbenium cation before it partially resolvates itself by establishing a hydrogen bond involving *another* water molecule at $\approx$11.4 ps. This temporary "drying" of O^5 is visualized in Figure 6, which depicts the distance from O^5 to the closest hydrogen atoms over the course of the simulation in comparison to the C^1–O^5 bond length, which monitors the oxocarbenium state.

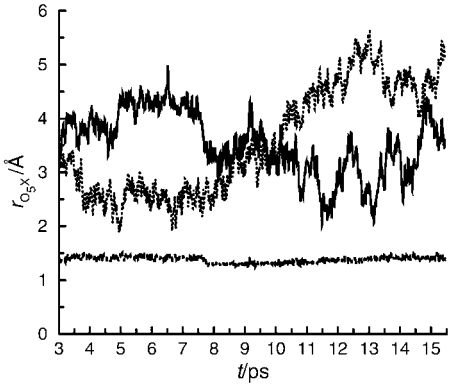

Figure 6. Distances between O^5 and the two closest solvent hydrogens (•••••, ——) in comparison to the C^1–O^5 bond length (dashed). The loss of hydrogen bonding involving O^5, between about 8 and 11 ps, correlates with the existence of an oxocarbenium cation which is characterized by a decrease in the C^1–O^5 bond length.

Between approximately 11.4 and 12.9 ps, at a constant C^1–O^m constraint distance of 1.69 Å, several deprotonation attempts of O^m occur, see dotted line in Figure 3b, each of which manifests itself by a slight but significant increase of the C^1–O^5 bond length. This period is characterized by electronic structure fluctuations of the oxocarbenium cation, see next section. The end of this period, at a constant C^1–O^m distance of 1.69 Å shows several reversible deprotonations of the O^m oxygen, each of which show a corresponding increase in the C^1–O^5 bond length. Finally, when the C^1–O^m distance is reduced to 1.54 Å, which lies at the edge of the fluctuating range of bond lengths in the subsequent unconstrained simulation (after 15.0 ps), the O^m oxygen is definitely deprotonated, that is, the dotted line in Figure 3b stays positive after 12.9 ps, at which time the C^1–O^5 bond length returns to the typical single-bond value, see Figure 4d, and a hydrogen bond is again fully accepted by O^5. Based on O^1–H^+, C^1–O^1, and C^1–O^5 distance criteria steps I and II of the reaction occur at 7.7 and 12.9 ps, respectively, with the period of reversible deprotonation of O^m occurring between 11.4 and 12.9 ps.

Condensed phase—Electronic structure: The chemistry responsible for these structural changes becomes obvious upon analyzing the electronic structure via the localization of the canonical Bloch orbitals in terms of Wannier functions.[48] Wannier functions (or orbitals) are the analogues of Boys' localized molecular orbitals for periodic systems such as a molecular liquid subject to periodic boundary conditions as simulated here. Thus, they allow a dynamical analysis of the electronic structure imposing a "Lewis-type viewpoint" based on concepts such as single bonds, double bonds and lone pairs. To this end the Wannier centers, being the average positions of corresponding orbitals, can be used as simplified indicators[49] as depicted in the right sequence of panels in Figure 4 where only the most important atoms and corresponding orbitals/centers are cut out of the trajectory for the sake of clarity. In Figure 4e, corresponding to the "normal" C^1–O^5 bond length, one can identify two Wannier centers connected exclusively to O^5 thus representing its two lone pairs. In addition there is one center each on the connecting axes to the neighboring carbon atoms (C^1 and C^5), which represent the two C^1–O^5 and C^5–O^5 single bonds. This bonding pattern changes *qualitatively* once the C^1–O^5 bond contracts around 7.7 ps: there is only *one* lone pair left at O^5 whereas *two* Wannier centers become aligned along the C^1–O^5 axis, see Figure 4f. At the same time the Wannier center between O^5 and C^5 stays invariant implying the existence of the C^5–O^5 single bond. In terms of a simple Lewis picture this change clearly signals the formation of a C^1=O^5 double bond, that is, of an oxocarbenium cation. This is the site for the incoming methanol molecule to form an ion–dipole complex which remains stable up to about 11.4 ps. In the interval between approximately 11.4 and 12.9 ps several deprotonation attempts of O^m occur and the Wannier centers of O^5, see Figure 4g, are found to oscillate between one or two lone pairs and a double or single C^1–O^5 bond, respectively, corresponding to sp^2- and sp^3-hybridization patterns of C^1. After reducing C^1–O^m to 1.54 Å the C^1–O^5 bond assumes again a clear single bond character as seen from the Wannier centers in Figure 4h which closely resemble the situation in Figure 4e.

This bonding analysis, which relies on a specific orbital localization procedure, was supplemented and confirmed by a more elaborate topological analysis[50–53] of the electron localization function (ELF) introduced by Becke and Edgecombe;[54] note that the present analysis focuses exclusively on the chemically relevant valence electrons. This function $\eta(\boldsymbol{r})$ is large in regions where two electrons with antiparallel spin are paired in space thus forming covalent bonds or lone pairs;[50,51] ELF is normalized between zero and unity and its value for the uniform electron gas is $^1/_2$. Positions in space where ELF attains maximum values (denoted by η^{max}) are called attractors which can be used to locate covalent bonds and/or lone pairs; conventional single bonds are characterized by one attractor on the axis connecting two nuclei

whereas double bonds feature two attractors above and below the respective symmetry plane; note that in the case of covalent bonds involving hydrogen, H–*X*, the corresponding bond attractors are shifted toward the position of the proton. In order to characterize the regions around attractors in more detail one can consider all the points in space with $\eta(\boldsymbol{r}) \geq f$, which define corresponding f localization domains; f is a positive constant smaller than η^{max}. These regions in space can be visualized by showing the isosurface $\eta(\boldsymbol{r}) = f$ such as done in Figure 7 for two important configurations sampled from the trajectory. For a given value of f, several closed isosurfaces surrounding different attractors can exist. As the value of f is lowered different spatially separated f localization domains are confluent and closed isosurfaces appear which contain more than one attractor.

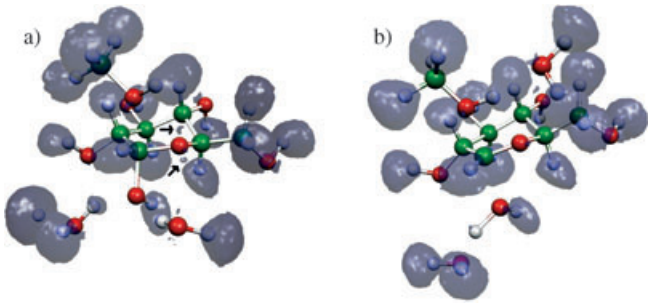

Figure 7. Electron localization function for a reduced system of reactants and two water molecules participating in the proton transfers, a) at 7.27 and b) at 7.75 ps.

In the present case, clear evidence for oxocarbenium formation is found in terms of the emergence of a double bond attractor between C^1 and O^5. At a time of 7.27 ps there are two pronounced lone pair ELF attractors (with $\eta^{max} = 0.94$ in both cases and confluence of the localization domains at $\eta^* = 0.90$, see arrows in Figure 7a) and two single bond attractors between O^5 and both C^1 and C^5 (in both cases $\eta^{max} = 0.85$). This topology is changed qualitatively at a time of 8.48 ps where O^5 clearly has only one lone pair left ($\eta^{max} = 0.94$), whereas a double bond attractor is now located between O^5 and C^1 ($\eta^{max} = 0.84$) while preserving the single bond attractor between O^5 and C^5 ($\eta^{max} = 0.86$), similar to Figure 7b. At 10.83 ps, O^5 is found to have one lone pair ($\eta^{max} = 0.95$), a sharp (in terms of its localization domain) single bond attractor between O^5 and C^5, whereas the attractor between O^5 and C^1 is of single-bond type but its localization domain is very large/diffuse. As already discussed above, this transition period up to about 11.4 ps is characterized by an unsolvated O^5 site. Finally, at 13.19 ps the standard bonding situation is fully re-established, that is, two lone pairs at O^5 ($\eta^{max} = 0.93$) and two ("sharp") single bond attractors between O^5 and C^5 as well as between O^5 and C^1 (in both cases $\eta^{max} = 0.85$).

In order to scrutinize and possibly complement the assignment of the two steps as inferred from simple structural indicators the Wannier analysis was performed for many system configurations both shortly before and after steps I and II. This dynamical electronic structure analysis allows to determine at which step the C^1–O^5 bond changed from single to double bond character and back. Based on the locations of Wannier centers step I clearly occurs at 7.7 ps where an O^5 oxygen "lone pair" center becomes oriented along the C^1–O^5 axis, whereas the center that was previously located *on* the axis connecting C^1 to O^5 moves off-axis which yields a C^1=O^5 double-bond pattern. Throughout the time between 11.4 and 12.9 ps where reversible deprotonation of O^m occurs the electronic structure "close to" O^5 oscillates between having one and two Wannier centers oriented along the C^1–O^5 axis, and thus two and one lone pair centers at O^5 respectively, finally remaining at one site after 12.9 ps. Thus, both the electronic and real-space structure lead to the same window of existence of the oxocarbenium cation and thus yield the same assignments of steps I and II.

Condensed phase—Energetics: The free energy of the reaction can be calculated[39,40] from the force acting on the constraint along the reaction coordinate ξ, that is, the C^1–O^m distance, by integration of the curve shown in Figure 8a. From the free energy profile the difference between the reactant and product energies is about 27 kcal mol^{-1} with a barrier height of about 35 kcal mol^{-1} to formation. Additional simulations at larger values of ξ were later added to account for a consistent initial and final force of about zero, causing the slight difference between this value of the free energy and our initially reported one.[29] As the constraint distance shortened, a negative constraint force is observed, see Figure 8a. However, the force finally approaches zero as the adduct becomes deprotonated when H^m is "taken" by a solvating water molecule as depicted in Figure 4c. In turn, the resulting hydronium ion H_3O^+ detaches from the adduct complex, becomes solvated in the liquid, and features standard Grotthuss structural diffusion.[55] The final value of the C^1–O^m distance constraint was 1.462 Å, which is close to the average value of about 1.47 ± 0.05 Å as determined from an unconstrained $\approx$1 ps simulation in continuation of the last constrained run, see Figure 3a.

Based on the shape of the free energy profile, see Figure 8b, the transition state can be identified from the maximum of the curve: it corresponds to a constraint value of about 1.95 Å or equivalently to $\approx$9 ps. Thus, the transition state is close to step I which is the situation depicted in Figure 4b and f. This is broadly consistent with both the assessment of structural rearrangements as the reaction progresses and the dynamics of the electronic structure extracted from Wannier orbital and ELF analysis. Furthermore, the conformation of the pyranose ring close to the transition state can be analyzed in detail, see for example, Figure 3c. It is found that its conformation is a flattened $^4C_1$ chair with some $^1S_5$ character as depicted in Figure 5b. This is similar to the proposed transition states for corresponding hydrolysis of methyl β-D-glucopyranoside.[1,2,8] At the transition state the C^1–O^1 distance was $\approx$1.9 Å indicating the degree of bond

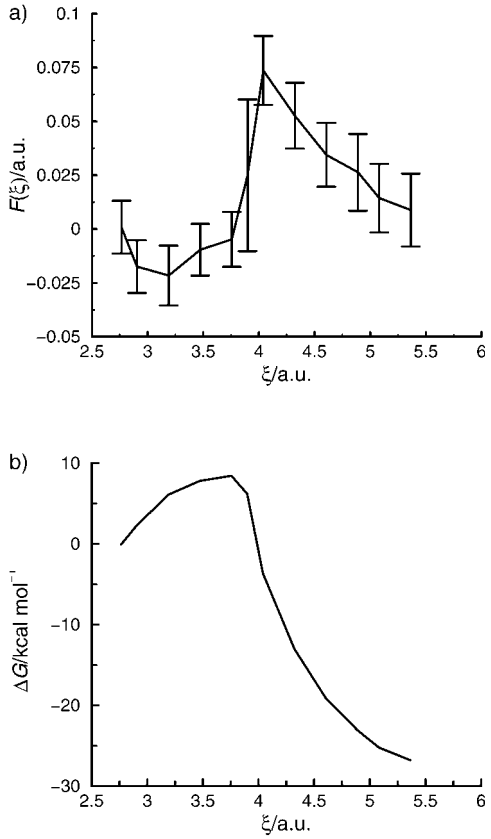

Figure 8. Constraint force (a) and corresponding free energy profile (b) as a function of the reaction coordinate ξ which is the constrained C^1–O^m distance. Standard deviations obtained from the fluctuations of the Lagrange multiplier during the sampling windows at each value of the constraint, see black line in Figure 3a, are shown as error bars.

cleavage is almost complete. As the starting conformation of the pyranose ring was also $^4C_1$ the transition state must be regarded as an early one, which is consistent with a late transition state for the back reaction.[2,5,6] The stability range of the oxocarbenium cation can also be roughly inferred from the shape of the free energy profile, see Figure 8, from values of ξ close to the maximum (where the transition state exists). Taking the maximum to be bounded by $3.9 < \xi < 3.2$ au corresponds to simulation times between 6 and 12 ps. This assignment of the oxocarbenium lifetime complements the previously discussed assignments based on atomic and electronic structure.

At this stage it is revealing to compare this scenario to what was observed in gas phase calculations[18] where the transition state of protonated methyl β-D-glucopyranoside was found to feature a stabilizing *intramolecular* interaction of the protonated O^m by O^6 instead of an *intermolecular* stabilization due to solvent molecules. In the light of the present "wet chemistry" results this intramolecular stabilization looks like an artefact—when compared with the situation in solution—that is caused by the lack of solvent. This conclusion is underscored by our previous observation that protonation of α-D-glucopyranose at the O^4 site in the gas phase leads to release of a water molecule and concurrent formation of a five-membered ring (i.e., the furanose depicted in Figure 2), which is not expected in solution. Thus, the corresponding gas-phase reaction probably follows another pathway from reactants to products than the condensed phase reaction.

Concluding Remarks

A detailed investigation of the mechanism of the specific acid-catalyzed reaction of methanol and α-D-glucopyranose, modeling the formation of a glycosidic bond in liquid water at ambient conditions, is presented based on a set of constrained molecular dynamics simulations. The constraint, being a one-dimensional reaction coordinate, was chosen to be the distance between the oxygen atom of the attacking methanol and the C^1 site of the pyranose ring. After preparing the system relying on classical molecular dynamics, the reactive simulations were carried out within the Car–Parrinello approach to ab initio simulation thus allowing to include efficiently the electronic structure within density functional theory. The solvent is included on the same footing as the reactants and the reaction mechanism is analyzed in terms of structural and energetic aspects as well as changes of the electronic structure. The latter is dissected by using localized orbitals of the Boys/Wannier type and the topology of the electron localization function ELF.

In a nutshell, the reaction is found to proceed—along the enforced reaction coordinate—via a two-step $D_N^*A_N$ mechanism as sketched in the sequences of left panels in Figure 4, because of the apparent necessity of close proximity of the methanol molecule to the successful protonation and subsequent dehydration of the glucopyranose ring. This scenario is akin to the hydrolysis of methyl α- and methyl β-D-glucopyranoside.[2] In particular, the D_N step consists of a concerted protonation of the O^1 hydroxyl group, breaking of the C^1–O^1 bond, and oxocarbenium ion formation involving a C^1=O^5 double bond. The second A_N part is the formation of the C^1–O^m glycosidic bond, deprotonation of the methanol hydroxyl group O^mH^m, and re-formation of the C^1–O^5 single bond. The formed hydronium H_3O^+ detaches from the adduct and diffuses away via Grotthuss structural diffusion in the liquid.

Based on detailed and complementary analyses of the electronic structure along the reaction coordinate it is shown that the intermediate can be classified to be an oxocarbenium cation, see Figure 4f for its Wannier function representation and Figure 7b for ELF. Within the reaction sequence it exists between the D_N and A_N steps, which is confirmed by structural parameters, chemical bonding analyses, and the features of the free energy profile of the entire reaction. The oxocarbenium is non-solvent equilibrated in light of the apparent necessity of a close approach of the methanol O^m before its formation. This was previously found to be the case for many glucopyranoside hydrolysis reactions,[1–3] as well as for the theoretical study of glycosyltransferase.[46]

Most interestingly, water is found to play a crucial role in the reaction sequence. It was possible to show that the formation of the oxocarbenium intermediate correlates with desolvation of the O^5 site. After formation of the glycosidic C^1–O^m bond and thus upon switching back from the C^1=O^5 double bond to the C^1–O^5 single bond the O^5 site becomes again re-solvated. This active role of water correlates well with the previous finding that an *intramolecular* stabilization of the transition state occurs in the gas phase. Beyond this specific example it is expected that a major cause for different reaction mechanisms occurring in the gas phase versus aqueous solution is caused by the missing influence of specific hydrogen-bonding interactions involving solvation shell water. In more general terms this implies that theoretical studies of reactions in associated liquids relying on continuum solvation models, even if hydrogen-bonding effects are approximately included, are likely to lead to erroneous results.

Acknowledgements

We are indebted to J. I. Siepmann for his support and we are grateful to R. Rousseau, T. Ryba, and N. Doltsinis for discussions. The calculations were carried out at MSI (Minneapolis) and at BOVILAB@RUB (Bochum). We thank DAAD, Frieda Martha Kunze Fellowship, University of Minnesota Graduate School Doctoral Dissertation Fellowship, NSF (CTS-0138393), DFG (FOR 436), and FCI for partial financial support.

[1] J. Zhu, A. J. Bennet, *J. Am. Chem. Soc.* **1998**, *120*, 3887.
[2] A. J. Bennet, M. L. Sinnott, *J. Am. Chem. Soc.* **1986**, *108*, 7287.
[3] X. Huang, C. Surry, T. Hiebert, A. J. Bennet, *J. Am. Chem. Soc.* **1995**, *117*, 10614.
[4] M. N. Namchuk, J. D. McCarter, A. Becalski, T. Andrews, S. G. Withers, *J. Am. Chem. Soc.* **2000**, *122*, 1270.
[5] Y. Zhang, J. Bommuswamy M. L. Sinnott, *J. Am. Chem. Soc.* **1994**, *116*, 7557.
[6] N. S. Banait, W. P. Jencks, *J. Am. Chem. Soc.* **1991**, *113*, 7951.
[7] A. Fürstner, F. Jeanjean, P. Razon, *Angew. Chem.* **2002**, *114*, 2203–2206; *Angew. Chem. Int. Ed.* **2002**, *41*, 2097.
[8] A. J. Bennet, T. E. Kitos, *J. Chem. Soc. Perkin Trans. 2* **2002**, 1207.
[9] J. P. Richard, W. P. Jencks, *J. Am. Chem. Soc.* **1984**, *106*, 1373.
[10] T. L. Amyes, W. P. Jencks, *J. Am. Chem. Soc.* **1989**, *111*, 7888.
[11] M. L. Sinnott, W. P. Jencks, *J. Am. Chem. Soc.* **1980**, *102*, 2026.
[12] A. J. Bennet, A. J. Davis, L. Hosie, M. L. Sinnott, *J. Chem. Soc. Perkin Trans. 2* **1987**, 581.
[13] J. Zhu, A. J. Bennet, *J. Org. Chem.* **2000**, *65*, 4423.
[14] R. D. Guthrie, *Pure Appl. Chem.* **1989**, *61*, 23.
[15] R. D. Guthrie, W. P. Jencks, *Acc. Chem. Res.* **1989**, *22*, 343.
[16] B. Capon, *Chem. Rev.* **1969**, *69*, 407.
[17] C. W. Andrews, B. Fraser-Reid, J. P. Bowen, *J. Am. Chem. Soc.* **1991**, *113*, 8293.
[18] N. Buckley, N. J. Oppenheimer, *J. Org. Chem.* **1996**, *61*, 8048.
[19] S. E. Barrows, J. W. Storer, C. J. Cramer, A. D. French, D. G. Truhlar, *J. Comput. Chem.* **1998**, *19*, 1111.
[20] B. J. Smith, *J. Am. Chem. Soc.* **1997**, *119*, 2699.
[21] A. Bérces, G. Enright, T. Nukada, D. M. Whitfield, *J. Am. Chem. Soc.* **2001**, *123*, 5460.
[22] A. R. Dinner, G. M. Blackburn, M. Karplus, *Nature* **2001**, *413*, 752.
[23] C. Molteni, M. Parrinello, *J. Am. Chem. Soc.* **1998**, *120*, 2168.
[24] K. J. Naidoo, J. W. Brady, *J. Am. Chem. Soc.* **1999**, *121*, 2244.
[25] A. Vishnyakov, G. Widmalm, J. Kowalewski, A. Laaksonen, *J. Am. Chem. Soc.* **1999**, *121*, 5403.
[26] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471.
[27] D. Marx, J. Hutter, *Ab Initio Molecular Dynamics: Theory and Implementation, in Modern Methods and Algorithms of Quantum Chemistry* (Ed.: J. Grotendorst), p. 301–449 (NIC, FZ Jülich, **2000**); for downloads see www.theochem.rub.de/go/cprev.html
[28] CPMD, J. Hutter et al., MPI FKF and IBM Zürich (Switzerland), **1995–2001**.
[29] J. M. Stubbs, D. Marx, *J. Am. Chem. Soc.* **2003**, *125*, 10960.
[30] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993.
[31] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.
[32] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.
[33] G. M. Brown, H. A. Levy, *Acta Crystallogr. Sect. B* **1979**, *35*, 656.
[34] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225.
[35] W. Damm, A. Frontera, J. Tirado-Rives, W. L. Jorgensen, *J. Comput. Chem.* **1997**, *18*, 1955.
[36] L. X. Dang, B. M. Pettitt, *J. Phys. Chem.* **1987**, *91*, 3349.
[37] TINKER, J. W. Ponder et al., Washington University School of Medicine, St. Louis, MO 63110 (USA), **1990–2001**, see http://dasher.wustl.edu/tinker/
[38] T. E. Timell, *Can. J. Chem.* **1964**, *42*, 1456.
[39] E. A. Carter, G. Ciccotti, J. T. Hynes, R. Kapral, *Chem. Phys. Lett.* **1989**, *156*, 472.
[40] M. Sprik, G. Ciccotti, *J. Chem. Phys.* **1998**, *109*, 7737.
[41] C. Molteni, M. Parrinello, *Chem. Phys. Lett.* **1997**, *275*, 409.
[42] M. Appell, G. Strati, J. L. Willett, F. A. Momany, *Carbohydr. Res.* **3004**, *339*, 537.
[43] B. S. Jursic, *J. Mol. Struct.* (THEOCHEM) **1999**, *490*, 1.
[44] "Thermodynamics Source Database" in *NIST Chemistry WebBook*, *NIST Standard Reference Database Number 69* (Eds.: P. J. Linstrom, W. G. Mallard), March **2003** (National Institute of Standards and Technology, Gaithersburg, MD 20899), see http://webbook.nist.gov
[45] K. A. Jebber, K. Zhang, C. J. Cassady, A. Chung-Phillips, *J. Am. Chem. Soc.* **1996**, *118*, 10515.
[46] I. Tvaroska, I. André, J. P. Carver, *J. Am. Chem. Soc.* **2000**, *122*, 8762.
[47] D. Indurugalla, A. J. Bennet, *J. Am. Chem. Soc.* **2001**, *123*, 10889.
[48] N. Marzari, D. Vanderbilt, *Phys. Rev. B* **1997**, *56*, 12847.
[49] P. L. Silvestrelli, N. Marzari, D. Vanderbilt, M. Parrinello, *Solid State Commun.* **1998**, *107*, 7.
[50] B. Silvi, A. Savin, *Nature* **1994**, *371*, 683.
[51] A. Savin, R. Nesper, S. Wengert, T. F. Fässler, *Angew. Chem.* **1997**, *109*, 1892; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1809; see also http://www.cpfs.mpg.de/ELF/
[52] D. Marx, A. Savin, *Angew. Chem.* **1997**, *109*, 2168; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2077.
[53] R. Rousseau, D. Marx, *Chem. Eur. J.* **2000**, *6*, 2982.
[54] A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397.
[55] a) D. Marx, M. E. Tuckerman, J. Hutter, M. Parrinello, *Nature* **1999**, *397*, 601; b) J. T. Hynes, *Nature* **1999**, *397*, 565.

Received: July 28, 2004
Revised: November 8, 2004
Published online: February 24, 2005

