JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Unraveling the Base Excision Repair Mechanism of Human DNA Glycosylase

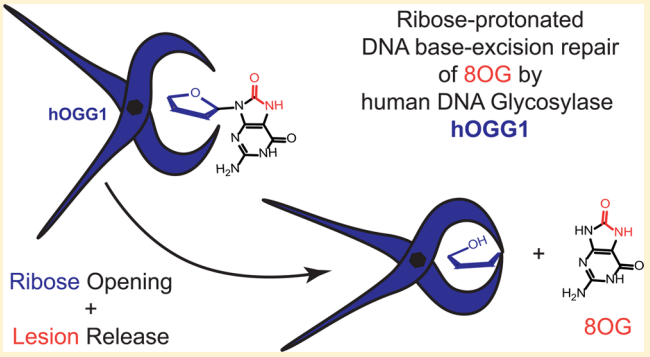
Keyarash Sadeghian and Christian Ochsenfeld*

Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstrasse 7, D-81377 Munich, Germany

Center for Integrated Protein Science Munich (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstrasse 5-13, D-81377 Munich, Germany

Supporting Information

ABSTRACT: Human DNA glycosylase, hOGG1, is known to perform DNA repair by cleaving oxidized guanine (8OG) from the DNA. Despite numerous experimental and theoretical investigations, the underlying selective molecular mechanism has remained a mystery. Here we present a mechanism that explains how hOGG1's catalytic pocket is able to host an undamaged guanine base, but is not able to cleave it from the DNA. Using linear-scaling quantum mechanics/molecular mechanics (QM/MM) techniques with more than 500 atoms in the QM-region, we have investigated previously proposed mechanisms that all rely on protonating the 8OG nucleobase. We have found that the repair mechanisms propagated in the literature to this date are not capable of differentiating between the G and 8OG nucleobase. Besides this nonselectivity, they also involve reaction barriers that are too high, hence rendering the corresponding reaction intermediates inaccessible. Instead, we present a completely different reaction mechanism, where hOGG1 initially targets the ribose moiety of the substrate and cleaves the glycosidic bond at the very last stage. Our ribose-protonated repair mechanism is not only energetically more preferable, but also explains the selectivity utilized by hOGG1 to block processing a guanine base.



INTRODUCTION

DNA is constantly exposed to byproducts of cellular aerobic respiration, such as reactive oxygen species (ROS), which often attack the guanine (G) nucleobase.[1,2] ROS oxidize G and form the 8-oxo-7,8-dihydroguanine lesion (8OG, see Figure 1(a)). The most harmful consequence of 8OG formation is that this lesion can mimic a thymine (T) base during DNA replication and therefore prefers base-pairing with an adenine (A). 8OG:A-containing DNA double-strands, if replicated, may lead to a G:C → T:A transversion mutation (Figure 1(b)) and possibly to cancer.[3] In human 8-oxoguanine DNA glycosylase 1 (hOGG1)[4−7] is a *bi*functional enzyme responsible for the base excision repair (BER) of 8OG. Here the nucleobase of the lesion is cleaved from the DNA (→*glycosylase* activity), followed by the removal of the remaining ribose ring of 8OG (→*AP-lyase* activity).

Although numerous experimental investigations[8−16] have been reported in the literature, a full understanding of the repair mechanism utilized by hOGG1 has remained a challenge. The crystal structures available for the hOGG1-DNA complexes provide us a valuable, but yet inconclusive picture of both the electronic nature as well as the order of the molecular steps involved in 8OG-excision repair. The hOGG1-DNA complex crystal structure, trapped as a reaction intermediate after sodium borohydride treatment (PDB-code: 1HU0, depicted in Figure 2(a)),[13] shows that the enzyme not only cleaves the intended glycosidic bond of 8OG, but also opens the ribose-ring of the lesion. As evident from this structure, besides ribose opening, the anomeric C1′ atom of the lesion undergoes nucleophilic attack by the Lys249 residue of hOGG1. Further support for the crucial role of the latter residue is provided by the crystal structure of 8OG found in the binding site of the catalytically inactive Lys249-Gln mutant of hOGG1 (Figure 2(b)).[14]

An important player that has been almost fully neglected in the discussion of the proposed repair mechanisms for hOGG1 is the Asp268 residue. Experimentally, it is known that the absence of this residue diminishes the catalytic activity of the enzyme.[9] The crystal structure of the Asp268-Asn mutant of hOGG1 with a 8OG-substrate reveals the same binding mode (*syn*) for the lesion as observed in the Lys249 mutant (PDB-code 1N3C,[9] for a close-up view see Figure 2(c)). These findings not only emphasize on the role of the interaction between the ribose ring of the 8OG lesion and the Asp268 residue, but also highlight the protonation state of the latter as a

Received: February 9, 2015
Published: July 30, 2015

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

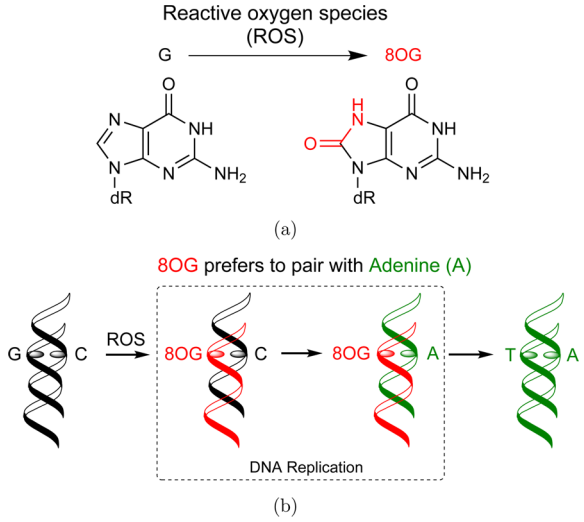


Figure 1. (a) Oxidation of G by reaction oxygen species (ROS) can lead to the formation of the 8OG-lesion. (b) 8OG mimics the thymine (T) base, pairs with adenine (A) and can lead to G:C → A:T mutation.

key issue in the discussion of the base-excision mechanism of the hOGG1 enzyme.

An experimental finding which contributes to the mystery behind the hOGG1 repair mechanism is the crystal structure reported by Crenshaw et al.[15] Here, an intact G base is extruded out of the DNA duplex and has found its way into the catalytic pocket of wild type hOGG1 (see Figure 2(d)). As evident from this crystal structure (PDB-code: 3IH7), the G base is also bound to the enzyme in the *syn* mode and its arrangement closely resembles that of the 8OG lesion in the binding site of the Asp268 and Lys249 hOGG1 mutants. The experimental data suggest that such a hOGG1-G complex is stable for many days and the mentioned G nucleobase is not cleaved by the enzyme. However, the force-field molecular dynamics (FF-MD) simulations performed by Karplus and co-workers[15] show that once the cross-linking groups are removed (in silico) the extruded G-base is no longer stable in the binding site. Hence, the FF-MD data suggest that the unexpected presence of a G-base in the catalytic center of hOGG1 is due the cross-linking procedure employed in the experiment for achieving crystallization. Nevertheless we expect a valid reaction mechanism to be able to explain why an undamaged G base, despite stability and occupation in the very similar position in the active site, is not cleaved by hOGG1. It is precisely for this reason that we decided to use the corresponding crystal structure (PDB: 3IH7) for our QM/MM investigations (see Computational Details, further below).

Theoretical investigations reported in the literature all pursue the idea of a base-protonated base excision repair mechanism, yet they do not clarify the specific role played by both catalytic residues, Lys249 and Asp268. For example, Calvaresi et al.[17] have proposed the possibility of protonating the N3 position of 8OG by the Lys249 residue while Asp268 accepts this proton at a later stage of the reaction mechanism. However, as evident from our results presented further below, the N3-protonated

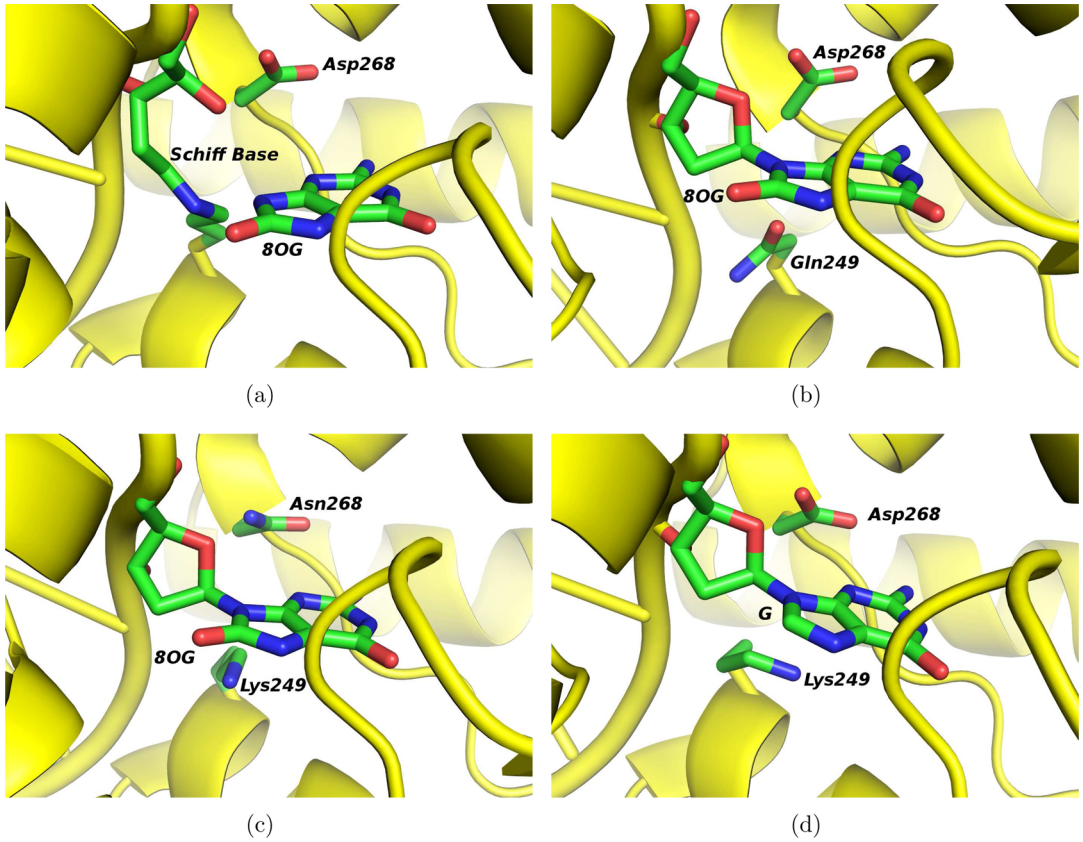


Figure 2. Close-up views of the crystal structures of hOGG1-DNA complexes. (a) Schiff base intermediate after trapped sodium borohydride treatment (PDB-code: 1HU0[13]). (b) 8OG bound to a Lys249-Gln mutant of hOGG1 (PDB-code: 1EBM[14]). (c) 8OG bound to a Asp268-Asn mutant of hoGG1 (PDB-code: 1N3C[9]). (d) G bound to wild type hOGG1 (PDB-code: 3IH7[15]).

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

mechanism does not provide any selectivity on behalf of hOGG1 to prevent the excision of an undamaged G base. In addition, the possibility of other protonation states for these two residues was not explored in ref 17. More recently, Šebera et al.[18] proposed an excision repair mechanism which involves protonation at the N9 position, where the Lys249 residue is assumed to be the proton donor. The authors hold the pyramidal coordination of the lesion's N9 atom toward the Lys249 residue to be responsible for the N9-protonation pathway.[18] Furthermore, their results indicate that the electronic structure of the five-member ring of 8OG is better suited for such N9-pyramidalization than that found in a normal G base.[18] However, the Asp268 residue was not included explicitly in the mechanistic calculations reported in ref 18. Overall, the Lys249 residue has been considered as the proton donor in both mentioned base-protonated mechanisms, whereas the critical role of the Asp268 residue is not directly addressed.

In the present work, we report the results of our QM/MM study on the BER mechanisms proposed for the glycosylase activity of hOGG1 and study the catalytic importance of both Lys249 and Asp268 residues. Our calculations rule out the so far propagated base-protonated mechanisms. On the contrary, we propose a reaction mechanism which is initiated by the protonation of the ribose ring of 8OG where the Asp268 residue provides the required proton. Furthermore, our QM/MM calculations show that the ribose-protonated pathway is much more selective than the other mechanisms. We have verified our data by performing QM/MM calculations with QM-region more than 600 atoms. The results are in line with our previous work on the repair mechanism of the bacterial enzyme, MutM.[19] On the basis of such calculations, we can provide a plausible explanation for the experimental fact that G can be enforced into the catalytic pocket of hOGG1, but not catalytically processed by the repair enzyme.

COMPUTATIONAL DETAILS

Force-Field Setup. For our investigations, we employed the crystal structure of the wild type hOGG1 in complex with a DNA double strand containing a G base, trapped in the active site of the enzyme (PDB-code: 3IH7).[15] For calculations involving the 8OG lesion, the G base was replaced manually by the 8OG lesion. The orientation of all Asn, Gln, and His residues were checked using the MolProbity server.[20] We have checked the protonation states using the H++ server.[21] The His270 was treated as positively charged (double-protonated) as it interacts with the phosphate group of 8OG and the solvent exposed Asp322 residue. We note that Calvaresi et al.[17] had also chosen the same protonation state for H270 in their model system. The LEAP-module of AMBER Tools[22] was used to neutralize the system by adding sodium ions. The neutralized system was then placed in a water box, thereby making sure that every solute molecule (enzyme, DNA, and counterions) is covered by a solvent layer of at least 12 Å thickness. The parameters for 8OG were taken from ref.[23] The standard AMBER-10 force-field parameters and AMBER Tools[22] were used to parametrize the amino- and nucleic acids of the system. The TIP3P water model[24] was used to describe the explicit solvation. The NAMD force-field engine[25] was used throughout this study. The solvated and neutralized structures, with either G or 8OG were each minimized in three steps. These steps involved sequential relaxation of (i) hydrogen atoms, (ii) water molecules, and finally (iii) all atoms of the system.

QM/MM Setup. The force-field minimized structure obtained from each of the 8OG-hOGG1 and G-hOGG1 setups discussed above was taken as the starting point for QM/MM calculations. In order to reduce the computation time of the QM/MM calculations, we decided to remove the outermost solvating water molecules. The fragments which were kept include: (i) the full protein, (ii), the full DNA, (iii) all counterions, (iv) all water molecules where at least one atom is within 5.0 Å of either protein, DNA, or counterions, and (vi) all water molecules where at least one atom is within 12 Å of the 8OG-lesion. These reductions lead to structures with almost 10 000 atoms.

We have used ChemShell[26] as QM/MM interface. The FermiONs++[27] package, developed in our group, and TURBOMOLE[28] were used in this work as QM-codes. The DL-POLY[29] force-field engine integrated in ChemShell was used for the MM part of the calculations. The same set of AMBER force-field parameters used for the classical MD simulations was also used for the MM part of the QM/MM calculations. We have used the additive QM/MM scheme with electrostatic embedding. The charge-shift scheme[30] and link atoms (hydrogen) were employed to deal with the QM-MM boundary region. During the QM/MM structure optimizations which were carried out using the DL-FIND[31] optimizer module of ChemShell, we relaxed all fragments (amino acids, nucleic acids, water counterions) that lie within 12 Å of the N9 atom of the 8OG-residue. Everything beyond this selection is kept frozen during QM/MM structure optimizations.

The reaction pathways, starting from different protonation states, were investigated at the DFT/MM level of theory using the dispersion-corrected BLYP-D3 functional.[32−34] The SVP basis set[35] was used throughout this study. Our previous experience with the QM/MM study of the bacterial repair enzyme, MutM,[19] justifies choosing this combination of method and basis set.

For the QM/MM structure optimization, we have chosen a QM-region, denoted here as QM_0, which consists of 64 atoms. This QM-region consists of the 8OG base and ribose, as well as the Asp268, Lys249, Ser147, and the Cys253 residues of hOGG1 (see Figure 3). For the QM/MM reaction profiles of the guanine base, the nucleobase atoms of the G base were included in the corresponding QM_0 region.

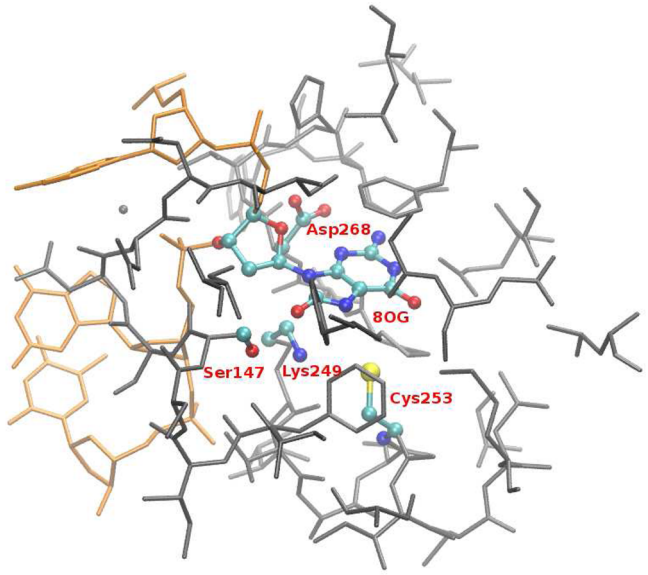


Figure 3. QM-region QM_0 used for the BLYP-D3/AMBER structure optimization is shown in CPK representation and atom color. Those residues which belong to the largest QM-region (containing 734 atoms) used in the QM/MM convergence study of this work, are shown in sticks representation (protein and water in gray, DNA in orange). Hydrogen atoms are omitted for clarity.

In addition to the QM/MM reaction profiles obtained using the QM_0 region, we have also performed a systematic QM/MM convergence study with respect to the QM-region size using linear-scaling QM-methods (see ref 36 for a recent review) implemented in the FermiONs++[27] program package. For this purpose we took the QM_0 region as reference and then selected those residues which have at least one atom within R (R = 2.0, 2.5, ..., 5.5 Å) of the QM_0 region.

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

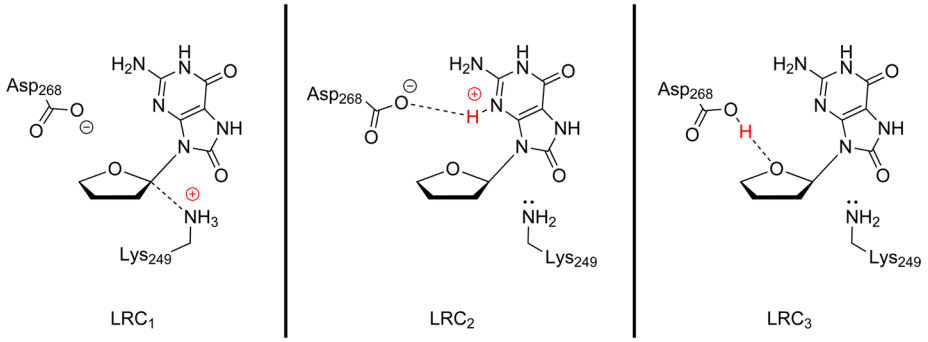


Figure 4. Protonation states, defined here as LRC_1, LRC_2, and LRC_3, which have been considered in this work as the starting point for the N9-, N3-, and ribose-protonated pathways, respectively. In LRC_1: N3 atom of 8OG is protonated, Lys249 side chain is neutral and Asp268 is treated as a carboxylate (charge −1). In LRC_2: Lys249 and Asp268 residues are treated in their standard protonation states (charge +1 and −1, respectively). In LRC_3: both Asp268 residue (protonated) Lys249 side chains are neutral.

The largest QM-region within this selection comprises 734 atoms (see Figure 3). Full details of the selected residues can be found in the SI. For the convergence study, single point B3LYP-D3/AMBER calculations on selected structures along the reaction profiles were performed (see further below). We also performed numerous attempts to obtain BLYP-D3/AMBER energies for QM-regions larger than QM_0, but we faced SCF convergence problems. This is caused by the inclusion of charged residues in the larger QM regions which lead to a small HOMO−LUMO gap in the BLYP-D3/AMBER calculations. This is why we switched to the B3LYP-D3 functional for QM/MM calculations with larger spheres.

RESULTS AND DISCUSSION

We present our QM/MM data in the following order. First we discuss the protonation states of the Asp268 and Lys249 residues in the active site of hOGG1. We then present our QM/MM reaction profiles for the N9-protonated mechanism (Šebera et al.[18] refer to this mechanism as σ-bond substitution), followed by the N3-protonation mechanism as suggested by Calvaresi et al.[17] We then present our QM/MM study of the alternative ribose-protonated pathway, which to our best knowledge has not been so far proposed for hOGG1. For all studied pathways, we also present the corresponding energy profile for a guanine substrate in the binding site of hOGG1.

Protonation State of the Catalytic Residues. We have investigated three different protonation states, LRC_{1-3}, for the catalytic pocket of hOGG1 (see Figure 4 for definition). These are three isomer states which differ in the position of a proton which is placed either at the Lys249 (LRC_1), Asp268 (LRC_3) residues, or at the N3-atom of the 8OG-lesion (LRC_2).

Our QM/MM isomerization energies (Table 1) indicate that for the 8OG-hOGG1 complex, the LRC_3 structure with a protonated Asp268 residue is energetically the most favorable state. The LRC_2 structure, with the N3 position of 8OG protonated, is with 2 kcal/mol above LRC_3 slightly less favorable. Finally the LRC_1 structure, with the protonated Lys249 residue, is 16 kcal/mol above LRC_3. Already, at this stage one could rule out the N9-protonated mechanism, since its starting structure (LRC_3) is not sufficiently populated at room temperature. We have nevertheless investigated this mechanism using QM/MM techniques and show that the reaction barriers involved are also too high to be of any relevance (see further below).

Table 1. BLYP-D3/AMBER Isomerization Energies (kcal/mol) for the Protonation States LRC_1 and LRC_2 Where the QM/MM Energy of the LRC_3 State Is Taken As Reference

substrate	LRC_1	LRC_2
8OG	16.2	1.5
G	12.9	−2.6

N9-Protonation Mechanism. On the basis of our QM/MM calculations we consider the N9-protonation (σ-bond substitution) mechanism to be rather unlikely. Our QM/MM calculations show that in order to reach the Schiff base intermediate state (SBI, closed form), a rate-determining barrier of about 54 kcal/mol would need to be surmounted (see Figure 5):[39] The oxacarbenium ion intermediate reached at the end of the first step is with +27 kcal/mol above LRC1 a high energy structure. From here the oxacarbenium ion undergoes a nucleophilic attack by the Lys249 residue and the SBI is formed. For this pathway we find the minimum energy path to involve a high energy structure (+54 kcal/mol above LRC1, Figure 5). All these energetics provide strong evidence against the N9-protonated mechanism.

We have also investigated a N9-protonated pathway for a G base. Our data, also shown in Figure 5, indicate a larger reaction barrier for the initial glycosidic bond of the G base than that observed for the 8OG (45 vs 35 kcal/mol). The oxacarbenium ion intermediate found for the G base (+24 kcal/mol) is higher than the corresponding LRC_3 structure. Interestingly, the pathway toward the SBI state involves a much lower barrier (high energy structure +33 kcal/mol) than that of 8OG (+54 kcal/mol, see above).

Overall, the N9-protonation mechanism is neither able to differentiate between damaged (8OG) and undamaged (G) bases, nor does it show favorable energetics. On the basis of these arguments, we can therefore rule out the N9-pathway as a plausible candidate for the repair mechanism.

N3-Protonation Mechanism. Our QM/MM investigations also rule out the N3-protonated mechanism as an energetically favorable pathway for hOGG1 (Figure 6). Starting from the N3-protonated structure (LRC_2) we have followed the minimum QM/MM energy reaction path along the cleavage of the glycosidic bond. This process is accompanied by the nucleophilic attack of the Lys249 residue on the anomeric C1′ of 8OG-ribose. Our QM/MM reaction profile indicates high energy structures above 60 kcal/mol along the pathway for such a scenario. We note a spontaneous proton transfer from the Lys249 residue to the N9 of 8OG after surmounting the transition state structure. The resulting

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

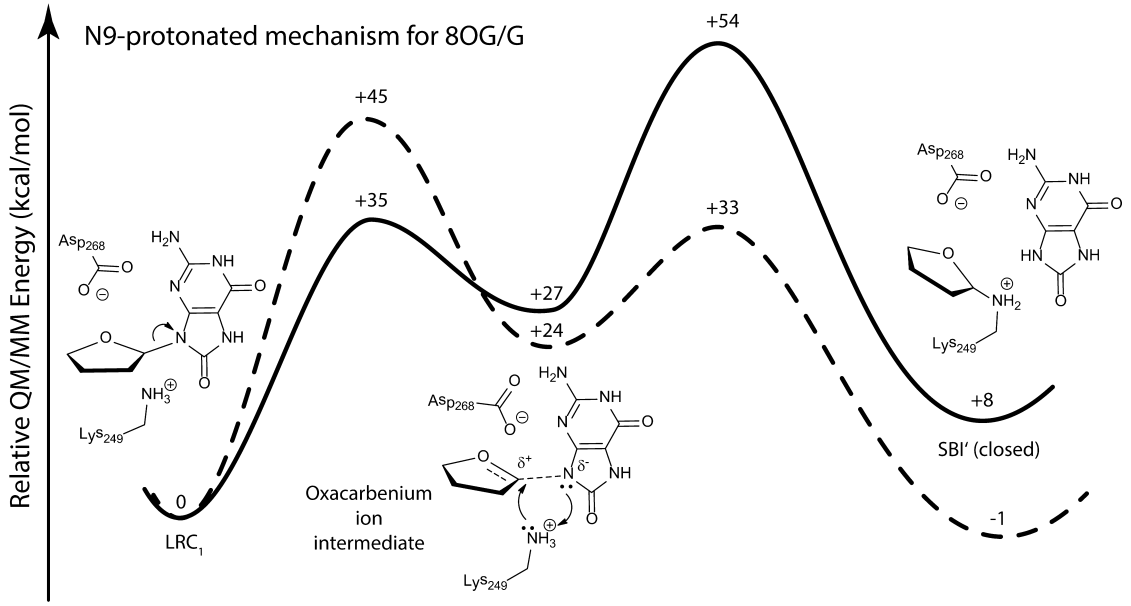


Figure 5. Relative BLYP-D3/AMBER QM/MM energies (kcal/mol) of the N9-protonated pathway for both 8OG (solid black line) and G (dashed red line) substrates. Note that for the 8OG and G substrates, the LRC_1 structures taken here as reference are 16 and 13 kcal/mol above their corresponding LRC_3 structures, respectively (see Figure 4).

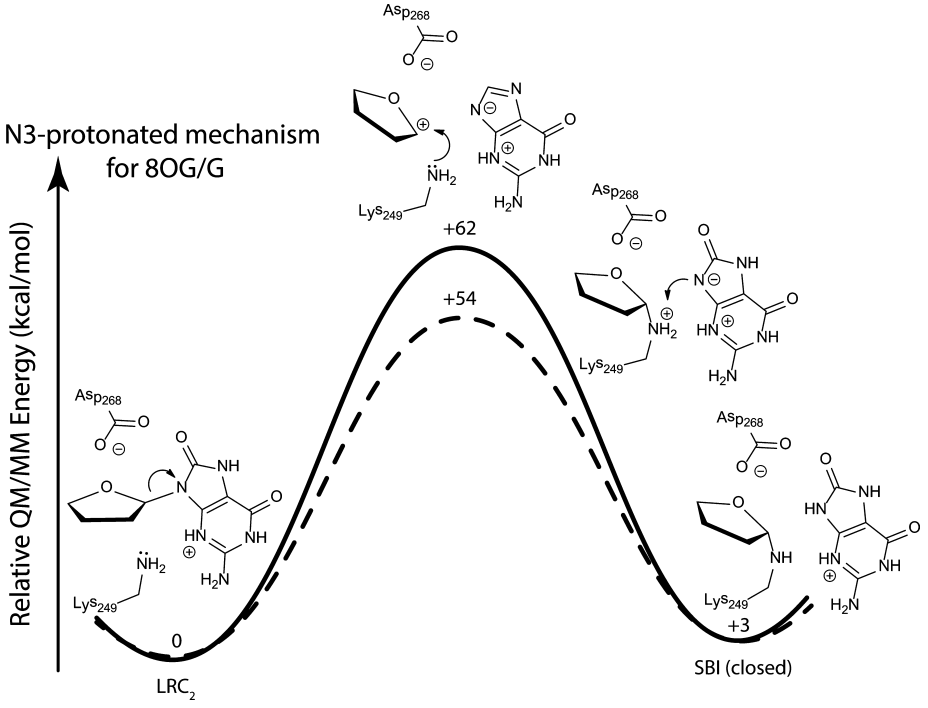


Figure 6. Relative BLYP-D3/AMBER QM/MM energies (kcal/mol) along the N3-protonated reaction pathway for both 8OG (solid black line) and G (dashed red line) substrates. Note that for the 8OG and G substrates, the LRC_2 structures taken here as reference are 2 kcal/mol above and 3 kcal/mol below the corresponding LRC_3 structures, respectively (see Figure 4).

product is a Schiff base intermediate with a closed form of the ribose is 3 kcal/mol above LRC_2.

Further evidence against the N3-protonated mechanism is provided by our QM/MM investigations of this pathway for a G base as substrate (Figure 6). A selective mechanism would imply a considerably higher barrier for the cleavage of a G base compared with that of 8OG. However, as our results indicate, the corresponding pathway for G involves a barrier of about 54 kcal/mol which is lower that the corresponding barrier of 62 kcal/mol for the 8OG excision. Overall, both barriers (62 and 54 kcal/mol for 8OG and G, respectively) are far too high to be of any relevance.

Ribose-Protonated Mechanism. In contrast to the base-protonated mechanism discussed above, we find a ribose-protonated pathway to be the energetically most favorable pathway. Our proposed mechanism together with the corresponding QM/MM reaction profile is presented in Figure 7. The first step of this mechanism is the opening of the ribose-

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

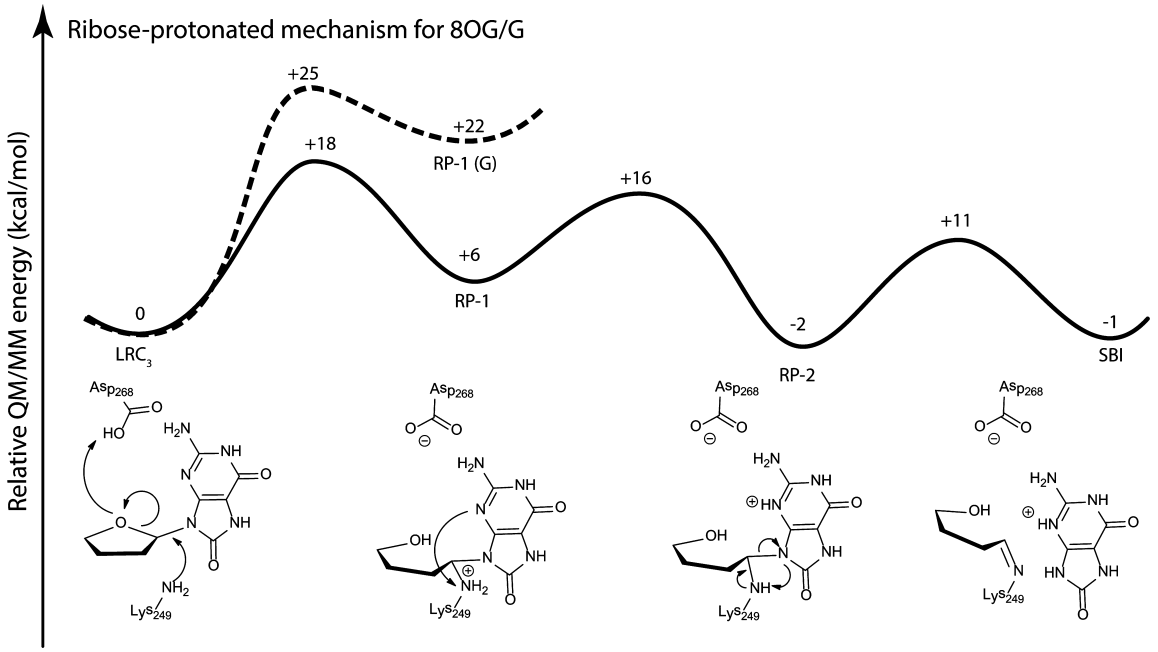


Figure 7. Relative BLYP-D3/AMBER QM/MM energies (kcal/mol, LRC_3 taken as reference) along the ribose-activated reaction pathway for 8OG and G substrates.

ring via the Asp268 residue in its protonated form (LRC_3 structure, Figure 4). We find this step to be rate-determining for the whole reaction mechanism and we estimate a barrier of about 18 kcal/mol. The first intermediate formed, RP-1 (+5 kcal/mol above LRC), also involves the nucleophilic attack of Lys249 residue on the anomeric C1′ of 8OG's ribose. Already the first step of this mechanism addresses the catalytic roles of the two Asp268 and Lys249 residues. Both the ribose-ring opening of 8OG and the Lys249-C1′ covalent-bond formation are the two features that are visible in the crystal structure of the Schiff base intermediate (Figure 2(a)). From RP-1 the reaction proceeds by a proton transfer from the covalently attached Lys249 to the N3-position of 8OG, leading to the intermediate, RP-2 (−2 kcal/mol). For the transition of RP-1 to RP-2 we find that highest energy structure is only 16 kcal/mol above the LRC. We also note that the reaction energy of the RP-1 → RP-2 is −8 kcal/mol and therefore thermodynamically favorable.

An interesting aspect of our alternative ribose-protonated mechanism is that in contrast to all previous mechanisms, the process of glycosidic cleavage occurs at the very last step of the reaction (RP-2 → SBI). It seems as if hOGG1 uses this order of events to lower the barrier drastically. We find the glycosidic bond cleavage to have a reaction barrier of about 15 kcal/mol. One has to compare this value to those for the glycosidic cleavage in the base-protonated mechanisms (35 and 62 kcal/mol for the N9- and N3-protonated pathways, respectively). Overall we find the ribose-protonated pathway to be energetically far more efficient than the base-protonated mechanisms proposed so far.

The important question raises at this stage whether our ribose-protonated mechanism can explain why hOGG1 does not cleave a G-base which has been trapped in its catalytic pocket? Interestingly we have found that our ribose-protonated pathway can indeed discriminate between the G and 8OG nucleobases. We have followed the same ribose-protonated pathway for a guanine base. Our QM/MM data, shown in Figure 7, indicate that the first step, i.e., ribose-opening and nucleophilic attack of Lys249 on C1′, involves a reaction barrier of about 25 kcal/mol. Compared with the barrier estimate of 18 kcal/mol for the 8OG base, we are dealing with a rather weak kinetic discrimination. However, the QM/MM relative energy of the RP-1 intermediate for G-base lies 22 kcal/mol above its corresponding LRC state (compared with +6 kcal/mol for the 8OG case). We can therefore conclude that the thermodynamic discrimination in the first step of the ribose-protonated repair mechanism allows hOGG1 to avoid cleaving a G base from the DNA. We also note that, as shown by Crenshaw and coworkers,[15] the G base is only held in the active site by the cross-linking procedure. It is therefore anticipated that the hOGG1, in vivo is able to enforce a successful substrate discrimination before allowing the DNA base to reach the catalytic site.

QM/MM Convergence Investigations. We have carried out a QM/MM convergence study by performing a series of single point QM/MM calculations to make sure that the energetics of the discussed mechanisms obtained with the rather small QM_0 region are also valid for larger QM-regions (see Computational Details and SI for further details). We consider the QM/MM energetics to be converged if the energy change upon QM-region extension is below 2 kcal/mol. We have found that this criterion to be fulfilled for QM-regions which include at least 734 atoms of the catalytic pocket. In the following we will discuss the QM/MM convergence profile for the protonation states LRC_{1-3} and the N9-, N3-, and ribose-protonated mechanisms (see Figure 8).

Our convergence study for the 8OG-hOGG1 complex, confirms the above-discussed finding (Table 1) that the protonation states LRC_1 and LRC_2 are both higher in energy than LRC_3. As shown in Figure 8(a), the isomerization B3LYP-D3/AMBER energies converge within 2 kcal/mol upon including 734 atoms of the catalytic pocket into the QM-region. In this case, we show that the LRC_1 and LRC_2 states are 4.1 and 11.8 kcal/mol above LRC_3, respectively. These results

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

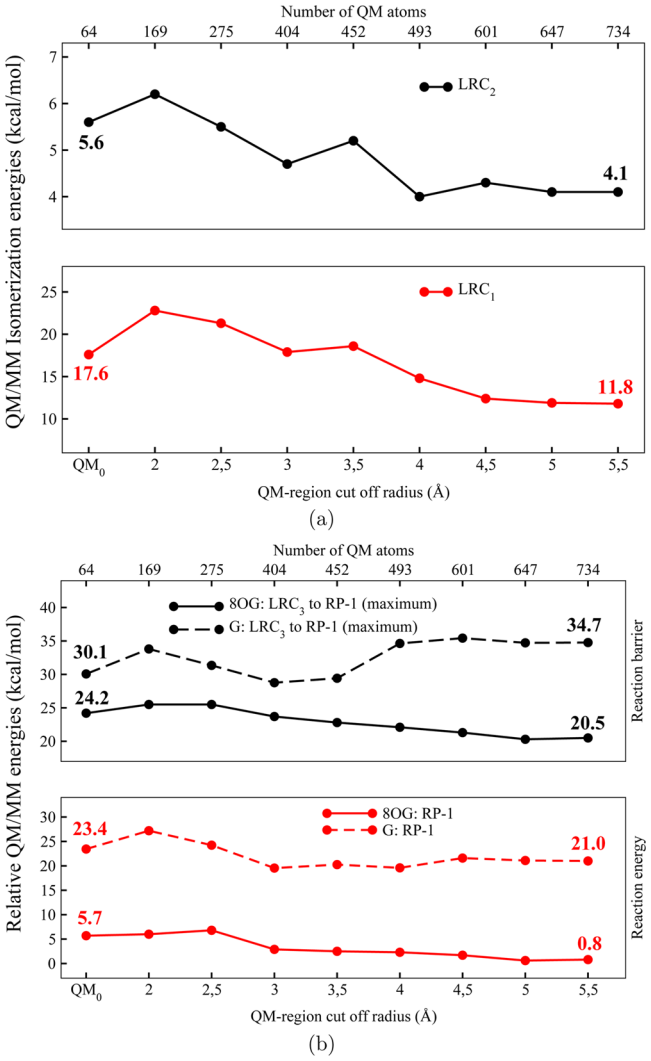


Figure 8. Convergence of the QM/MM (B3LYP-D3/AMBER) energies (in kcal/mol) with respect to the QM-region. (a) Isomerization energies for the protonation states LRC_{1-2}, where the energy of the LRC_3 is taken as reference. (b) Reaction barriers (labeled as maximum, top) and reaction energies (bottom) for first step of the ribose-protonated pathway. Note that the corresponding data presented in Table 1 and Figure 7 are obtained via BLYP-D3/AMBER QM/MM structure optimization (QM_0).

confirm our findings that the N3- and N9-protonated pathways both start from an energetically unfavorable structure.

For the ribose-protonated pathway, the QM/MM convergence profile for the reaction-energies and -barriers in the first rate-determining step ($LRC_3 \rightarrow RP1$) is shown in Figure 8(b). Here, we obtain a converged reaction barrier of 20.5 kcal/mol for a QM-region consisting of 647 atoms. The relative QM/MM energy of the RP1 state, i.e., the reaction energy, also reaches its converged value of 0.8 kcal/mol (above LRC_3) for the same QM-region size. The overall trend obtained here is in agreement with the QM/MM BLYP-D3/Amber results presented in Figure 7.

We have also applied our QM/MM convergence calculation scheme on the ribose-protonated pathway with a G-base in the catalytic pocket. The corresponding QM/MM convergence profile, shown in Figure 8(b), supports the BLYP-D3/Amber trend already presented in Figure 7. For the G-substrate, we find the converged barrier for the first step of this mechanism to be 34.7 kcal/mol above the corresponding LRC_3 state. The corresponding QM/MM converged reaction energy is 21.0 kcal/mol (G-RP-1, see Figure 8(b)). Overall, our QM/MM convergence study confirms the idea that the ribose-protonated pathway is able to discriminate between the 8OG and G substrates.

For the sake of comparison, we have performed a QM/MM study on the N9- and N3-protonated pathways. As for the 8OG-hOGG1 complex, the QM/MM data (shown in Figure SI.2) confirm the trend observed with the QM_0-region. For example, the cleavage of the glycosidic bond after protonation at the N3-position, is still an energetically demanding process. We find the converged QM/MM B3LYP-D3/Amber barrier for this process to be 67.0 kcal/mol (Figure SI.2). Also the formation of the oxacarbenium ion intermediate during the N9-protonation is found to be energetically not favorable. In this case, reaction barriers and energies converge to 40.3 and 31.5 kcal/mol, respectively (Figure SI.2).

CONCLUSION

In this work, we have unraveled the repair mechanism utilized by the human DNA glycosylase, hOGG1, to perform excision of the 8OG lesion from DNA. We show that mechanisms proposed in the past relying on the idea of the initial protonation of the 8OG-nucleobase, involve highly unfavorable reaction energies and barriers. Furthermore, our calculations show that base-protonated mechanisms could also not discriminate between the 8OG and G nucleobases.

We propose a ribose-protonated mechanism that not only exhibits much more preferable energetics, but can also provide a basis for the discrimination of the enzyme. In this alternative mechanism, both the opening of the ribose-ring and the nucleophilic attack of the enzyme initiate the whole reaction cascade. Already at this very first step, hOGG1 imposes an energetic discrimination on the G base, which leads to the destabilization of the corresponding reaction intermediate. For the 8OG lesion, however, the reaction mechanism proceeds and the intended glycosidic cleavage takes place at the very last step. Overall, our QM/MM calculations provide a plausible explanation for the experimental findings including the catalytic role of the conserved amino acids in the reaction center of hOGG1.

The present findings for the human enzyme are also in line with our recent theoretical investigations[19,37] on the bacterial DNA glycosylase, MutM (also known as Fpg). There, we have also proposed ribose-protonated pathways for the base excision of the 8OG[19] and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapydG)[37] lesions from the DNA. Noteworthy is that the binding mode of the 8OG toward hOGG1 and MutM differs. While the 8OG lesion bind to MutM in the *syn*[38] mode, its binding mode to hOGG1 is *anti*.

Overall, both our theoretical data and the experimental results available indicate that bacterial (hOGG1) and bacterial (MutM) enzymes employ a common strategy to catalyze the excision of oxidized nucleobases from the DNA.

ASSOCIATED CONTENT

Supporting Information

Detailed information on the force-field and QM/MM setup as well as the coordinates of the reaction intermediates. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b01449.

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831

AUTHOR INFORMATION

Corresponding Author

*christian.ochsenfeld@uni-muenchen.de

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge financial support by the DFG funding initiatives SFB749 (C7) and the Excellence Cluster EXC114 (CIPSM).

REFERENCES

(1) Lindahl, T. *Nature* **1993**, *362*, 709.
(2) Friedberg, E. C. *Nature* **2003**, *421*, 436−440.
(3) Hollstein, M.; Shomer, B.; Greenblatt, M.; Soussi, T.; Hovig, E.; Montesano, R.; Harris, C. C. *Nucleic Acids Res.* **1996**, *24*, 141−6.
(4) Roldan-Arjona, T. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 8016−8020.
(5) Bjoras, M. *EMBO J.* **1997**, *16*, 6314−6322.
(6) Rosenquist, T. A.; Zharkov, D. O.; Grollman, A. P. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 7429−7434.
(7) David, S. S.; O'Shea, V. L.; Kundu, S. *Nature* **2007**, *447*, 941−50.
(8) Norman, D. P. G.; Bruner, S. D.; Verdine, G. L. *J. Am. Chem. Soc.* **2001**, *123*, 359−360.
(9) Norman, D. P. G.; Chung, S. J.; Verdine, G. L. *Biochemistry* **2003**, *42*, 1564−72.
(10) Chung, S. J.; Verdine, G. L. *Chem. Biol.* **2004**, *11*, 1643−9.
(11) Morland, I.; Luna, L.; Gustad, E.; Seeberg, E.; Bjørås, M. *DNA Repair* **2005**, *4*, 381−7.
(12) Vidal, A. E. *Nucleic Acids Res.* **2001**, *29*, 1285−1292.
(13) Fromme, J. C.; Bruner, S. D.; Yang, W.; Karplus, M.; Verdine, G. L. *Nat. Struct. Biol.* **2003**, *10*, 204−11.
(14) Bruner, S. D.; Norman, D. P.; Verdine, G. L. *Nature* **2000**, *403*, 859−866.
(15) Crenshaw, C. M.; Nam, K.; Oo, K.; Kutchukian, P. S.; Bowman, B. R.; Karplus, M.; Verdine, G. L. *J. Biol. Chem.* **2012**, *287*, 24916−28.
(16) McKibbin, P. L.; Kobori, A.; Taniguchi, Y.; Kool, E. T.; David, S. S. *J. Am. Chem. Soc.* **2012**, *134*, 1653−61.
(17) Calvaresi, M.; Bottoni, A.; Garavelli, M. *J. Phys. Chem. B* **2007**, *111*, 6557−70.
(18) Šebera, J.; Trantírek, L.; Tanaka, Y.; Sychrovský, V. *J. Phys. Chem. B* **2012**, *116*, 12535−44.
(19) Sadeghian, K.; Flaig, D.; Blank, I. D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C. *Angew. Chem., Int. Ed.* **2014**, *53*, 10044−8.
(20) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., III; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, 375−83.
(21) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368.
(22) Case, D.; Darden, T.; Cheatham, T. E., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D.; Seetin, M.; Sagui, C.; Babin, V.; Kollman, P. *AMBER 10*; University of California: San Francisco, 2008.
(23) Miller, J. H.; Fan-Chiang, C.-C. P.; Straatsma, T. P.; Kennedy, M. a. *J. Am. Chem. Soc.* **2003**, *125*, 6331−6.
(24) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(25) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(26) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct.: THEOCHEM* **2003**, *1*, 632.
(27) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 134114.
(28) *TURBOMOLE*, version 6.5; TURBOMOLE GmbH: Karlsruhe, Germany, 2013; http://www.turbomole.com.
(29) Smith, W.; Yong, C. W.; Rodger, P. M. *Mol. Simul.* **2002**, *28*, 385−471.
(30) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133−6141.
(31) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. *J. Phys. Chem. A* **2009**, *113*, 11856−65.
(32) Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.
(33) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(34) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(35) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571−2577.
(36) Kussmann, J.; Beer, M.; Ochsenfeld, C. *WIREs Comp. Mol. Sci.* **2013**, *3*, 614−636.
(37) Blank, I. D.; Sadeghian, K.; Ochsenfeld, C. *Sci. Rep.* **2015**, *5*, 10369.
(38) Fromme, J. C.; Verdine, G. L. *J. Biol. Chem.* **2003**, *278*, 51543−51548.
(39) We found it much more likely that the glycosidic is broken first, thus forming the oxacarbenium ion, before N9 can be protonated.

DOI: 10.1021/jacs.5b01449
J. Am. Chem. Soc. 2015, 137, 9824−9831