

# Comparison of ab Initio, DFT, and Semiempirical QM/MM Approaches for Description of Catalytic Mechanism of Hairpin Ribozyme

Vojtěch Mlýnský,<sup>†</sup> Pavel Banáš,<sup>†</sup> Jiří Šponer,<sup>‡,§</sup> Marc W. van der Kamp,<sup>||</sup> Adrian J. Mulholland,<sup>\*,||</sup> and Michal Otyepka<sup>\*,†</sup>

<sup>†</sup>Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, tr. 17 listopadu 12, 771 46, Olomouc, Czech Republic

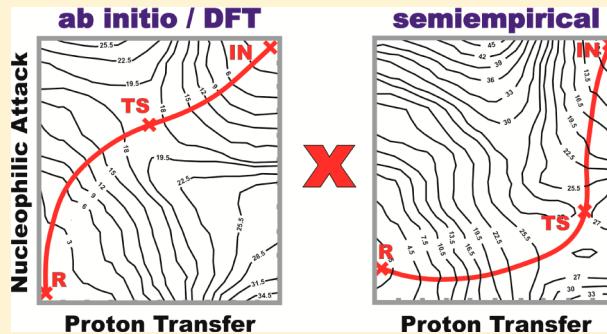
<sup>‡</sup>Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic

<sup>§</sup>CEITEC—Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

<sup>||</sup>Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, U.K.

## S Supporting Information

**ABSTRACT:** We have analyzed the capability of state-of-the-art multiscale computational approaches to provide atomic-resolution electronic structure insights into possible catalytic scenarios of the hairpin ribozyme by evaluating potential and free energy surfaces of the reactions by various hybrid QM/MM methods. The hairpin ribozyme is a unique catalytic RNA that achieves rate acceleration similar to other small self-cleaving ribozymes but without direct metal ion participation. Guanine 8 (G8) and adenine 38 (A38) have been identified as the catalytically essential nucleobases. However, their exact catalytic roles are still being investigated. In line with the available experimental data, we considered two reaction scenarios involving protonated A38H<sup>+</sup> as a general acid which is further assisted by either canonical G8 or deprotonated G8<sup>-</sup> forms. We used the spin-component scaled Møller–Plesset (SCS-MP2) method at the complete basis set limit as the reference method. The semiempirical AM1/d-PhoT and SCC-DFTBPR methods provided acceptable activation barriers with respect to the SCS-MP2 data but predicted significantly different reaction pathways. DFT functionals (BLYP and MPW1K) yielded the same reaction pathway as the SCS-MP2 method. The activation barriers were slightly underestimated by the GGA BLYP functional, although with accuracy comparable to the semiempirical methods. The SCS-MP2 method and hybrid MPW1K functional gave activation barriers that were closest to those derived from experimentally measured rate constants.



## INTRODUCTION

A group of small self-cleaving ribozymes (also called nucleolytic ribozymes) catalyzes reversible phosphodiester bond cleavage and ligation, i.e., internal transesterifications, which are considered to be fundamental nucleic acid reactions.<sup>1,2</sup> The catalytic reaction of the hairpin ribozyme is unique among small self-cleaving ribozymes. It achieves comparable rate acceleration but without the presence of a metal ion in the active site,<sup>3,4</sup> which makes the hairpin ribozyme an especially useful model for understanding how ribozymes achieve catalysis.<sup>5,6</sup> The scissile phosphate is located between residues adenine –1 (A-1) and guanine +1 (G+1). Structural and biochemical data have identified two other nucleobases, guanine 8 (G8) and adenine 38 (A38), as the key nucleobases involved in the catalysis. The cleavage reaction is initiated by nucleophilic attack of O2' on the adjacent scissile phosphate and proceeds through a pentacoordinated phosphorane transition state,

generating 2',3'-cyclic phosphate and 5'-OH terminated RNA fragments.<sup>1</sup> The exact mechanism by which small self-cleaving ribozymes realize that cleavage is still debated, although several catalytic strategies have been identified, including (i) active site organization required for the in-line attack (the angle between O2', P, and O5' atoms is close to 180°), (ii) activation of the 2'-OH nucleophile by a general base, (iii) transition state stabilization, and (iv) protonation of O5' by a general acid.<sup>1,6,7</sup>

The potential roles of G8 and A38 in hairpin ribozyme catalysis have been addressed in numerous experimental and theoretical studies. In precleavage (containing 2'-O-methyl substitution to inhibit the reaction), product (P), and transition state (TS) analogue crystal structures, G8 is positioned close to the 2'-OH nucleophile and is predisposed for donating a

Received: November 23, 2013

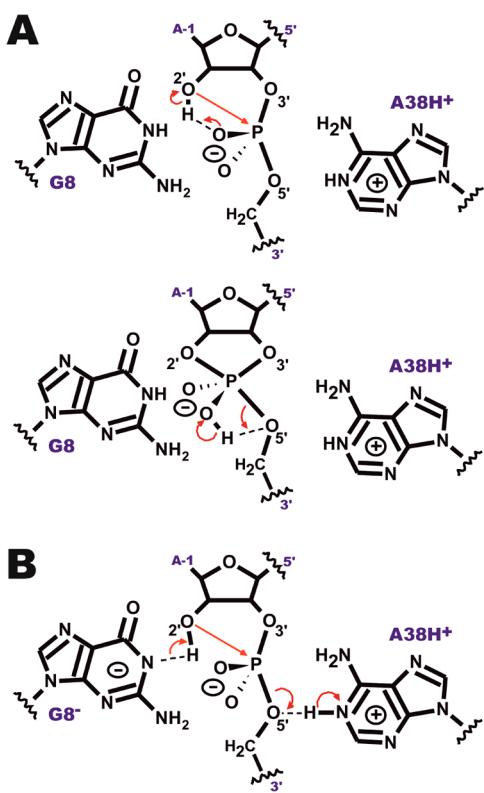
Published: February 13, 2014



hydrogen bond from N1 to the 2'-oxygen of A-1.<sup>8–11</sup> These observations, together with exogenous nucleobase rescue experiments, have led to the proposal that G8 electrostatically stabilizes the TS and/or acts as a general base (which is deprotonated before the reaction).<sup>6,12–15</sup>

Extensive hybrid quantum mechanical/molecular mechanical (QM/MM) computations using the AM1/d-PhoT semiempirical method<sup>16</sup> by Nam and co-workers have pointed toward a mechanism in which the A-1(2'-OH) group is deprotonated by the nonbridging oxygen of the scissile phosphate and canonical G8 participates indirectly in the reaction by electrostatic stabilization.<sup>17</sup> A mechanism in which a nonbridging oxygen acts as a proton shuttle, i.e., the monoanionic scenario (Scheme 1A), was also supported by

**Scheme 1.** Two Reaction Scenarios for the Self-Cleavage Reaction Pathway of the Hairpin Ribozyme Used in This Study<sup>a</sup>



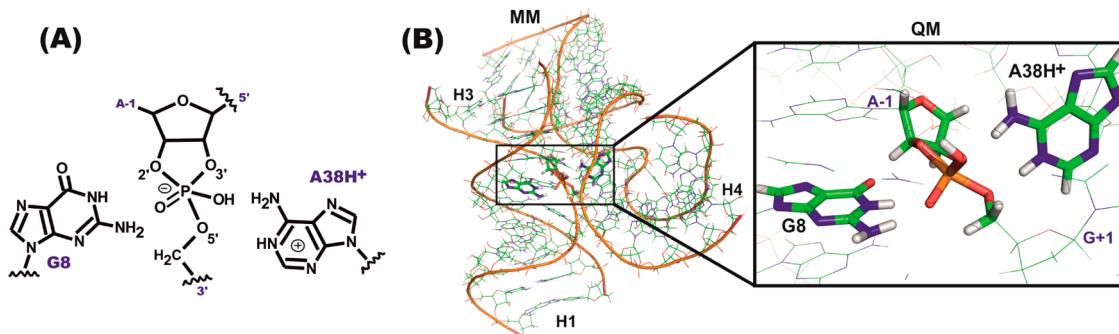
<sup>a</sup>(A) The monoanionic mechanism (proton shuttling scenario), where a nonbridging oxygen of the scissile phosphate accepts a proton from the A-1(2'-OH) nucleophile and transfers it to the leaving G+1(O5') group. The nucleobases (G8 and A38H<sup>+</sup>) are not directly involved in any proton transfer. (B) The dianionic mechanism (combined general base/general acid scenario), in which the G8<sup>-</sup> acts as a general base accepting a proton from the A-1(2'-OH) nucleophile and A38H<sup>+</sup> acts as a general acid protonating the leaving G+1(O5') alcoholate.

our subsequent density functional theory (DFT) QM/MM calculations.<sup>18</sup> Note that these two QM/MM approaches differ considerably (large sampling with lower-level QM method vs limited sampling with higher-level QM method), and thus their consistent suggestion of the nonbridging oxygen shuttle mechanism as a viable scenario is significant. However, the monoanionic mechanism is not supported by experimentally measured pH profiles<sup>3,4,19–21</sup> and is not widely accepted due to the low pK<sub>a</sub> ( $\sim 1^{22}$ ) of an intact phosphate group. Nevertheless,

the significantly higher pK<sub>a</sub> (from 11 to 15<sup>22,23</sup>) of the phosphorane intermediate suggests that the increased basicity of nonbridging oxygen upon nucleophilic attack may promote proton transfer. Therefore, the monoanionic mechanism should be considered as a plausible scenario, competing with and/or accompanying the actual reaction mechanism in small ribozyme catalysis.<sup>18</sup>

Unlike the semiempirical AM1/d-PhoT QM/MM approach used by Nam and co-workers,<sup>24</sup> our previous QM/MM calculations using the hybrid DFT MPW1K functional showed that the mechanism with deprotonated G8<sup>-</sup> acting as a general base (the dianionic scenario, Scheme 1B) is also energetically feasible, giving a comparable activation barrier to the monoanionic mechanism and resulting in pH profiles that are in agreement with experimental pH rate profiles.<sup>18</sup> Namely, the recently estimated pK<sub>a</sub> of G8 (9.5<sup>25</sup>) in the active site of the hairpin ribozyme fits well with experimental profiles showing rate constants as functions of pH.<sup>13,26</sup> The possibility of a general base mechanism with deprotonated G8<sup>-</sup> was suggested earlier by Bevilacqua<sup>13</sup> and most recently by Wilson and Lilley.<sup>27,28</sup> However, computational studies have one considerable uncertainty when dealing with this mechanism. The QM/MM calculations have to be initiated from geometries where the deprotonated G8<sup>-</sup> is in an active conformation, ready for catalysis. They do not deal with the mechanism of the G8 deprotonation. In large-scale classical (MM) molecular dynamics (MD) simulations, deprotonated G8<sup>-</sup> forms only a transient A-1(2'-OH)...G8<sup>-</sup>(N1) H-bond, and subsequently departs from the active site due to electrostatic repulsion from the negatively charged scissile phosphate.<sup>29</sup> Since the guanine can hardly be deprotonated when donating the H-bond to the 2'-hydroxyl nucleophile, its deprotonation is likely to occur when it is oriented away the 2'-hydroxyl nucleophile. However, previous simulations have not suggested any rearrangement that could subsequently lead to a catalytically potent geometry. Instead, they indicate strong repulsion between G8<sup>-</sup> and the scissile phosphate, preventing G8<sup>-</sup> from easily acting as a general base. However, it should be noted that the results from MD simulations should be interpreted with care. The MM description is based on an approximation of permanent partial charges, which do not change (are not polarizable) in classical force fields. Thus, the MM description of structural dynamics of the active site with two closely spaced molecular anions may be inaccurate.<sup>30,31</sup>

Protonated A38H<sup>+</sup> has been suggested to directly participate in the reaction pathway and/or play an indirect role in structural and electrostatic stabilization. The N1 nitrogen of A38 is located near the leaving G+1(O5') oxygen in crystal structures of TS analogues.<sup>9,32–34</sup> Exogenous nucleobase rescue experiments have suggested that A38(N1) is directly involved in the cleavage reaction.<sup>19,35</sup> Raman crystallography has revealed an elevated pK<sub>a</sub> of A38 (5.5,<sup>20</sup> up from 4.3<sup>36</sup>), leading to proposals that A38 might be protonated under physiological pH  $\sim 7$  prior to cleavage.<sup>20,21</sup> Other roles of A38 may include alignment of reactive groups and electrostatic stabilization of the negatively charged phosphorane TS.<sup>19,35,37</sup> MD simulations have shown that protonated A38H<sup>+</sup> is the protonation form most consistent with available X-ray structures.<sup>29</sup> QM/MM computations with the semiempirical AM1/d-PhoT method have concluded that either canonical A38 stabilizes the TS or protonated A38H<sup>+</sup> transfers the proton to the leaving G+1(O5') group (acting as general acid).<sup>17,24</sup> Our subsequent DFT QM/MM calculations revealed that A38H<sup>+</sup> can either act



**Figure 1.** (A) Scheme of the QM region for the phosphorane intermediate state in the monoanionic mechanism (containing the A-1 ribose ring, G+1 phosphate, canonical G8, and protonated A38H<sup>+</sup> species) used in the QM/MM computations. (B) Snapshot from QM/MM calculations (shown for the phosphorane intermediate state), where the MM region and QM core are rendered as wires and thicker sticks, respectively. Water molecules and sodium counterions are not shown for clarity.

as a general acid or be involved in electrostatic stabilization of the TS.<sup>18</sup> Canonical A38 also provided a feasible activation barrier but required a properly structured active site derived from simulations with a protonated A38H<sup>+</sup> form.<sup>18</sup>

Multiscale QM/MM calculations could, in principle, allow an ultimate description of the enzymatic chemistry at an atomic level of resolution.<sup>38–44</sup> However, the accuracy of such calculations heavily depends on the quality of the QM method used<sup>45–51</sup> and affordable sampling.<sup>52,53</sup> In general, these two factors have opposing effects: the better the quality of the QM method, the more limited is the sampling that can be achieved. Semiempirical methods are often used because of their relatively low computer cost and possibility of extended sampling. Some such methods have been carefully parametrized to provide reaction profiles in agreement with advanced quantum chemistry, e.g., the AM1/d-PhoT<sup>16</sup> and SCC-DFTBPR<sup>54,55</sup> methods. More expensive DFT methods have also been successfully applied to study (RNA) catalytic reactions.<sup>18,56–58</sup> In the past, gradient corrected (GGA) functionals (e.g., BLYP) were often favored because they allow a convenient balance of accuracy and computational cost. However, reaction barriers computed by BLYP are slightly underestimated due to the electron self-interaction error inherent to gradient corrected (GGA) functionals.<sup>48,59–61</sup> The self-interaction error can be reduced by using DFT functionals containing some portion of the exact Hartree–Fock electron exchange, e.g., hybrid functionals. Among them, the MPW1K functional has been optimized for reaction kinetics.<sup>62,63</sup> The activation barriers provided by such hybrid DFT functionals are in better agreement with experimental data.<sup>64</sup> Finally, post Hartree–Fock wave function theory methods (e.g., MP2, CCSD(T), etc.) can also be used, but their application is hampered by large computational costs.

Most QM/MM computations evaluate potential energy surfaces (PESs). In PES computations, corrections (involving zero point energy, enthalpy correction for finite temperature, and entropy contribution) need to be included to obtain activation barriers that are directly comparable to experimentally observed data. Another option is to sample free energy surfaces (FESs) along the reaction pathway by constructing an entire FES of the enzymatic reaction. However, FES calculations are prohibitively expensive for high quality QM methods. Combined QM/MM–MD methods using Hartree–Fock, DFT BLYP, and B3LYP approaches are also available for optimization of minimum free energy pathways.<sup>65–71</sup> In order to investigate more complex FESs of enzymatic reactions, only

empirical (e.g., EVB)<sup>72–74</sup> and semiempirical methods are affordable with the current computer equipment.<sup>44,48</sup>

In the present study, we analyzed PES and (when possible) FES of the self-cleavage reaction catalyzed by the hairpin ribozyme using various QM/MM approaches. For comparison of these approaches, we considered two reaction scenarios involving the canonical and deprotonated form of G8 and used various hybrid QM/MM methods differing in the QM core. We applied the advanced QM spin-component scaled Møller–Plesset (SCS-MP2) method as a source of reference QM data for the description of the PES. Comparison with this reference data showed that the semiempirical AM1/d-PhoT and SCC-DFTBPR Hamiltonians provided acceptable activation barriers compared to the SCS-MP2 level (agreeing within a few kilocalories per mole). However, this comparison is somewhat misleading because the semiempirical methods favored a different reaction pathway because of the significantly different shapes of the PESs in comparison with those calculated by the benchmark SCS-MP2 method. In contrast, the DFT BLYP and MPW1K functionals predicted (at the level of PES description) the same reaction pathway as the SCS-MP2 method. The hybrid MPW1K functional gave a similar activation barrier to the SCS-MP2 data, which agreed well with experiment, while the BLYP barrier was slightly underestimated with a value comparable to values obtained from semiempirical methods. Currently, construction of FESs for enzymatic reactions is affordable only with semiempirical methods. Our calculations revealed that shapes of the FESs are to a large extent determined by the PESs, i.e., the semiempirical FES and PES predict the same reaction pathway differing from the benchmark SCS-MP2 PES. The results thus suggest that the semiempirical methods are prone to introduce errors in deciphering the reaction mechanisms of RNA catalysis.

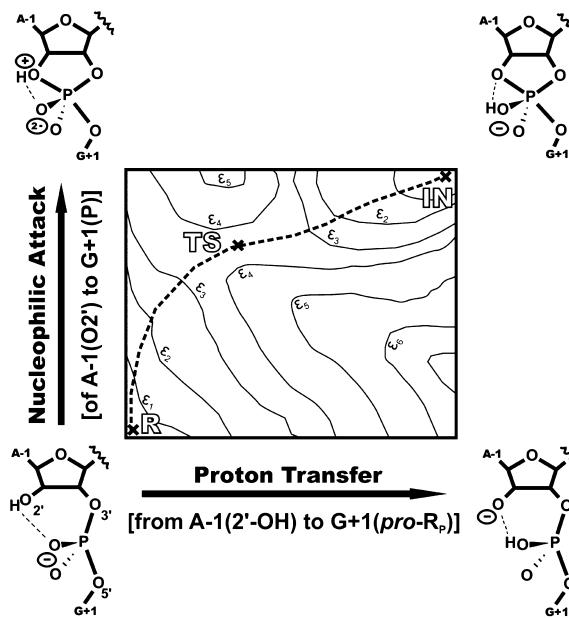
## METHODS

We used three different approaches for description of the QM region of QM/MM computations: (i) DFT, represented by BLYP<sup>75,76</sup> and MPW1K<sup>62,63</sup> functionals; (ii) correlated ab initio calculations, represented by the spin-component scaled MP2 (SCS-MP2) method;<sup>77</sup> and (iii) semiempirical methods, represented by semiempirical molecular orbital AM1/d-PhoT<sup>16</sup> and SCC-DFTBPR<sup>54,55</sup> Hamiltonians; the latter semiempirical methods have relatively low computational costs, and thus allow free energy calculations by umbrella sampling molecular dynamics (MD) simulations.

**Potential Energy Surface Calculations.** We used a hybrid QM/MM scheme<sup>38</sup> with electronic embedding, where QM energies are electronically polarized by MM point charges of the AMBER (Cornell et al.) molecular mechanical force field ff99.<sup>78,79</sup> Valences of the QM atoms at the QM/MM boundary were satisfied by the addition of link hydrogen atoms. The DFT approach was used with the QM/MM subtractive scheme by a two-layer ONIOM method<sup>80</sup> implemented in Gaussian 09,<sup>81</sup> where the MM region was treated by the Sander module of AMBER 12.0.<sup>82</sup> Geometry optimizations were carried out with the QM region described by the BLYP/6-31G(d,p) method. The PES was further refined by single point energies (using the BLYP-optimized geometries) with the MPW1K/6-31+G(d,p) method. Discussion of the applicability and performance of these methods for modeling RNA catalysis can be found elsewhere.<sup>48</sup> The ab initio approach was applied using the QM/MM additive scheme employed by the QoMMMa program,<sup>83</sup> where the MM region was treated by Tinker software.<sup>84</sup> The QM calculations were performed using the MOLPRO 2012 software package<sup>85</sup> at the SCS-MP2/aug-cc-pVDZ and SCS-MP2/aug-cc-pVTZ levels to estimate the SCS-MP2/complete basis set limit (CBS).<sup>86,87</sup> Energies were obtained as single points by again using ONIOM BLYP-optimized geometries. Note that although the SCS-MP2 method generally provides rather similar activation barriers to the standard MP2 method, SCS-MP2 activation barriers are often in slightly better agreement with the reference data.<sup>88–90</sup> In addition, the SCS-MP2 method describes sufficiently well dispersion interactions.<sup>91,92</sup> On the other hand, the dispersion interactions are neglected by standard DFT functionals (here both BLYP and MPW1K) but a posteriori correction schemes might improve DFT accuracy for enzyme catalyzed reactions.<sup>93</sup> The semiempirical approach was considered using the QM/MM additive scheme with the Sander module of AMBER 12.0<sup>82</sup> and QM/MM PME implementation.<sup>94,95</sup> The QM region was described by two semiempirical Hamiltonians: (i) a combined AM1/d and MNDO/d model with new parameters for H, O, and P atoms (AM1/d-PhoT<sup>16</sup>), and (ii) a self-consistent charge version of a DFT-based tight binding Hamiltonian (SCC-DFTB)<sup>54</sup> parametrized for phosphate reactions (SCC-DFTBPR).<sup>55</sup> The semiempirical PESs were calculated over (i) the BLYP-optimized geometries as well as (ii) fully optimized semiempirical structures.

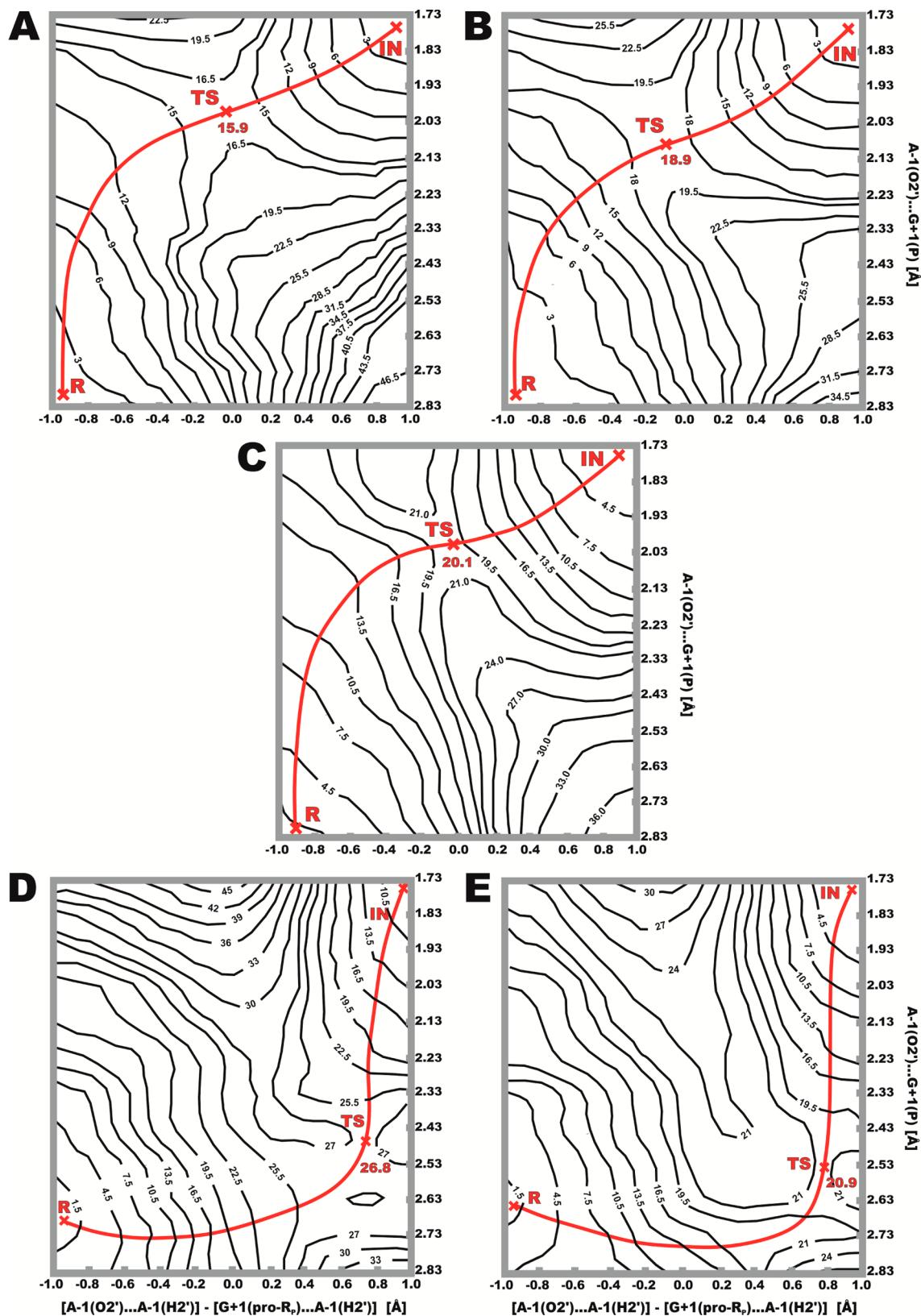
**Free Energy Surface Calculations.** The QM/MM approach with both semiempirical Hamiltonians was subsequently used in combination with classical explicit solvent MD simulations. The semiempirical QM/MM–MD protocol used 1 fs integration time steps. A series of MD simulations at constant pressure with isotropic ( $T = 300$  K) position scaling were carried out to sample all points across the FES surface with 50-ps-long dynamics for each window. Harmonic restrained potentials were applied with force constants reaching 200 kcal/mol for each geometry point (key distances differing by 0.1 Å) in the two-dimensional (2-D) diagram. Outputs (as differences in restrained distances) were subsequently analyzed by the weighted histogram analysis method (WHAM)<sup>96</sup> in order to obtain a 2-D semiempirical QM/MM–MD FES.

**Selection of Geometries and Definition of the QM Region.** The QM region (49 atoms) comprised two nucleobases (G8 and A38) and the A-1 ribose ring with the G+1 scissile phosphate (Figure 1). In the QM/MM DFT and ab initio calculations, the ribozyme was immersed in a water droplet with an ~10 Å thick layer of water molecules

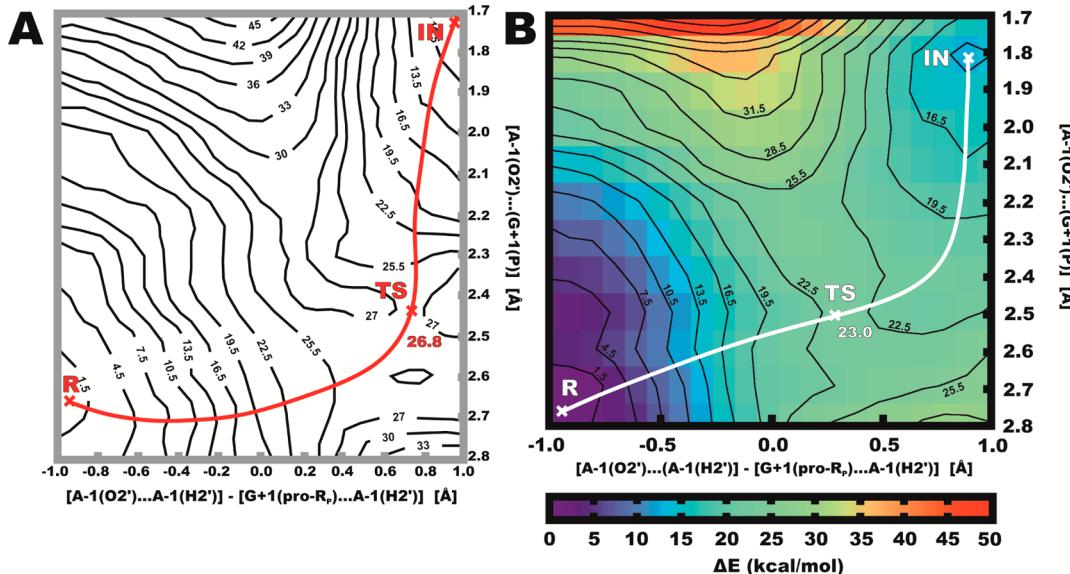


**Figure 2.** Scheme of 2-D projection, known as a More O'Ferrall–Jencks plot.<sup>99</sup> Each axis represents steps of the respective reaction, i.e., nucleophilic attack and nucleophile activation by proton transfer in the hairpin ribozyme (shown as the simplest models). Energy levels are displayed as contours computed for particular structures at defined intervals on the grid. The most convenient path (requiring the lowest energies) from the initial state (reactant, R) across the saddle point (transition state, TS) to the final state (intermediate, IN, or product, P) is called the reaction coordinate (dashed line). Note that the displayed 2-D profile is illustrative.

surrounding the RNA molecule. An ~5 Å thick layer of waters on the surface of the droplet and counterions outside the droplet were fixed during all QM/MM calculations to prevent any changes in energy caused by H-bond network reorganization at the water–vacuum interface.<sup>18,56</sup> The whole system contained ~13 000 atoms, of which ~6900 were fixed. As initial structures for the QM/MM calculations (DFT and ab initio approaches), we used BLYP/6-31G(d,p) optimized geometries from our previous QM/MM study on the hairpin ribozyme containing a slightly larger QM region (64 atoms).<sup>18</sup> Initial geometries for the QM/MM computations with semiempirical methods were prepared with the aid of MD simulations (monoanionic mechanism, Scheme 1A) or using the BLYP/6-31G(d,p) optimized geometry of the R state determined in our previous QM/MM calculations<sup>18</sup> as the initial snapshot (dianionic mechanism, Scheme 1B). MD simulations were carried out using the AMBER 12.0 program<sup>82</sup> with the ff99bsc0XOL3 force field (Cornell et al. force field ff99 with recently implemented refined backbone torsion parameters),<sup>97,98</sup> modeling immersed structures in a rectangular water box with an ~10 Å thick layer of TIP3P water molecules around the RNA solute (~30 000 atoms overall). The starting snapshots for the semiempirical QM/MM computations were chosen based on geometries achieving the best orientations of key residues (mainly those included in the QM region) in the active site, where key H-bond distances (G8(N1H)…G+1(pro-S<sub>P</sub>), G8(N2H)…G+1(pro-S<sub>P</sub>), A38H<sup>+</sup>(N1H)…G+1(OS'), and A38H<sup>+</sup>(N6H)…G+1(pro-R<sub>P</sub>)) were below 3.0 Å with an in-line attack angle (A-1(O2')…G+1(P)…G+1(OS')) greater than 160°.



**Figure 3.** Two-dimensional QM/MM maps of the PES for the monoanionic mechanism obtained using (A) BLYP, (B) MPW1K, (C) SCS-MP2, (D) AM1/d-PhoT, and (E) SCC-DFTBPR methods. Energies (kcal/mol) were calculated as single points on BLYP/6-31G(d,p)/AMBER(ff99) geometries (see Methods). Red lines indicate reaction paths with R, TS, and IN states indicated.



**Figure 4.** QM/MM semiempirical PES for the monoanionic mechanism computed by the AM1/d-PhoT method. (A) AM1/d-PhoT single points over BLYP/6-31G(d,p) geometries (same as Figure 3D) and (B) AM1/d-PhoT fully optimized PES. Red and white lines indicate reaction paths with R, TS, and IN states marked.

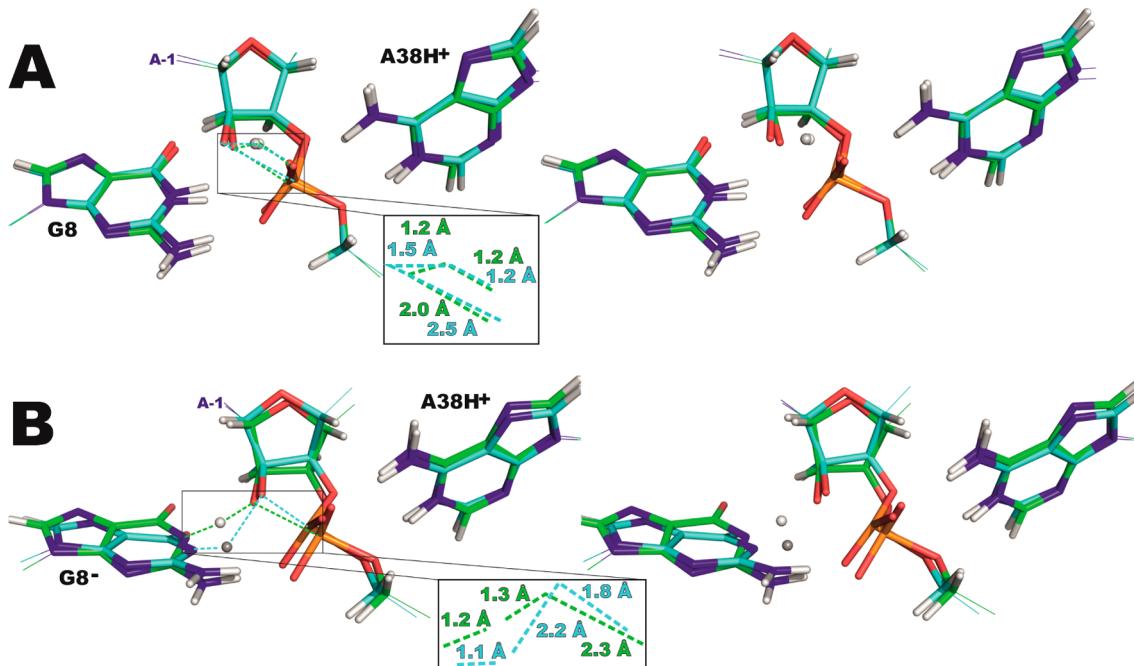
## RESULTS

We used diverse QM/MM methods and constructed multiple More O’Ferrall–Jencks diagrams (Figure 2) to represent the PES and FES of the self-cleavage reaction pathway of the hairpin ribozyme. We explored two possible reaction scenarios (Scheme 1) that were considered to be the most plausible based on the available mechanistic and computational data. The first system contained canonical guanine 8 (G8) and adenine 38 (A38H<sup>+</sup>) in the active site, and the second system contained deprotonated G8 (G8<sup>-</sup>) and A38H<sup>+</sup> nucleobases. We primarily focused on the first step of the self-cleavage reaction, i.e., nucleophilic attack of the A-1(2'-OH) group to the scissile phosphate with simultaneous proton transfer from A-1(2'-OH) to either one of the nonbridging oxygens of the scissile phosphate or to G8<sup>-</sup>(N1) nitrogen. This first step was suggested to be rate limiting in our preceding PES QM/MM calculations.<sup>18</sup> The computed More O’Ferrall–Jencks diagrams mapped energy changes along two tracked distance-based parameters. For the monoanionic “shuttle” mechanism, the first parameter was the difference between the A-1(O2')…A-1(H2') and G+1(pro-R<sub>p</sub>)…A-1(H2') distances, corresponding to proton transfer from the A-1(2'-OH) group to the G+1(pro-R<sub>p</sub>) nonbridging oxygen of the scissile phosphate. In the reaction scenario with G8<sup>-</sup>, the first parameter was the difference between the A-1(O2')…A-1(H2') and G8<sup>-</sup>(N1)…A-1(H2') distances, corresponding to proton transfer from A-1(2'-OH) to G8<sup>-</sup>(N1). The second tracked parameter, i.e., the distance between A-1(O2')…G+1(P), was identical in both scenarios and represented the nucleophilic attack step.

**Monoanionic Reaction Mechanism. QM/MM PES Calculations Using the DFT and ab Initio Methods Reveal Transition States with Concerted Nucleophilic Attack and Proton Transfer Steps.** The position of the TS in the PES (Figure 3A) calculated by BLYP/6-31G(d,p) indicated that the nucleophilic attack (of A-1(O2') oxygen to the scissile phosphate) and proton transfer (from A-1(2'-OH) group to G+1(pro-R<sub>p</sub>) nonbridging phosphate oxygen) are concerted events because the TS geometry corresponded to an

A-1(O2')…G+1(P) distance of ~2.0 Å and the proton was located in the middle between the A-1(O2') and G+1(pro-R<sub>p</sub>) atoms. The difference between the A-1(O2')…A-1(H2') and G+1(pro-R<sub>p</sub>)…A-1(H2') distances was roughly zero. The computed BLYP/6-31G(d,p) activation barrier was 15.9 kcal/mol. BLYP (and generally all GGA DFT functionals) is known often to underestimate activation barriers due to the electron self-interaction error.<sup>48,61</sup> Therefore, we recalculated the energy profile using the more accurate hybrid DFT MPW1K/6-31+G(d,p) (Figure 3B) and ab initio SCS-MP2/CBS methods (Figure 3C), which yielded barriers of 18.9 and 20.1 kcal/mol, respectively. We note that the higher MPW1K barrier (with respect to the BLYP barrier) is almost equally caused by usage of hybrid functional and the diffuse basis set. To compare the barriers with experimental data, additional correction for the rare occurrence of protonated A38H<sup>+</sup> form at the active site has to be included (discussed later). Both 2-D profiles showed that the PES resembled the shape of the BLYP PES and both DFT and the ab initio methods yielded the same concerted reaction pathway.

**QM/MM Profiles with a Semiempirical Scheme Constructed as Single Points Using DFT-Optimized Geometries Favor a Sequential Reaction Pathway.** The semiempirical AM1/d-PhoT and SCC-DFTBPR single point recalculations using the QM/MM BLYP geometries provided PES profiles that were significantly different from those calculated by DFT and ab initio methods. The semiempirical Hamiltonians have been parametrized for reaction with phosphates<sup>16,55</sup> and extensively used to decipher reaction mechanisms catalyzed by proteins and RNA enzymes.<sup>17,24,100–110</sup> The AM1/d-PhoT method has previously been used in QM/MM studies of ribozyme catalysis.<sup>17,24,104</sup> Both the AM1/d-PhoT and SCC-DFTBPR PESs favor a sequential reaction pathway, in which proton transfer precedes the nucleophilic attack step (Figure 3D,E). Proton transfer is realized with the A-1(O2')…G+1(P) distance greater than ~2.6 Å and is well separated from the nucleophilic attack step. The calculated barriers were 20.9 kcal/mol (SCC-DFTBPR) and 26.8 kcal/mol (AM1/d-PhoT). The AM1/d-PhoT barrier was significantly higher than those



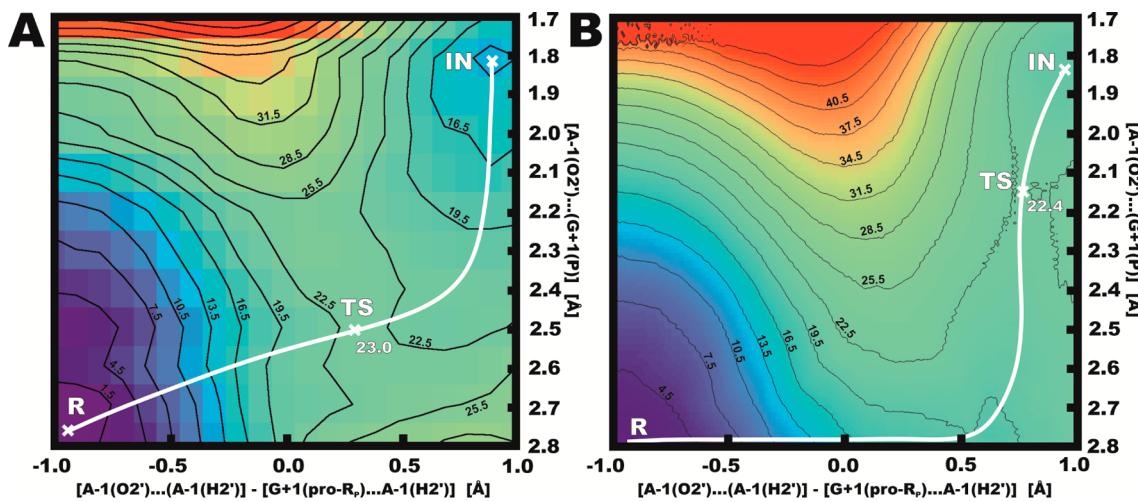
**Figure 5.** Stereo view of snapshots of the active site in the TS geometries for the (A) monoanionic and (B) dianionic reaction mechanisms. Green sticks represent the TS taken from the QM/MM BLYP/6-31G(d,p) calculations, whereas cyan sticks correspond to the TS obtained with the QM/MM semiempirical AM1/d-PhoT computations. Key distances involving proton transfer and nucleophilic attack steps are highlighted in the insets (green labels and cyan labels correspond to BLYP and AM1/d-PhoT structures, respectively).

calculated by DFT and ab initio methods (18.9–20.1 kcal/mol) and was also different from the AM1/d-PhoT barrier reported in the original work by Nam and co-workers (20.5 kcal/mol).<sup>17</sup>

**Semiempirical PES Obtained by Full Optimization Is Consistent with PES Obtained as Single Points Using DFT-Optimized Geometries.** As noted above, our AM1/d-PhoT barrier differed from earlier AM1/d-PhoT computations. This indicates that the results may also be affected by geometries. Thus, for both semiempirical methods, we obtained the PES using fully optimized geometries in order to verify overall barriers and positions of the TS on the 2-D diagrams. The reaction pathways localized on semiempirical PES surfaces again corresponded to the sequential pathway. Thus, the results confirmed that semiempirical methods prefer the scenario in which the proton transfer from the A-1(2'-OH) group to the G+1(*pro-R<sub>P</sub>*) nonbridging oxygen of the scissile phosphate precedes nucleophilic attack of A-1(O2') to G+1(P) (Figure 4B and Figures S1 and S2 in the Supporting Information). The TSs on the fully optimized PESs were localized in the same region as those predicted with semiempirical single point calculations on DFT-optimized geometries (Figure 4 and Figure S1 in the Supporting Information) and corresponded to proton transfers rather than nucleophilic attack steps. The TS geometries were characterized by an A-1(O2')...G+1(P) distance of ~2.5 Å (AM1/d-PhoT) or ~2.7 Å (SCC-DFTBPR) and the difference between A-1(O2')...A-1(H2') and G+1(*pro-R<sub>P</sub>*)...A-1(H2') distances of ~0.3 Å (AM1/d-PhoT) or ~0.2 Å (SCC-DFTBPR). A detailed comparison of the TS geometries corresponding to the sequential reaction pathway (predicted by semiempirical methods) and the concerted pathway (DFT and ab initio methods) is shown in Figure 5. The overall barriers were 22.1 and 23.0 kcal/mol for the SCC-DFTBPR and AM1/d-PhoT methods, respectively. Both barriers were still slightly higher than those calculated by QM/MM DFT and ab initio methods (18.9–20.1 kcal/mol). In contrast to single point

calculations on DFT-based geometries, the AM1/d-PhoT results are more consistent with the earlier study by Nam and co-workers.<sup>17</sup>

**Reaction Pathways on Semiempirical FESs Also Display Separated Proton Transfer and Nucleophilic Attack Steps.** For the final comparisons, we computed QM/MM semiempirical FES using restrained MD simulations, followed by subsequent analysis by WHAM (see Methods). Both the AM1/d-PhoT and SCC-DFTBPR FES calculations again showed that proton transfer from the A-1(2'-OH) group to G+1(*pro-R<sub>P</sub>*) precedes the nucleophilic attack step. The FES calculations thus provided comparable scenarios with the same sequential reaction pathway as predicted by the equivalent PES calculations (Figure 6 and Figure S3 in the Supporting Information). In other words, at least in this particular case, the FES profiles were primarily determined by the PES. The free energy corrections derived from the biased MD simulations on the semiempirical PES contributed by less than 3 kcal/mol (Figures S2 and S3 in the Supporting Information). The small free energy corrections were rather consistent with previously published free energy corrections estimated from the harmonic approximation on model reactions.<sup>18,48,56,90,111–114</sup> Obviously, we cannot guarantee that our sampling was sufficiently representative. However, we compared results obtained using 5 and 50 ps windows and found they were almost identical. This indicates good convergence of the results. Nevertheless, we still cannot rule out that a lack of convergence might become apparent on longer time scales, which are not achievable at present. Despite the rather good agreement of the reaction pathways on the PES and FES, the TS in the FES generated by AM1/d-PhoT was considerably closer (geometrically and energetically) to the position of the reaction intermediate (IN, top-right corner of the diagram in Figure 6B) with a free energy barrier of 22.4 kcal/mol. The AM1/d-PhoT TS was connected only with the nucleophilic attack step,



**Figure 6.** QM/MM AM1/d-PhoT (A) PES and (B) FES for the monoanionic mechanism. White lines indicate reaction paths with R, TS, and IN states marked. The color scale (energies in kcal/mol) is the same as in Figure 4. See Figures S2 and S3 in the Supporting Information with extended ranges of both coordinates.

in contrast to the PES profile. Notably, the position of the TS in the SCC-DFTBPR FES was also closer to the IN state (the  $A-1(O_2')\cdots G+1(P)$  distance of  $\sim 2.1 \text{ \AA}$ ), with an overall barrier of 16.7 kcal/mol (Figure S3B in the Supporting Information). The bottom-right-corner area on the FES diagram, where the proton is already located on  $G+1(\text{pro-}R_p)$  nonbridging oxygen with deprotonated  $A-1(2'\text{-O}^-)$  group, reveals a surprisingly stable state with significantly lower energies than in the corresponding SCC-DFTBPR PES profile (Figure S2B in the Supporting Information). In summary, both semiempirical FES profiles predicted reaction scenarios with the sequential pathway. This prediction was consistent with the semiempirical PES prediction description but differed significantly from the higher-level QM/MM DFT and ab initio PES calculations.

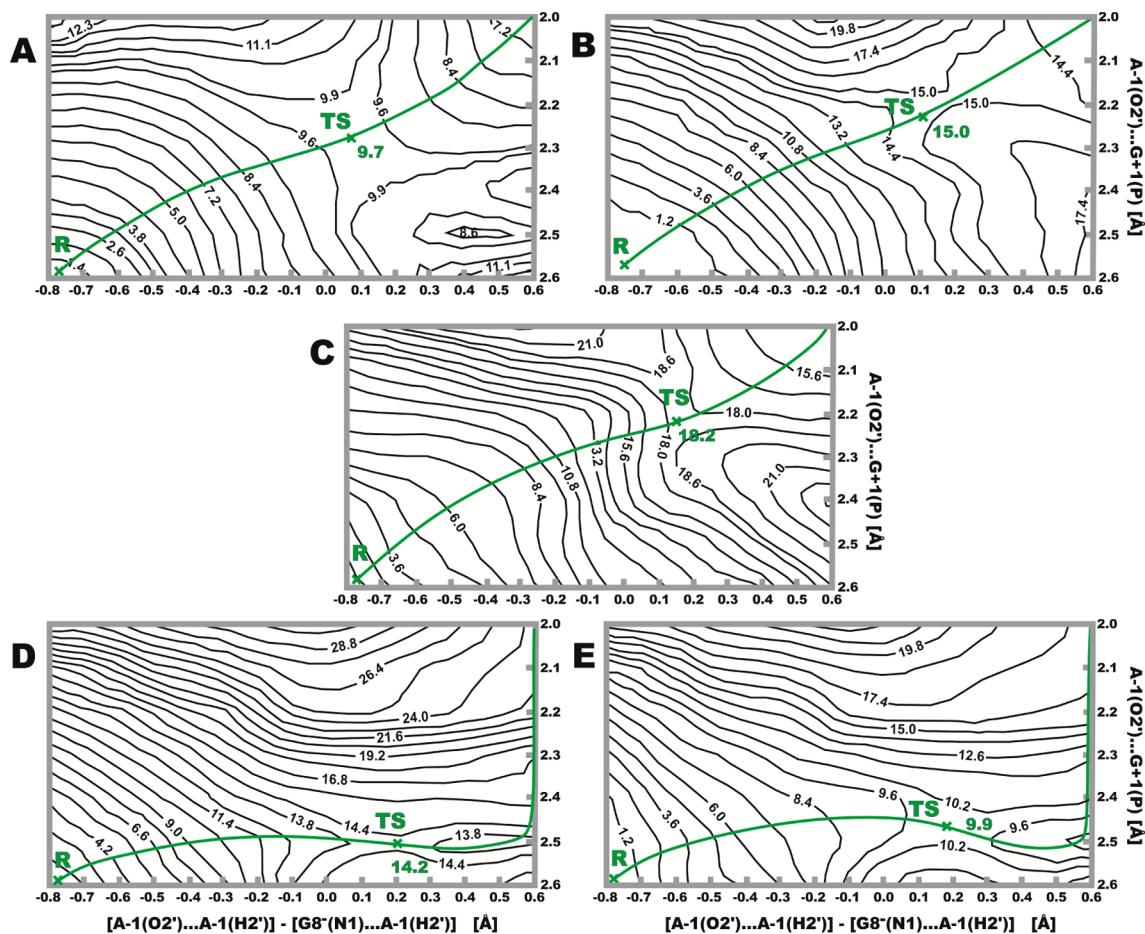
**Dianionic Reaction Mechanism. DFT and ab Initio Approaches Predict a TS with Concerted Nucleophilic Attack and Proton Transfer Steps for the Dianionic Reaction Mechanism.** The QM/MM BLYP/6-31G(d,p) PES profile revealed a single TS. The distance between the  $A-1(O_2')$  and  $G+1(P)$  atoms was  $\sim 2.3 \text{ \AA}$  and the proton from the  $A-1(2'\text{OH})$  group was localized in the middle between the  $A-1(O_2')$  and  $G8^-(N1)$  atoms. The difference between the  $A-1(O_2')\cdots A-1(H_2')$  and  $G8^-(N1)\cdots A-1(H_2')$  distances was  $\sim 0.1 \text{ \AA}$  (Figure 5B). The activation barrier was 9.7 kcal/mol (Figure 7A). The PES calculated by the BLYP method did not reveal any IN state, resulting in a barrierless profile between the TS and P states (data not shown). Subsequently, we recalculated the single points by MPW1K/6-31+G(d,p) and SCS-MP2/CBS methods. The barrier increased to 15.0 kcal/mol (MPW1K functional) and 18.2 kcal/mol (SCS-MP2 method), whereas the PES shape remained similar to that at the BLYP level (Figure 7B,C). It should be noted that, in order to compare with experiments, the overall barriers have to be corrected for the rare occurrence of both deprotonated  $G8^-$  and protonated  $A38H^+$  forms at the active site (discussed later).

**QM/MM Semiempirical Profiles Constructed as Single Points Using DFT Geometries Predict a Sequential Reaction Pathway for the Dianionic Mechanism.** Both semiempirical AM1/d-PhoT and SCC-DFTBPR Hamiltonians provided PESs that supported a sequential pathway. Proton transfer was accomplished immediately (with the  $A-1(O_2')\cdots G+1(P)$  distance larger than  $\sim 2.4 \text{ \AA}$ ) and was separated from the

subsequent nucleophilic attack steps (Figure 7D,E). We observed energetically stable IN states after the proton from the  $A-1(2'\text{-OH})$  group was transferred to the  $G8^-(N1)$  nitrogen. The corresponding barriers were 14.2 kcal/mol (AM1/d-PhoT) and 9.9 kcal/mol (SCC-DFTBPR).

**QM/MM Semiempirical Fully Optimized PES Also Reveals a Sequential Reaction Pathway for the Dianionic Mechanism with Rate-Limiting Steps Shifted toward an Exocyclic Cleavage Step.** The above QM/MM semiempirical profiles indicate that additional TSs (with the rate-limiting step) may be localized outside the region included in the grid used for the QM/MM BLYP geometry optimizations (Figure 7). Therefore, the semiempirical methods were used to extend the grid to include the complete PES reaching up to the P state. The extended AM1/d-PhoT relaxed PES predicted a sequential pathway, in which proton transfer from the  $A-1(2'\text{-OH})$  group to  $G8^-(N1)$  (the first step of the reaction) was separated from the nucleophilic attack step. The proton transfer occurred immediately (distance between  $A-1(O_2')$  and  $G+1(P)$  is  $\sim 2.8 \text{ \AA}$ , Figure 8A) and led to a stable IN state with a deprotonated  $A-1(2'\text{-O}^-)$  group and canonical  $G8$  form. We scanned the whole PES up to the P state in order to identify the TS with the overall barrier. Indeed, the TS was localized near the middle of the diagram's left edge (Figure 8A), indicating that it was associated with the nucleophilic attack, exocyclic cleavage step, and with simultaneous second proton transfer from  $A38H^+$  to  $G+1(O_5')$ . This behavior differed from the QM/MM profiles obtained by the DFT and ab initio methods, where the single TS (the rate-limiting step) was characterized by coordinated nucleophilic attack and the first proton transfer step from  $A-1(2'\text{-OH})$  to  $G8^-(N1)$ . The calculated overall AM1/d-PhoT barrier (32.6 kcal/mol) appeared to be very high (Table 1). This indicates that the sampling may have been insufficient. Thus, we carried out additional scans, combining nucleophilic and exocyclic cleavage steps with proton transfer from  $A38H^+$  to  $G+1(O_5')$ , but no significant decrease in energies along the reaction pathway was achieved.

The optimized PES obtained by SCC-DFTBPR indicated a similar profile with a sequential reaction pathway as described for the AM1/d-PhoT method. However, we observed spurious bond formation between the  $G+1(\text{pro-}S_p)$  oxygen and one  $\text{Na}^+$



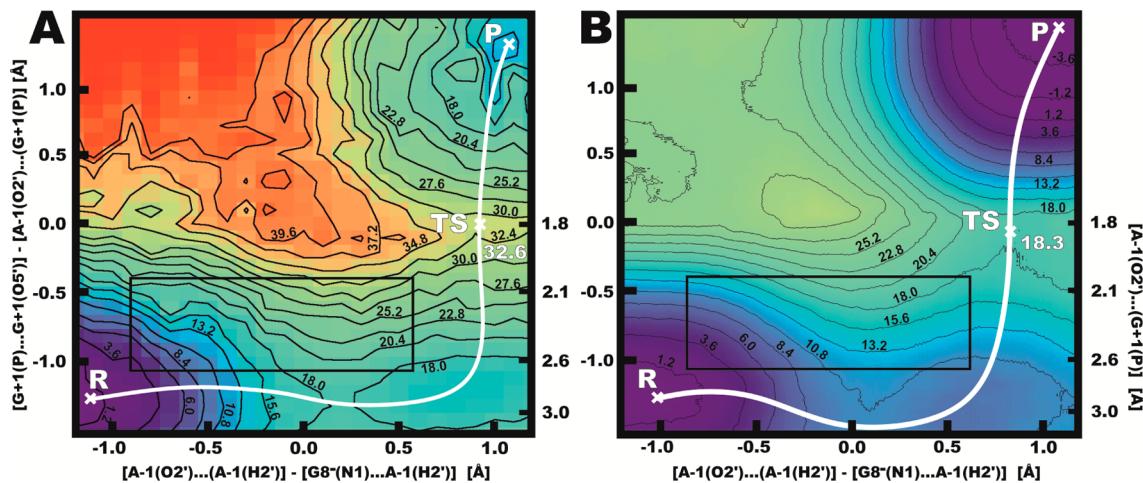
**Figure 7.** Two-dimensional maps of the PES for the dianionic mechanism obtained using (A) BLYP, (B) MPW1K, (C) SCS-MP2, (D) AM1/d-PhoT, and (E) SCC-DFTBPR QM/MM methods. Energies (kcal/mol) are calculated as single points on BLYP/6-31G(d,p)/AMBER(ff99) geometries (see Methods). Green lines indicate reaction paths with R and TS states marked.

counterion prior to the exocyclic cleavage, which substantially distorted the PES (data not shown).

**AM1/d-PhoT FES Calculation of the Dianionic Reaction Mechanism Confirms Separated Proton Transfer and Nucleophilic Attack Steps.** We also constructed a semi-empirical QM/MM FES profile for the dianionic mechanism, using restrained MD simulations followed by WHAM analysis. The AM1/d-PhoT FES showed that proton transfer from the A-1(2'-OH) group to the G8<sup>-</sup>(N1) nitrogen proceeds immediately (A-1(O2')...G8<sup>-</sup>(N1) distance of ~3.0 Å) and generates a stable IN state with canonical G8 and deprotonated A-1(2'-O<sup>-</sup>) group (Figure 8B). The second TS corresponding to the overall barrier was localized near the left middle edge of the diagram and was connected with both nucleophilic attack and exocyclic cleavage steps. The overall shape of the FES was consistent with the reaction pathway identified in the AM1/d-PhoT PES (Figure 8). However, the activation barrier in the FES (18.3 kcal/mol) was considerably lower than the PES barrier (32.6 kcal/mol). We inspected geometries generated in the course of 50-ps-long MD simulations by analyzing all umbrella sampling windows in terms of maximal root-mean-square deviation from the starting structure (corresponding to PES conformation). Constructed 2-D maps reveal in principle homogeneous profiles without any structural deviations (Figure S4 in the Supporting Information). The QM/MM AM1/d-PhoT FES calculation thus indicates that the free energy correction (~14 kcal/mol, Figure 8) plays a very important role

in this particular case. This result contrasts with the free energy correction suggested by previous QM calculations on a model reaction in the absence of the ribozyme in water (reference reaction), which was around 1 kcal/mol.<sup>18,56,111,112,114</sup> It is not straightforward to explain this large difference, which can be caused by several factors. On one hand, the pure QM calculations of the reference reaction may be limited by the use of small model systems, with the free energy correction derived within the ideal gas, rigid rotor, and harmonic oscillator approximation. On the other hand, we cannot rule out (although we do not have any specific evidence) that the QM/MM MD simulations may be affected by the rather short sampling times used for each window (50 ps in our simulations). This could indicate that the sampling of the QM/MM MD simulations was not fully convergent. In any case, the close agreement between the PES and FES barriers reported above for the monoanionic mechanism may be coincidental. Thus, further work is needed to clarify the relation between PES and FES descriptions in studies of RNA catalysis. Still, when considering all available data, the overall agreement between the PES and FES is encouraging. It implies that the reaction can be usefully studied by calculating the PES, justifying the higher-level QM/MM calculations used here and in previous studies.

It should be noted that we were not able to calculate the FES profile using the SCC-DFTBPR method owing to instabilities during the MD simulations, most likely caused by electrostatic



**Figure 8.** QM/MM fully optimized (A) PES and (B) FES for the dianionic mechanism computed by the AM1/d-PhoT method. White lines indicate the reaction paths with R, TS, and P states marked. The color scale (showing energies in kcal/mol) is the same as in Figure 4. The black boxes within the diagrams correspond to the 2-D maps captured in Figure 7.

interaction of the scissile phosphate with  $\text{Na}^+$  counterions that enter the active site. Thus, one of the nonbridging oxygen atoms detached from the scissile phosphate in the simulations. This underscores some of the challenges of semiempirical description of RNA catalysis.

**Differences in Estimated Reaction Pathways between Semiempirical and Higher-Level Methods Are Apparent Also in QM Gas Phase Calculations.** In order to decipher the origin of the differences in reaction paths derived from semiempirical and higher-level methods, we calculated 2-D PES profiles using QM calculations on the hairpin ribozyme active site model in the gas phase. We used a small model system that matched the QM region used in our QM/MM calculations (53 atoms in total with added hydrogen link atoms). We performed gas phase optimizations using the AM1/d-PhoT method and constructed the PES of the dianionic mechanism. The heavy atoms bonded to hydrogen link atoms ( $\text{A-1(C1')}$ ,  $\text{A-1(C4')}$ ,  $\text{G+1(C5')}$ ,  $\text{G8}^-(\text{N9})$ , and  $\text{A38H}^+(\text{N9})$ ) were constrained in order to retain the conformation resembling the hairpin ribozyme active site conformation. In addition, we constrained  $\text{A38H}^+(\text{N1-H1})$ ,  $\text{A38H}^+(\text{N6-H6})$ , and  $\text{G8}^-(\text{N2-H2})$  bonds to avoid spurious proton transfers which, e.g., may lead to adenine N6-imino rare tautomer (no such tautomers were observed in similar DFT calculations). Subsequently, single point calculations on these AM1/d-PhoT geometries were performed by the MPW1K/6-31+G(d,p) method. As expected, the QM and QM/MM PES profiles differed significantly. We observed differences between AM1/d-PhoT and MPW1K QM PES profiles similar to those in the full QM/MM computations; i.e., the AM1/d-PhoT and MPW1K methods predicted sequential and concerted reaction pathways, respectively (Figure S7 in the Supporting Information). Thus, the profound differences between semiempirical and higher-level QM methods discussed in our study originate primarily in QM Hamiltonians rather than in some specific aspects of the QM/MM methodology, such as QM/MM coupling.

## DISCUSSION

We have analyzed the PES and FES of the RNA backbone self-cleavage reaction catalyzed by the hairpin ribozyme using various QM/MM methods. The main aim of our study was a side-by-side comparison of the performance of different QM/

MM methods. In particular, we aimed to compare methods based on a less rigorous treatment of the QM region (which allows more conformational sampling) with higher-level QM/MM methods. Three types of methods, based on fundamentally different concepts of quantum chemistry, were used for the investigation of activation barriers, reaction pathways, and overall shapes of energy surfaces: (i) DFT, (ii) ab initio calculations with inclusion of electron correlation, and (iii) parametrized semiempirical methods based on either molecular orbital or approximate DFT methods. Using the more reliable but more expensive ab initio and DFT approaches, we calculated 2-D PESs of the reactions, constructed by using geometry optimizations. With the more approximate semiempirical treatments, we were also able to obtain FESs using QM/MM-MD simulations.

We considered two reaction mechanisms, labeled as the monoanionic and dianionic scenarios (Scheme 1), which have been shown to be the mechanisms most compatible with the available biochemical data and previous QM/MM calculations.<sup>17,18,24</sup> In the case of the monoanionic mechanism, the hairpin ribozyme active site contains canonical G8 and protonated  $\text{A38H}^+$ . In the dianionic mechanism, deprotonated  $\text{G8}^-$  and protonated  $\text{A38H}^+$  are involved.

**Corrections for Rare Protonation States.** As G8 and A38 have  $pK_a$ 's of 9.5<sup>25</sup> and 5.5,<sup>20</sup> respectively, the energy profiles calculated for the considered active site protonation states have to be corrected for the presence of minor protonation states at pH 7 as follows. If  $\text{G8}^-$  is present, the energy is shifted by 3.4 kcal/mol, if  $\text{A38H}^+$  is present, the energy is shifted by 2.1 kcal/mol, and if both  $\text{G8}^-$  and  $\text{A38H}^+$  are involved, the energy is shifted by 5.5 kcal/mol.<sup>18</sup> Such energy corrections have to be taken into account for comparison of calculated activation barriers with the experimental value of 20–21 kcal/mol derived from kinetic measurements.<sup>3</sup>

Note that, to compare the calculated potential energy barriers with experimental data, the free energy correction (to give the activation free energy) should be considered as well. The simplest approach is based on estimates using a reference reaction (calculations on a small model system). The free energy corrections were estimated to be negligible, i.e., 0.0 and 0.1 kcal/mol for model reactions of the monoanionic and

dianionic mechanisms, respectively.<sup>18</sup> Obviously, the free energy barriers obtained in these small-model computations may not necessarily be representative of the ribozyme-catalyzed reaction. The other option is to calculate the FES.

**Comparison of Activation Barriers Obtained by Different Methods.** We first obtained PESs for the two reaction pathways using the BLYP DFT functional. The BLYP method is (relatively) fast, and it is expected to be sufficiently accurate to obtain realistic geometries along the reaction coordinates. However, our results (in line with the literature data for other reactions)<sup>61,115</sup> indicate that the BLYP energy barriers might be underestimated due to the electron self-interaction error inherent to all gradient-corrected (GGA) functionals. The self-interaction error is reduced with DFT functionals containing some part of the exact Hartree–Fock exchange, as in the case of hybrid DFT methods. Such DFT functionals have been shown to be in better agreement with experimental reaction barrier data.<sup>64</sup> Thus, we recalculated (using the BLYP geometries) the PES using single point calculations with the hybrid MPW1K DFT functional, which was parametrized to provide reliable estimates of reaction barriers. For the hairpin ribozyme, MPW1K significantly increased the barriers compared to BLYP (which is almost equally caused by usage of the hybrid functional and the diffuse basis set) and the data became consistent with the experimentally delineated range for both the monoanionic and dianionic reaction scenarios.

The ab initio approach was represented by the SCS-MP2 method, utilized with complete basis set extrapolation. We consider this method as the most reliable in our set of computations, and likely the most accurate QM level that has been used so far in QM/MM studies of RNA catalysis. The SCS-MP2 single point energies were again derived using geometries obtained by the BLYP QM/MM approach. The SCS-MP2 activation barriers were 22.2 and 23.6 kcal/mol for the monoanionic and dianionic mechanisms, respectively (Table 1). These values are slightly higher than the experimentally derived barriers and are quite consistent with the MPW1K data (Table 1).

Semiempirical methods are usually considerably faster than DFT and ab initio methods but also less rigorous. In the present work, we used two semiempirical Hamiltonians specifically parametrized for reactions with phosphates: AM1/d-PhoT and SCC-DFTBPR. From FES simulations, the AM1/d-PhoT method (carefully parametrized to capture energy changes along the phosphodiester cleavage pathway) provided slightly higher free energy barriers than experiment, i.e., 24.5 and 23.8 kcal/mol for the monoanionic and dianionic mechanisms, respectively. Note that, from the PESs, there was a visible difference between the overall barriers calculated using the semiempirical and higher-level methods for the dianionic mechanism, as discussed in detail under Results.

The computational efficiency of the AM1/d-PhoT method allowed us to compare PES and FES descriptions of the reaction pathways. Comparison of the potential and free energy barriers showed that the difference, i.e., free energy correction, was typically less than 3 kcal/mol along the monoanionic mechanism and up to 14 kcal/mol for the dianionic mechanism. This finding may indicate a significant effect of thermal and entropy contributions, at least for the exocyclic cleavage part of the dianionic reaction mechanism (which was not studied for the monoanionic reaction scheme). Whereas the activation entropy contribution of 3 kcal/mol for the

**Table 1. QM/MM Activation Barriers (kcal/mol) and Corrections Required for Relevant Comparison to the Experimentally Measured Barrier (20–21 kcal/mol)**

	$\Delta E^\ddagger$ <sup>a</sup>	$\Delta G^\ddagger$ <sup>b</sup>	$\Delta G^\ddagger$ <sup>c</sup>	$E_{\text{cor}}$ <sup>d</sup>	$\Delta E^\ddagger$ <sup>e</sup>	$\Delta G^\ddagger$ <sup>f</sup>
Monoanionic Mechanism						
BLYP	15.9	15.9	—	2.1	18.0	18.0
MPW1K	18.9	18.9	—	2.1	21.0	21.0
SCS-MP2	20.1	20.1	—	2.1	22.2	22.2
AM1/d-PhoT	23.0	23.0	22.4	2.1	25.1	24.5/25.1
SCC-DFTBPR	22.1	22.1	16.7	2.1	24.2	18.8/24.2
Dianionic Mechanism						
BLYP	9.7	9.6	—	5.5	15.2	15.1
MPW1K	15.0	14.9	—	5.5	20.5	20.4
SCS-MP2	18.2	18.1	—	5.5	23.7	23.6
AM1/d-PhoT	32.6	32.5	18.3	5.5	38.1	23.8/38.0
SCC-DFTBPR <sup>g</sup>	—	—	—	5.5	—	—

<sup>a</sup>Directly computed uncorrected energy barriers from QM/MM PES.

<sup>b</sup>Free energy corrections (including zero point energies, enthalpy correction to finite temperature, and entropies) extrapolated from calculations of the reference reaction using small model system<sup>18</sup> are added to QM/MM energy barriers (0.0 and −0.1 kcal/mol for monoanionic and dianionic mechanisms, respectively). <sup>c</sup>Free energy barriers taken directly from QM/MM semiempirical FES calculations.

<sup>d</sup> $pK_a$  corrections for the presence of G8<sup>−</sup> and/or A38H<sup>+</sup> at pH ~7.

<sup>e</sup>Energies incorporating the respective  $pK_a$  correction. <sup>f</sup>Final  $pK_a$  corrected activation barriers calculated directly from the FES (the first number, where available) or estimated with extrapolated (reference reaction) free energy corrections (italics). <sup>g</sup>These calculations failed owing to accuracy limitations of the SCC-DFTBPR method (see the text).

monoanionic mechanism is within the typical range for activation entropy contributions of enzyme reactions, the contribution of 14 kcal/mol for the dianionic mechanism is outside it.<sup>116</sup> Further research is needed to understand the exact origin of such a large difference. As noted above, free energy corrections derived for model reactions in water in the absence of the ribozymes are typically around 1 kcal/mol.<sup>18,56,111,112,114</sup> It is worth noting that the utilized QM/MM scheme with electronic embedding might be artificially affected by overpolarization of the QM region, especially for systems with double negatively charged phosphates in the QM region as is the case of the dianionic reaction mechanism. The spurious overpolarization of the QM region could be overcome, e.g., by a three-layer ONIOM scheme<sup>80</sup> or the Klopman–Ohno formalization.<sup>117</sup> In order to decipher the extent of overpolarization effects in our calculations, we tested both electronic and mechanical QM/MM couplings using two distinct (MPW1K/6-31+G(d,p) and AM1/d-PhoT) approaches for both monoanionic and dianionic reaction scenarios. In case of the monoanionic reaction mechanism, the shapes of PESs were almost identical in both electronic and mechanical embedding. As expected, the effect of QM/MM coupling was more visible in the dianionic reaction mechanism; however, the extent of this effect was still rather small (Figures S5 and S6 in the Supporting Information). In other words, the overpolarization effects do not seem to be significant in this particular case and we are not capable to fully explain the large free energy correction obtained for the dianionic scenario.

**Comparison with Previous Data.** Previous large-scale free energy studies employing the same AM1/d-PhoT semiempirical method have reported different activation barriers for both scenarios.<sup>17,24</sup> For the monoanionic reaction

mechanism, Nam and co-workers reported barriers of 20.5 and 18.1 kcal/mol in the presence of A38 and A38H<sup>+</sup>, respectively, which were in agreement with the experimental data. In contrast, they observed a significantly higher overall barrier of 26 kcal/mol for the dianionic mechanism, which did not support the possible role of G8<sup>-</sup> in the reaction mechanism.<sup>24</sup> In this study, we report comparable AM1/d-PhoT barriers for both the monoanionic and dianionic mechanisms (24.5 and 23.8 kcal/mol, respectively, cf. Table 1), which are, however, slightly higher than the experimentally measured data (20–21 kcal/mol). We can hypothesize that Nam and co-workers obtained lower activation barriers for the monoanionic reaction mechanism owing to addition of DFT-based corrections (ranging between −8.0 and +2.5 kcal/mol) based on calculations of the active site models.<sup>17</sup> On the other hand, the higher activation barrier for the dianionic reaction mechanism calculated by Nam and co-workers might originate from the presence of A38 (instead of A38H<sup>+</sup> used in our work) in the active site.<sup>24</sup>

The other semiempirical Hamiltonian, SCC-DFTBPR, underestimated the free energy barrier (18.8 kcal/mol) for the monoanionic mechanism. The calculations for the dianionic scenario failed completely due to the occurrence of spurious geometries (for both PES and FES profiles) that were possibly caused by overestimated interactions between the scissile phosphate in the QM region and Na<sup>+</sup> counterions in the MM region. We even observed loss of the covalent bond between one of the nonbridging oxygens and the scissile phosphate group. All these observations underline the challenges faced in obtaining accurate semiempirical descriptions of enzymatic reactions.<sup>18</sup>

### Comparison of Dianionic and Monoanionic Reaction Mechanisms.

For the comparison of the performance of different QM/MM methods, we used two mechanisms of hairpin ribozyme self-cleavage suggested in our recent study,<sup>18</sup> where a detailed discussion of these mechanisms including their comparison can be found. The agreement found in the current study between highly accurate ab initio and hybrid DFT calculations further supports the main conclusions of our previous study, where the monoanionic and dianionic mechanisms showed similar activation barriers.<sup>18</sup> Thus, both mechanisms appear viable and may occur/compete during the hairpin ribozyme catalysis. Only the dianionic scenario involving deprotonated G8<sup>-</sup> as the general base and protonated A38H<sup>+</sup> as the general acid can apparently explain the pH rate profiles; our results however, do not rule out the possibility of coexistence of both mechanisms based on their rather similar reaction barriers (see the Supporting Information for more details).

### Semiempirical Methods Predict Different Reaction Pathways.

There is an interesting difference between the higher-level (ab initio and DFT) QM and semiempirical descriptions of the cleavage reaction. Whereas the former methods consistently predict that the reaction involves a concerted pathway for both the monoanionic and dianionic mechanisms, the semiempirical treatments favor sequential pathways.

The ab initio and DFT methods showed that the first proton transfer from the A-1(2'-OH) to either the G+1(*pro-R<sub>p</sub>*) nonbridging oxygen or G8<sup>-</sup>(N1) nitrogen occurs simultaneously with nucleophilic attack of A-1(O2') to G+1(P). In the monoanionic reaction mechanism, the corresponding TS is the rate-limiting step. Subsequent rotation of the G+1(*pro-R<sub>p</sub>H*)

group followed by exocyclic cleavage with departure and protonation of the G+1(O5') group requires significantly lower energy activation.<sup>18</sup> The combined nucleophilic attack and proton transfer step of the dianionic reaction mechanism result in a single TS along the reaction pathway. The additional proton transfer from A38H<sup>+</sup> to G+1(O5') and exocyclic cleavage step (departure of the G+1(5'-OH) group) are barrierless events.<sup>18</sup>

In contrast, both semiempirical methods favored reaction scenarios with a sequential (stepwise) pathway for both the monoanionic and dianionic mechanisms. The first proton transfer from A-1(2'-OH) to either the G+1(*pro-R<sub>p</sub>*) nonbridging oxygen (monoanionic path) or G8<sup>-</sup>(N1) nitrogen (dianionic path) is separated from the subsequent nucleophilic attack of A-1(O2') to G+1(P). In the monoanionic mechanism, the TS is connected to either the first proton transfer from A-1(2'-OH) to the nonbridging oxygen (the A-1(O2')...G+1(P) distance is longer than 2.4 Å, PES profiles; Figures 3D,E and 4B), or subsequent nucleophilic attack steps (and is closer to IN states, FES profiles; Figure 6B and Figure S3 in the Supporting Information). The rate-limiting step in the dianionic mechanism is at the beginning of the exocyclic cleavage, with departure of the G+1(5'-OH) group (Figure 8). Moreover, the dianionic FES computed by AM1/d-PhoT revealed that the first proton transfer from A-1(2'-OH) to the G8<sup>-</sup>(N1) nitrogen forms a stable IN state containing canonical G8 with a deprotonated ribose A-1(2'-O<sup>-</sup>) group (Figure 8B). These significant differences between semiempirical and higher-level QM methods may stem from the substantial approximations of the semiempirical formalisms.

Semiempirical computations are an important tool for investigating the FESs of reactions catalyzed by biomacromolecules because of their affordable computer cost.<sup>17,24,100–110</sup> This allows restrained semiempirical QM/MM–MD simulations required for the construction of FESs. Such simulations are presently not tractable with higher-quality QM methods. However, semiempirical methods heavily rely on their parametrization. Therefore, considerable attention should be paid to careful parametrization of semiempirical methods when modeling ribozyme-catalyzed reactions, i.e., the sugar–phosphate backbone self-cleavage in our case. Until now, direct QM/MM comparison using various semiempirical, DFT, and ab initio approaches has been lacking in the literature. Our study shows that semiempirical methods provide activation barriers in reasonable agreement with QM data. On the other hand, both utilized semiempirical methods (SCC-DFTBPR and AM1/d-PhoT) mapped reaction profiles that were considerably different from those predicted by DFT (BLYP, MPW1K) and ab initio (SCS-MP2) methods, suggesting that different reaction pathways were favored by the semiempirical methods compared to the DFT and ab initio profiles. Consequently, the QM/MM scheme with semiempirical methods may introduce substantial uncertainties when deciphering the reaction mechanisms of RNA catalyzed reactions. On the other hand, because of the enormous computer costs of hybrid DFT and post Hartree–Fock (electron correlation) ab initio methods, calculations of FES profiles of enzymatic reactions using such methods remain impractical. In summary, the quantitative QM/MM description of RNA catalysis remains a formidable challenge for future developments in computational chemistry. Still, different approaches, each with their own limitations, may provide useful preliminary insights. Nevertheless, semiempirical

studies of reaction mechanisms of RNA catalyzed reactions should be interpreted with great care.

## CONCLUSIONS

We constructed PESs and FESs of the self-cleavage reaction catalyzed by the hairpin ribozyme using diverse hybrid QM/MM methods. Three different levels of quantum theory were used for computations of More O'Ferrall–Jencks diagrams, from which activation barriers and reaction pathways were derived. We compared the performances of semiempirical (AM1/d-PhoT and SCC-DFTBPR approaches), DFT (BLYP and MPW1K functionals), and correlated ab initio (SCS-MP2) methods. For the comparison, we considered two reaction scenarios with protonated A38H<sup>+</sup> and either canonical or deprotonated forms of G8 in the active site, termed the monoanionic and dianionic mechanisms.

As summarized in Table 1, all QM/MM methods provided activation barriers with acceptable agreement with the experimental data. On the other hand, there were large differences in the nature of the reaction scenario predicted by the different methods. Both semiempirical methods suggested sequential reaction pathways. This was in sharp contrast to the concerted pathway favored by the SCS-MP2 calculations (our reference method) and both DFT methods. As the approximate semiempirical methods supported the sequential pathway at both the PES and FES levels of description, the preference for the sequential scenario can be primarily attributed to the semiempirical potential rather than application of the free energy MD approach. In addition, gas phase QM calculations ruled out a possibility that the above-noted discrepancy reflected some QM/MM specific setting, such as QM/MM coupling. The inconsistent estimation of PES shapes originated in semiempirical Hamiltonians. Since the results obtained by the robust ab initio and DFT methods were mutually consistent, it appears that semiempirical methods, although clearly superior in terms of sampling, may visibly distort the calculated PESs and FESs of ribozyme reactions.

On the basis of our computations, we recommend using the BLYP functional to construct FESs in the foreseeable future because this DFT provides a correct shape of the PES and is computationally less demanding than hybrid DFT (such as MPW1K) or correlated ab initio methods. However, it should be taken into account that BLYP slightly underestimates the reaction barriers. The BLYP FES may be a posteriori tweaked by adding a correction between the PES constructed by BLYP and a more accurate method (e.g., SCS-MP2, MPW1K, etc.) to the FES constructed by BLYP. In such a case, one may obtain FES profiles that can be interpreted in terms of both the reaction kinetics and reaction mechanism.

## ASSOCIATED CONTENT

### Supporting Information

Detailed description of the dianionic and monoanionic reaction mechanisms. Figures S1–S7 showing additional 2-D profiles are also included. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Authors

\*Tel.: +420 585634769. E-mail: michal.otyepka@upol.cz.

\*Tel.: +440 1179289097. E-mail: adrian.mulholland@bristol.ac.uk.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Dr. Richard Lonsdale for his assistance with setting up the QoMMa calculations. This work was funded by Grants P208/12/1878 (J.S. and M.O.) and P301/11/P558 (P.B.) from the Czech Science Foundation. This work was further supported by the projects “CEITEC—Central European Institute of Technology” CZ.1.05/1.1.00/02.0068 from the European Regional Development Fund (J.S.), Operational Program Research and Development for Innovations—European Regional Development Fund (Project CZ.1.05/2.1.00/03.0058), and the Operational Program Education for Competitiveness—European Social Fund (CZ.1.07/2.3.00/20.0017 and CZ.1.07/2.3.00/20.0058) of the Ministry of Education, Youth and Sports of the Czech Republic (V.M., M.O., P.B.), and the student project PrF\_2013\_028 from Palacky University (V.M.). A.J.M. is an EPSRC Leadership Fellow (Grant EP/G007705/01) and (together with M.W.v.d.K.) thanks the EPSRC for support.

## REFERENCES

- (1) Lilley, D. M. *Trends Biochem. Sci.* **2003**, *28*, 495–501.
- (2) Fedor, M. J.; Williamson, J. R. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 399–412.
- (3) Fedor, M. J. *J. Mol. Biol.* **2000**, *297*, 269–291.
- (4) Fedor, M. J. *Annu. Rev. Biophys.* **2009**, *38*, 271–299.
- (5) Walter, N. G.; Burke, J. M. *Curr. Opin. Chem. Biol.* **1998**, *2*, 24–30.
- (6) Cochrane, J. C.; Strobel, S. A. *Acc. Chem. Res.* **2008**, *41*, 1027–1035.
- (7) Kraut, D. A.; Carroll, K. S.; Herschlag, D. *Annu. Rev. Biochem.* **2003**, *72*, 517–571.
- (8) Rupert, P. B.; Ferre-D'Amare, A. R. *Nature* **2001**, *410*, 780–786.
- (9) Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferre-D'Amare, A. R. *Science* **2002**, *298*, 1421–1424.
- (10) Alam, S.; Grum-Tokars, V.; Krucinska, J.; Kundracik, M. L.; Wedekind, J. E. *Biochemistry* **2005**, *44*, 14396–14408.
- (11) Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J. E. *Biochemistry* **2006**, *45*, 686–700.
- (12) Lebruska, L. L.; Kuzmine, I. I.; Fedor, M. J. *Chem. Biol.* **2002**, *9*, 465–473.
- (13) Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2259–2265.
- (14) Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. *J. Mol. Biol.* **2004**, *340*, 233–251.
- (15) Walter, N. G. *Mol. Cell* **2007**, *28*, 923–929.
- (16) Nam, K.; Cui, Q.; Gao, J. L.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486–504.
- (17) Nam, K. H.; Gao, J. L.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 4680–4691.
- (18) Mlynsky, V.; Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. J. *Phys. Chem. B* **2011**, *115*, 13911–13924.
- (19) Kuzmin, Y. I.; Da Costa, C. P.; Cottrell, J. W.; Fedor, M. J. *J. Mol. Biol.* **2005**, *349*, 989–1010.
- (20) Guo, M.; Spitale, R. C.; Volpini, R.; Krucinska, J.; Cristalli, G.; Carey, P. R.; Wedekind, J. E. *J. Am. Chem. Soc.* **2009**, *131*, 12908–12909.
- (21) Cottrell, J. W.; Scott, L. G.; Fedor, M. J. *J. Biol. Chem.* **2011**, *286*, 17658–17664.
- (22) Bevilacqua, P. C.; Brown, T. S.; Nakano, S.; Yajima, R. *Biopolymers* **2004**, *73*, 90–109.
- (23) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 432–450.
- (24) Nam, K.; Gao, J. L.; York, D. M. *RNA* **2008**, *14*, 1501–1507.
- (25) Liu, L.; Cottrell, J. W.; Scott, L. G.; Fedor, M. J. *Nat. Chem. Biol.* **2009**, *5*, 351–357.

- (26) Nahas, M. K.; Wilson, T. J.; Hohng, S. C.; Jarvie, K.; Lilley, D. M. J.; Ha, T. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1107–1113.
- (27) Wilson, T. J.; Lilley, D. M. *RNA* **2011**, *17*, 213–221.
- (28) Kath-Schorr, S.; Wilson, T. J.; Li, N. S.; Lu, J.; Piccirilli, J. A.; Lilley, D. M. *J. Am. Chem. Soc.* **2012**, *134*, 16717–16724.
- (29) Mlynšky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114*, 6642–6652.
- (30) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. *Acc. Chem. Res.* **2010**, *43*, 40–47.
- (31) Sponer, J.; Sponer, J. E.; Mladek, A.; Banas, P.; Jurecka, P.; Otyepka, M. *Methods* **2013**, *64*, 3–11.
- (32) Torelli, A. T.; Krucinska, J.; Wedekind, J. E. *RNA* **2007**, *13*, 1052–1070.
- (33) Torelli, A. T.; Spitale, R. C.; Krucinska, J.; Wedekind, J. E. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 154–158.
- (34) Macelrevy, C.; Salter, J. D.; Krucinska, J.; Wedekind, J. E. *RNA* **2008**, *14*, 1600–1616.
- (35) Cottrell, J. W.; Kuzmin, Y. I.; Fedor, M. J. *J. Biol. Chem.* **2007**, *282*, 13498–13507.
- (36) CRC Handbook of Chemistry and Physics, 83rd ed.; Lide, D. R., Ed.; CRC: Boca Raton, FL, 2003.
- (37) Ditzler, M. A.; Sponer, J.; Walter, N. G. *RNA* **2009**, *15*, 560–575.
- (38) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (39) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.
- (40) Gao, J. L.; Ma, S. H.; Major, D. T.; Nam, K.; Pu, J. Z.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188–3209.
- (41) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (42) Kamerlin, S. C. L.; Warshel, A. *Proteins* **2010**, *78*, 1339–1375.
- (43) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142–151.
- (44) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *Chem. Soc. Rev.* **2012**, *41*, 3025–3038.
- (45) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.
- (46) Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.
- (47) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (48) Banas, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *Methods* **2009**, *49*, 202–216.
- (49) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708–2728.
- (50) Steinbrecher, T.; Elstner, M. *Methods Mol. Biol.* **2013**, *924*, 91–124.
- (51) Groenhof, G. *Methods Mol. Biol.* **2013**, *924*, 43–66.
- (52) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2011**, *133*, 8934–8941.
- (53) Lans, I.; Medina, M.; Rosta, E.; Hummer, G.; Garcia-Viloca, M.; Lluch, J. M.; Gonzalez-Lafont, A. *J. Am. Chem. Soc.* **2012**, *134*, 20544–20553.
- (54) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (55) Yang, Y.; Yu, H. B.; York, D.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2008**, *4*, 2067–2084.
- (56) Banas, P.; Rulisek, L.; Hanosova, V.; Svozil, D.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2008**, *112*, 11177–11187.
- (57) Ganguly, A.; Bevilacqua, P. C.; Hammes-Schiffer, S. *J. Phys. Chem. Lett.* **2011**, *2*, 2906–2911.
- (58) Thaplyal, P.; Ganguly, A.; Golden, B. L.; Hammes-Schiffer, S.; Bevilacqua, P. C. *Biochemistry* **2013**, *52*, 6499–6514.
- (59) Csonka, G. I.; Johnson, B. G. *Theor. Chem. Acc.* **1998**, *99*, 158–165.
- (60) Mori-Sanchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2006**, *125*, 201102.
- (61) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
- (62) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811–4815.
- (63) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936–2941.
- (64) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. T. *Chem. Rev.* **2012**, *112*, 289–320.
- (65) Wang, S. L.; Hu, P.; Zhang, Y. K. *J. Phys. Chem. B* **2007**, *111*, 3758–3764.
- (66) Hu, P.; Wang, S.; Zhang, Y. *J. Am. Chem. Soc.* **2008**, *130*, 3806–3813.
- (67) Hu, P.; Wang, S. L.; Zhang, Y. K. *J. Am. Chem. Soc.* **2008**, *130*, 16721–16728.
- (68) Zhou, Y. Z.; Wang, S. L.; Zhang, Y. K. *J. Phys. Chem. B* **2010**, *114*, 8817–8825.
- (69) Ke, Z. H.; Guo, H.; Xie, D. Q.; Wang, S. L.; Zhang, Y. K. *J. Phys. Chem. B* **2011**, *115*, 3725–3733.
- (70) Ganguly, A.; Thaplyal, P.; Rosta, E.; Bevilacqua, P. C.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2014**, *136*, 1483–1496.
- (71) Rosta, E.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2014**, DOI: 10.1021/ja411408x.
- (72) Warshel, A.; Weiss, R. M. *Ann. N. Y. Acad. Sci.* **1981**, *367*, 370–382.
- (73) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210–3235.
- (74) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley and Sons: New York, 1991.
- (75) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (76) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (77) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- (78) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (79) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
- (80) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.
- (81) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Comperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (82) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12.0*; University of California: San Francisco, 2012.
- (83) Harvey, J. N. *Faraday Discuss.* **2004**, *127*, 165–177.
- (84) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016–1024.
- (85) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.;

- Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Koppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pfluger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. MOLPRO, version 2012.1, a package of ab initio programs, 2012.
- (86) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (87) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437–446.
- (88) Goumans, T. P. M.; Ehlers, A. W.; Lammertsma, K.; Wurthwein, E. U.; Grimme, S. *Chem.—Eur. J.* **2004**, *10*, 6468–6475.
- (89) Schwabe, T.; Grimme, S. *Acc. Chem. Res.* **2008**, *41*, 569–579.
- (90) van der Kamp, M. W.; Zurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. *J. Phys. Chem. B* **2010**, *114*, 11303–11314.
- (91) Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5291–5303.
- (92) Pitonak, M.; Neogrady, P.; Cerny, J.; Grimme, S.; Hobza, P. *ChemPhysChem* **2009**, *10*, 282–289.
- (93) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Chem. Theory Comput.* **2012**, *8*, 4637–4645.
- (94) Nam, K.; Gao, J. L.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 2–13.
- (95) Walker, R. C.; Crowley, M. F.; Case, D. A. *J. Comput. Chem.* **2008**, *29*, 1019–1031.
- (96) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339–1350.
- (97) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.
- (98) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., 3rd; Jurecka, P. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.
- (99) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, CA, 2006.
- (100) Lopez-Canut, V.; Marti, S.; Bertran, J.; Moliner, V.; Tunon, I. *J. Phys. Chem. B* **2009**, *113*, 7816–7824.
- (101) Lopez-Canut, V.; Ruiz-Pernia, J.; Tunon, I.; Ferrer, S.; Moliner, V. *J. Chem. Theory Comput.* **2009**, *5*, 439–442.
- (102) Lopez-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tunon, I. *J. Am. Chem. Soc.* **2010**, *132*, 6955–6963.
- (103) Lopez-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tunon, I. *J. Am. Chem. Soc.* **2011**, *133*, 12050–12062.
- (104) Wong, K. Y.; Lee, T. S.; York, D. M. *J. Chem. Theory Comput.* **2011**, *7*, 1–3.
- (105) Yang, Y.; Yu, H.; Cui, Q. *J. Mol. Biol.* **2008**, *381*, 1407–1420.
- (106) Yang, Y.; Cui, Q. *J. Phys. Chem. A* **2009**, *113*, 12439–12446.
- (107) Zhang, C. H.; Gao, J. Y.; Chen, Z. Q.; Xue, Y. *J. Mol. Graphics Modell.* **2010**, *28*, 799–806.
- (108) Capoferri, L.; Mor, M.; Sirirak, J.; Chudyk, E.; Mulholland, A. J.; Lodola, A. *J. Mol. Model.* **2011**, *17*, 2375–2383.
- (109) Hou, G.; Cui, Q. *J. Am. Chem. Soc.* **2012**, *134*, 229–246.
- (110) Hou, G.; Cui, Q. *J. Am. Chem. Soc.* **2013**, *135*, 10457–10469.
- (111) Torres, R. A.; Himo, F.; Bruice, T. C.; Noddleman, L.; Lovell, T. *J. Am. Chem. Soc.* **2003**, *125*, 9861–9867.
- (112) Lopez, X.; Dejaegere, A.; Leclerc, F.; York, D. M.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 11525–11539.
- (113) Claeysens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schutz, M.; Thiel, S.; Thiel, W.; Werner, H. *J. Angew. Chem., Int. Ed.* **2006**, *45*, 6856–6859.
- (114) Chval, Z.; Chvalova, D.; Leclerc, F. *J. Phys. Chem. B* **2011**, *115*, 10943–10956.
- (115) Zhao, Y.; Gonzalez-Garcia, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.
- (116) Wolfenden, R.; Snider, M. *J. Acc. Chem. Res.* **2001**, *34*, 938–945.
- (117) Hou, G.; Zhu, X.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2012**, *8*, 4293–4304.
- (118) Zurek, J.; Bowman, A. L.; Sokalski, W. A.; Mulholland, A. J. *Struct. Chem.* **2004**, *15*, 405–414.