Exploring chemical reaction mechanisms through harmonic Fourier beads path optimization

Ilja V. Khavrutskii,[a] Jason B. Smith, and Anders Wallqvist
Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, United States Army Medical Research and Materiel Command, Fort Detrick, Maryland 21702, USA

(Received 7 August 2013; accepted 8 October 2013; published online 25 October 2013)

Here, we apply the harmonic Fourier beads (HFB) path optimization method to study chemical reactions involving covalent bond breaking and forming on quantum mechanical (QM) and hybrid QM/molecular mechanical (QM/MM) potential energy surfaces. To improve efficiency of the path optimization on such computationally demanding potentials, we combined HFB with conjugate gradient (CG) optimization. The combined CG-HFB method was used to study two biologically relevant reactions, namely, L- to D-alanine amino acid inversion and alcohol acylation by amides. The optimized paths revealed several unexpected reaction steps in the gas phase. For example, on the B3LYP/6-31G(d,p) potential, we found that alanine inversion proceeded via previously unknown intermediates, 2-iminopropane-1,1-diol and 3-amino-3-methyloxiran-2-ol. The CG-HFB method accurately located transition states, aiding in the interpretation of complex reaction mechanisms. Thus, on the B3LYP/6-31G(d,p) potential, the gas phase activation barriers for the inversion and acylation reactions were 50.5 and 39.9 kcal/mol, respectively. These barriers determine the spontaneous loss of amino acid chirality and cleavage of peptide bonds in proteins. We conclude that the combined CG-HFB method further advances QM and QM/MM studies of reaction mechanisms. [http://dx.doi.org/10.1063/1.4826470]

I. INTRODUCTION

Our understanding of complex chemical transformations relies on the identification of elementary reactions involved in the process: the reaction mechanism. Each elementary reaction involves a reactant, a product, and a connecting transition state. Computational chemistry maps these reactions onto potential energy surfaces, providing both structures and energetics for analysis. The mapping relies on energy optimization to identify a reaction path that starts at the reactant, passes through the transition state, and ends at the product.

Transition states (first-order saddle points on potential energy surfaces) determine the rates of elementary reactions[1,2] but are hard to find.[3] The search for transition states requires specialized optimization techniques that climb potential energy ridges and succeed only when the starting structures are sufficiently close to the transition states.[4–9] Furthermore, slight differences in starting structures can lead to different, often irrelevant transition states.[10]

In contrast to transition states, reactant, product, and intermediate states (local minima on potential energy surfaces) do not require specialized optimization techniques. The local minima can be efficiently located using conventional downhill energy minimization techniques.[3,7,9]

Before the reactant, transition, and product states can be used to represent a reaction path on a corresponding three-point energy diagram, the transition state must be validated. First, the gradient of the potential at the transition state must be zero, just like at a local minimum. Second, a separate normal mode calculation must confirm that the transition state has a single imaginary frequency, i.e., is a first-order saddle point. The magnitude of the imaginary frequency is representative of the potential energy curvature at the transition state. Displacements along the normal mode vector with imaginary frequency should move key atoms in the direction of either the reactant or the product. Ideally, an additional calculation, namely, the intrinsic reaction coordinate (IRC),[11–13] should be performed to ensure that the transition state connects the reactant with the product and not with any other unexpected intermediate state. Both the normal mode and IRC calculations become prohibitively expensive as the system size grows, necessitating other means of transition state validation.

Because finding the correct transition states using specialized optimization techniques is fraught with challenges, the process becomes more of an art than science, hindering studies of reaction mechanisms. The situation could be improved with the introduction of computational tools that robustly pinpoint correct transition states with minimal user input. We think that full reaction path optimization is one of the best solutions to this problem. In contrast to optimization techniques that operate on a single state, reaction path optimization techniques operate on a "chain of states" that connect the reactant and product states.[2–4,14–24] This connection between the states in the path should be sufficient to accurately and robustly pinpoint the transition states using simple energy minimizations.

[a]Author to whom correspondence should be addressed. Electronic mail: ikhavrutskii@bhsai.org

Here, to study chemical reactions we use an approach that can simultaneously locate and validate transition states in a robust manner. The approach is based on the reaction path optimization method, called harmonic Fourier beads (HFB).[25–27] The HFB method builds on such popular path optimization methods as the nudged elastic band (NEB)[2,23] and string methods.[4,10,14–17,19,28,29] Unlike NEB and string method, HFB uses the Fourier parametrization of the path pioneered by Doll, Freeman, and their co-workers.[30,31] Originally, the HFB method was used with molecular mechanical (MM) potentials that did not allow covalent bond breaking or forming.[25–27,32–34] The present work extends the applicability of the HFB method to chemical reactions involving covalent bond breaking and forming on quantum mechanical (QM) and hybrid QM/MM potentials.

Because QM and QM/MM potentials are computationally more demanding than MM potentials, we combined the HFB method with the conjugate gradient (CG) optimizer[35] for added efficiency. The combination is, thus, referred to as CG-HFB. Almost two decades ago, Doll, Freeman, and Matro devised a method to search for transition states using the Fourier parametrization of the path with CG optimizer.[30,31] In particular, they attempted to locate transition states by optimizing the Fourier amplitudes of the path – an approach that, as the authors acknowledged, was not entirely reliable.[31] In contrast, the CG-HFB optimizes the Cartesian coordinates of the snapshots or beads along the path and uses the Fourier parametrization to string them. Furthermore, CG-HFB uses adaptive harmonic restraints during optimization to prevent the beads from collapsing into the nearest minima. These restraints transform the rugged energy landscapes into simpler, funnel-like landscapes,[36,37] which simplify the optimization of all the beads and make the CG-HFB method more reliable.

With the help of the combined CG-HFB method we studied mechanisms of two biologically relevant reactions. Specifically, we studied monomolecular L- to D-alanine amino acid inversion and bimolecular alcohol acylation by amides. All reactions were studied in the gas phase using Hartree-Fock self-consistent field (HF-SCF)[38] and a Becke three-parameter hybrid functional combined with Lee-Young-Parr correlation functional (B3LYP)[39–41] with 6-31G(d,p)[42] basis set. CG-HFB has successfully identified multiple reaction paths and located corresponding transition states. Importantly, CG-HFB required only minimal initial preparation and a few trivial adjustments over the course of path optimizations, thus significantly simplifying the search for and validation of transition states.

II. METHODS

A. Brief overview of the CG-HFB method

The key ideas and features of the combined CG-HFB implementation and its differences from earlier HFB implementations[25–27,32–34] and other related methods[2,14,15,23,28,43–46] are provided in the supplementary material.[64]

In brief, CG-HFB performs optimization in Cartesian coordinates, which avoids costly transformations and complicated corrections associated with internal coordinates. The CG optimization is performed on the landscape that is modified by adaptive positional harmonic restraints, which simplify the optimization and prevent abrupt structural changes along the path. Optionally, the CG optimizations can be performed orthogonally to the path. The CG optimizations in Cartesian coordinates require the least amount of storage and make the CG-HFB method particularly suited for large systems.

Each CG-HFB path optimization step begins with generating coordinates of equidistant beads. With the exception of the initial step, the equidistant coordinates are obtained from the coordinates of the beads optimized at the previous step with the help of the Fourier parametrization. The equidistant coordinates are used as the new reference structures or anchors of the harmonic restraints. The CG optimizations of the harmonically restrained beads are then performed. The optimization step ends when all of the CG optimizations of the harmonically restrained beads finish. The CG optimizations of the beads between the path optimization steps are independent and, therefore, can run in parallel for speed.

Importantly, CG-HFB can extract transition states from optimized paths using continuous energy, gradient, and coordinate path curves. The transition states are expected to be accurate when the continuous energy profiles reconstructed from the gradients at the beads match the actual energies of the beads along the path. To fulfill this condition, additional beads can be inserted into the path. The CG-HFB optimized paths automatically satisfy the IRC condition[11–13] and, hence, validate the extracted transition states.

B. L- to D-alanine inversion

To study the inversion of an asymmetric α-carbon of alanine in the gas phase, we used the neutral alanine molecule shown in Fig. 1.

1. Generating initial paths

First, we generated the coordinates of L-alanine, the reactant. The reactant was optimized to the nearest local minimum using single-bead HFB optimization with soft harmonic restraints on all the atoms. The coordinates of the optimized L-alanine were then inverted to yield D-alanine, the product.

To explore alanine inversion, we prepared three crude intermediates that would bias the reaction paths to go via three different routes. In particular, by displacing the H_α proton along the covalent bonds connecting the C_α carbon atom to the three chemical groups, namely, $–NH_2$, –C(O)OH, and

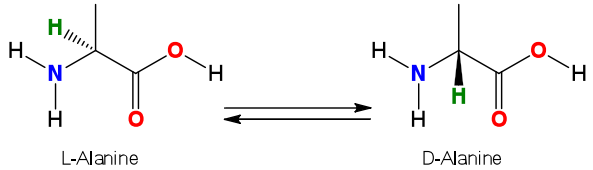

FIG. 1. Alanine inversion. The H_α proton that is involved in the alanine inversion is shown in bold green.

–CH_3, we created initial coordinates for the intermediates. The coordinates of each intermediate were then inserted into the input *xyz* file (refer to supplementary material[64] for details) between the reactant and product coordinates and passed to CG-HFB to generate three initial paths.

2. Optimization of the paths

Path optimizations began at the HF/6-31G level of theory and progressed through HF/6-31G(d) to HF/6-31G(d,p) and finally to B3LYP/6-31G(d,p). The bulk of optimization was performed using 24 beads/path. If necessary, the number of beads in each path was increased to obtain accurate integral energy profiles and transition states.

Paths with more than one transition state were split into elementary segments, each containing a single transition state. Segments were then optimized separately. The force constants and numbers of beads could be adjusted for each segment according to the transition state energy and curvature. Ultimately, this flexibility made the optimizations more efficient and allowed for improved integral energy profiles and transition state structures.

Transition states extracted from the CG-HFB optimized paths were used as initial guesses to obtain what we call "exact transition states." The exact transition states were obtained using the best transition state optimization method[7,9] available in GAUSSIAN09[47] with the tight convergence criteria. Typically, the optimizations to get the exact transition states began with computing analytical Hessians and then continued with reduced maximum step size of 0.015 Å. The majority of the optimizations completed in less than 10 steps.

C. Acylation by peptide

We studied the acylation reaction with the CG-HFB method using both QM and QM/MM potentials. Two systems, namely, "small" and "large" were considered as shown in Fig. 2. The small system comprised methanol and formamide. The large system substituted *N*-methylacetamide for formamide.

1. ONIOM QM/MM setup

The large system was partitioned such that the two methyl groups of *N*-methylacetamide were in the MM level (see Fig. 2). With this partitioning, the QM layer of the large system in ONIOM QM/MM is identical to that of the small system.

Any ONIOM QM/MM reaction path calculation depends somewhat on the MM parameters that span the QM region.[48] This is why careful partitioning of the system between QM and MM layers is crucial. A problematic partitioning scheme would terminate the QM layer too close to a covalent bond being broken, formed, or altered in a way that reduces its bond order.[48] MM atoms up to two bonds away from the affected bond are influenced by the MM dihedral terms involving the affected bond. Moreover, atoms one bond away from the affected bond are involved in the corresponding MM angle bending terms. None of these terms can be properly cancelled by the ONIOM QM/MM scheme.[48] Whenever possible, it is best to ensure that no purely MM atom is influenced by the affected bonds.

Due to the small size of the system being studied in this work, we could not avoid the partitioning issues[48] described above and needed a compromise. Because our goal was to assess the utility of the CG-HFB method for QM/MM applications, we favored simplicity over accuracy in the ONIOM QM/MM calculations. To remove spurious dihedral and angular MM terms associated with the modified bonds, we simply removed such bonds from the MM connectivity tables. To assess the effect of the atom types and of the vdW parameters on energy profiles, we performed path optimizations using both the reactant state, i.e., methanol and *N*-methylacetamide, and the product state, i.e., the methyl ester of acetic acid and *N*-methylamine. In these calculations, we allowed the partial charges to be computed on the fly with the charge equilibration method.[49]

2. Generating initial paths

To initialize reaction path optimization, we prepared several intermediate states using single-bead optimizations with added distance restraints (see supplementary material[64]).

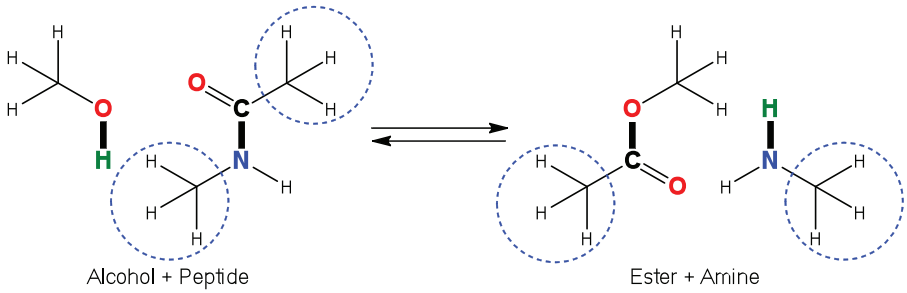

FIG. 2. Schematic representation of alcohol acylation by a peptide. Covalent bonds involved in the reaction are shown in bold. The hydrogen of the alcohol that transfers to the amide is shown in bold green. The methyl groups that are partitioned into the MM layer of "our own *N*-layer integrated molecular orbital MM" method [ONIOM(QM:MM)] calculation are circled. The reactant and product as drawn correspond to the larger model. The small system was derived from the large system by substituting the circled methyl groups for a hydrogen atom.

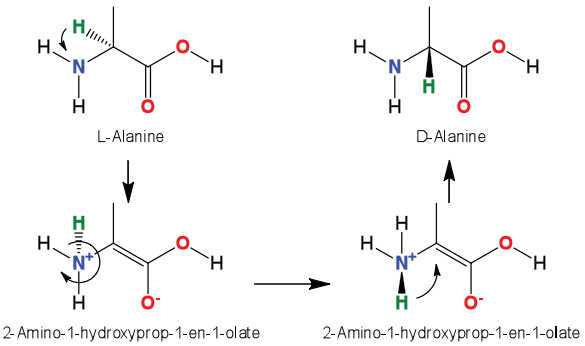

FIG. 3. Key steps of the C_α-NH_2 path. The H_α proton that is involved in the alanine inversion is shown in bold green.

Coordinates of reactant and product states for bimolecular reactions were obtained by optimization on the corresponding sides of the reaction barrier.

3. CG-HFB optimization of the paths

Path optimizations of the large system began at the HF/6-31G level of theory and progressed through HF/6-31G(d) to HF/6-31G(d,p) and finally to B3LYP/6-31G(d,p). For the small system and for the large system treated with ONIOM QM/MM, the highest level of QM used was HF/6-31G(d,p). Optimizations of the paths were performed starting with 24 beads/path. The number of beads was gradually increased to provide accurate integral energy profiles and transition states. The exact transition states were obtained as described for the alanine inversion.

III. RESULTS AND DISCUSSION

In this work, we explored two biologically relevant model reactions. First, we studied the inversion of the asymmetric carbon of alanine α-amino acid, converting it from the L-stereoisomer to the D-stereoisomer. Second, we studied the one-step acylation of an alcohol by an amide, a peptide mimic. Both reactions require QM description as they involve covalent bond breaking and forming. The barrier for alanine inversion provides the upper limit to spontaneous racemization of α-amino acids that comprise all known life forms. On the other hand, the barrier for the alcohol acylation serves as the limit to the spontaneous acylation that would disrupt peptide bonding between amino acids in proteins.

A. L- to D-alanine inversion

To study the inversion of an asymmetric α-carbon of alanine in the gas phase, we used the neutral alanine molecule shown in Fig. 1. Technical details of path preparation and optimization are provided in Sec. II. Here, we provide energetics at the B3LYP/6-31G(d,p) level of theory. The supplementary material[64] also provides results at the HF/6-31G(d,p) level of theory. Although the HF/6-31G(d,p) results differed from B3LYP/6-31G(d,p) results quantitatively and, sometimes, qualitatively, they helped identify intermediates that we would have missed otherwise.

1. The C_α-NH_2 path

The optimized C_α-NH_2 path comprises six segments, featuring seven local minima and six transition states. The key steps of the path are shown in Fig. 3. The energetics and other parameters of the actual path are shown in Table I and Fig. 4. The path identified three different conformations of alanine, namely, S1/R, S1/P (S2/R), and S5/P (S6/R). Here, S1/R and S1/P refer to the reactant and product of the first segment, respectively. Interestingly, the S1/P conformation was 1.8 kcal/mol lower in energy than the starting S1/R conformation. Accidentally, the S1/P conformation corresponds to the global minimum of alanine in the gas phase that was the subject of earlier theoretical[50–52] and experimental[53,54] studies.

The C_α-NH_2 inversion began with $-NH_2$ group rotation into the lowest energy conformation of alanine. The barrier for the rotation was 0.3 kcal/mol in the forward direction. Next, the H_α proton transferred to the nitrogen of the $-NH_2$ group, turning it into positively charged $-NH_3^+$. The barrier for the proton transfer was the highest energy point along the path at 65.7 kcal/mol from the lowest energy conformation, S1/P.

The six atoms of the resulting meta-stable intermediate S2/P lie in a prochiral plane with the sp^2 C_α atom in the

TABLE I. Summary of the C_α-NH_2 path for alanine inversion.[a]

Path segment	E_R (kcal/mol)	E_P (kcal/mol)	E_{TS} (kcal/mol)	$\Delta E_{HFB\text{-}TS}$ (kcal/mol)	*Imag Freq*$_{TS}$ (Δ*Freq*$_{HFB\text{-}TS}$) (cm^{-1})	*TS-HFB RMSD*$_{All}$ (*RMSD*$_{\text{-}HMethyl}$) (Å)	k (kcal/mol/Å^2)	$M(K)$
S1	0.0	−1.8	0.3	0.00	168 (0)	0.002 (0.002)	70	23 (12)
S2	−1.8	32.9	63.8	0.02	1627 (−4)	0.018 (0.014)	1400	64 (64)
S3	32.9	32.9	39.4	0.00	300 (−1)	0.002 (0.001)	70	31 (16)
S4	32.9	−1.8	63.8	0.02	1627 (−4)	0.022 (0.017)	1400	64 (64)
S5	−1.8	−0.2	0.2	0.00	177 (−1)	0.003 (0.003)	70	17 (9)
S6	−0.2	0.0	4.0	0.08	653 (1)	0.045 (0.044)	500	20 (10)

[a]E_R, E_P, and E_{TS} are B3LYP/6-31G(d,p) energy values relative to the original conformation of alanine (−323.754250 Hartree) of the exact reactant, product, and transition state (TS), respectively. $\Delta E_{HFB\text{-}TS}$ is the energy difference between the TS extracted from the conjugate gradient-harmonic Fourier beads (CG-HFB)-optimized path and the exact TS. The force constant (k) is different for different segments. The M (K) column shows the total number of beads (M) with the number of Fourier terms in the series (K) used in the final optimized path in parentheses. All segments were optimized without projecting the forces. The *Imag Freq*$_{TS}$ column shows the exact imaginary frequency values for the TSs and their differences with CG-HFB values in parentheses. The *TS-HFB RMSD*$_{All}$ column shows all-atom root mean square values, with *RMSD*$_{\text{-}HMethyl}$ values excluding the hydrogen atoms of the methyl groups in parentheses.

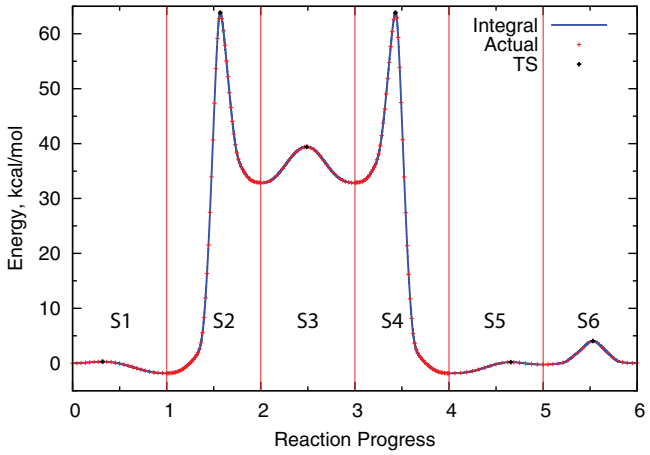

FIG. 4. Energy profile for the full C_α-NH_2 path. The integral energy profile reconstructed from the forces is shown in solid blue. The actual energy points offset to match the integral profile are shown as red plus signs. For clarity, the elementary segments are separated by vertical lines. Each segment contains a single transition state (TS) marked by a black "+" and covers one unit of the progress variable.

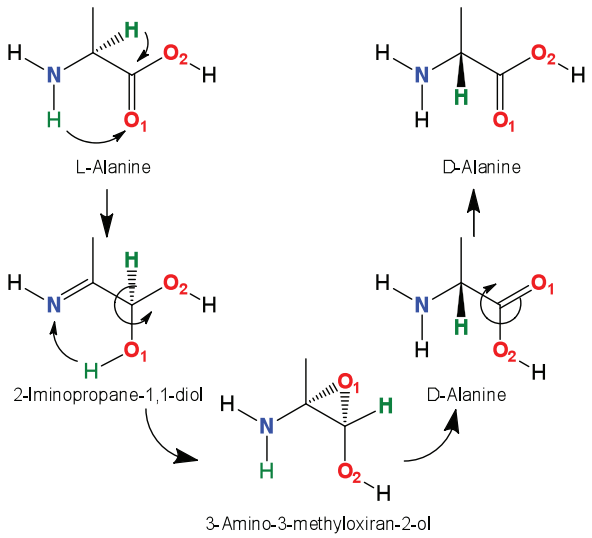

FIG. 5. Key steps of C_α-C(O)OH *path* 1. The H_α proton that is involved in the alanine inversion is shown in bold green.

center. The planar configuration was stabilized by an allylic-like resonance with the carbonyl oxygen of the –C(O)OH group, which delocalized the negative charge from the C_α atom. The metastable intermediate S2/P was 34.7 kcal/mol above the lowest energy conformation of alanine.

Subsequently, the $–NH_3^+$ group rotated followed by the symmetric proton transfer back to the C_α atom on the other side of the prochiral plane. Rotation of the $–NH_3^+$ group in the prochiral, meta-stable intermediate had a symmetric barrier of 6.5 kcal/mol. In principle, the $–NH_3^+$ rotation is not required as the three hydrogen atoms are equivalent. However, by construction, our path follows a particular hydrogen atom all the way through the inversion. The barrier for the H_α proton transfer back to the C_α atom on the other side of the prochiral meta-stable intermediate was 31.0 kcal/mol.

The final stage of the C_α-NH_2 inversion could be as simple as the mirror image of the initial stage, i.e., rotation of the $–NH_2$ group. However, our optimized path concluded differently. Specifically, starting from the lowest energy conformation of D-alanine, namely, S4/P, the $–NH_2$ group rotated into yet another conformation of alanine, S5/P, with an energy of 1.6 kcal/mol and a forward barrier of 2.0 kcal/mol. From this intermediate, the $–NH_2$ group inverted into the final product configuration, S6/P, with a forward barrier of 4.2 kcal/mol. The $–NH_2$ inversion brought the transformation to a close with the final product, S6/P, precisely mirroring the S1/R conformation of the reactant L-alanine.

2. C_α-C(O)OH paths

The exploratory C_α-C(O)OH path surpassed the C_α-NH_2 path in complexity and barrier heights. Despite the higher barriers in certain segments of the exploratory C_α-C(O)OH path, our analysis identified an important segment with a relatively low barrier. Specifically, the path included a proton transfer along the C_α-C(O)OH bond with a barrier of 50.5 kcal/mol. After the transfer, the path detoured through much higher energy regions. Still, this finding suggested that a lower energy C_α-C(O)OH path could exist. Therefore, we rationally redesigned the C_α-C(O)OH path with the hope of finding a route with barrier heights below 65.7 kcal/mol.

a. H_α proton transfer to the –C(O)OH group. The H_α proton transfer to the –C(O)OH group transformed the carboxyl group into a geminal diol $–C(OH)_2H$ group, as shown in Fig. 5. Even when forced to start from the lowest energy conformation of alanine, the reaction proceeded through the original conformation.

Interestingly, the H_α proton transfer concluded with a secondary proton transfer. The secondary proton transferred from the $–NH_2$ group to the negatively charged oxygen of the $–C(O^-)(OH)H$ group. The latter group existed in a meta-stable intermediate on the HF/6-31G(d,p) potential (see supplementary material[64]) but not on the B3LYP/6-31G(d,p) potential. The secondary proton transfer required the $–O^-$ as an acceptor. Relative to the lowest energy conformation of alanine, the barrier for the double proton transfer, S1/TS, was 52.3 kcal/mol and the product diol was 21.0 kcal/mol. This elementary reaction corresponds to the first segment, S1, of the redesigned C_α-C(O)OH path, as shown in Fig. 6 and Table II.

The double proton transfer was followed by the rotation of one of the geminal –OH groups that stabilized the diol by 0.2 kcal/mol. The barrier for this rotation was only 0.3 kcal/mol. This rotation was not observed at the HF/6-31G(d,p) level and added a separate segment to the C_α-C(O)OH path, S2.

From the geminal diol the inversion could proceed via two different paths, path 1 and path 2, which are discussed below.

b. C_α-C(O)OH path 1. Segments S3 and S4 of *path* 1 (Fig. 6) involved an oxiran-2-ol intermediate, as shown in Fig. 5. This path was originally found on the HF/6-31G(d,p) potential energy surface (see supplementary material[64]), and

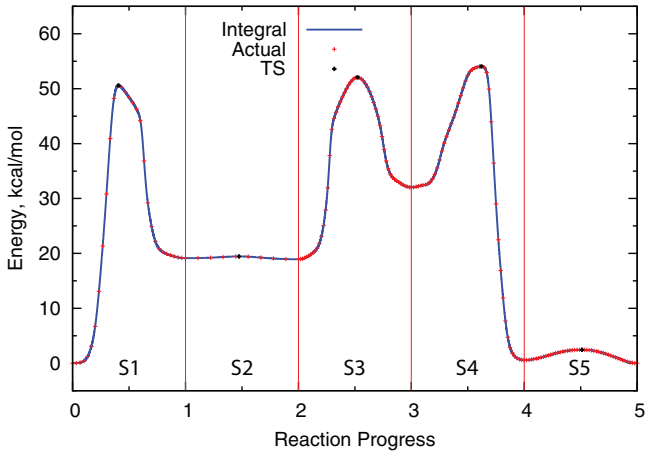

FIG. 6. Energy profile for full C_α-C(O)OH *path* 1. The integral energy profile reconstructed from the forces is shown in solid blue. The actual energy points offset to match the integral profile are shown as red plus signs. The elementary segments are separated by vertical lines for clarity. Each segment contains a single TS marked by a black "+" and covers one unit of the progress variable.

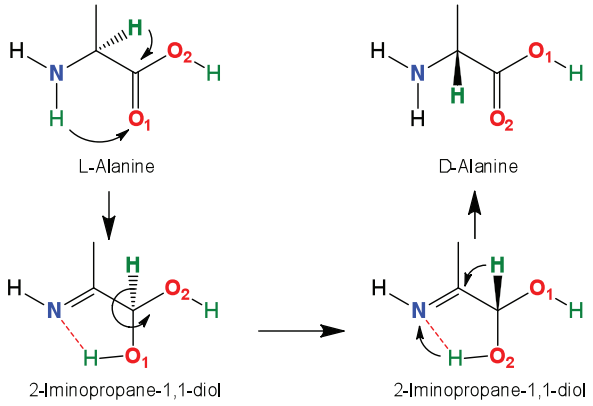

FIG. 7. Key steps of C_α-C(O)OH *path* 2. The H_α proton that is involved in the alanine inversion is shown in bold green.

instead of the geminal diol, HN=C_α(CH_3)–C(OH)$_2$H involved its zwitterionic form, H_2N^+=C_α(CH_3)–C(O^-)(OH)H. The zwitterionic form was not stable on the B3LYP/6-31G(d,p) potential energy surface.

b.1. Formation of oxiran-2-ol from the geminal diol. Two oxiran-2-ol intermediates have been identified, as described in the supplementary material.[64] One of the oxiran-2-ol intermediates could be used to complete the inversion, as schematically shown in Fig. 5. Relative to the lowest conformation of alanine, the energy of the productive oxiran-2-ol was 33.9 kcal/mol. The transition state S3/TS to form the oxiran-2-ol was 53.9 kcal/mol above the lowest conformation of alanine. Interestingly, oxiran-2-ol formation was also coupled to a proton transfer. Specifically, the proton from one of the –OH groups of the geminal diol transferred back to the =NH group. This reaction corresponds to the S3 segment shown in Fig. 6.

b.2. H_α proton transfer back to the C_α atom. Finally, we optimized proton transfer from the productive oxiran-2-ol back to the C_α atom, effecting the inversion in the S4 segment, as shown in Fig. 6. Relative to the lowest energy conformation of alanine, the energy of the transition state for the final proton transfer, S4/TS, shown in Fig. 6 was 55.8 kcal/mol. Considering the S4 reaction segment in reverse, it becomes clear that the oxiran-2-ol can be formed from alanine directly, bypassing the diol intermediate. However, this requires a rotation of the protonated carboxyl group.

b.3. Rotation of the carboxyl group. The carboxyl group in the inverted alanine was flipped compared with the original alanine. Indeed, the hydroxy (–OH) group rather than the carbonyl (=O) group interacted with the –NH_2 group in the inverted alanine. The final segment, S5, of the redesigned C_α-C(O)OH *path* 1 shown in Fig. 6 is the rotation of the carboxyl group. The original conformation of alanine, S1/R, was only 0.6 kcal/mol lower than the flipped conformation, S5/R. The carboxyl group can rotate in two opposing directions. The lowest barrier for the carboxyl group rotation from the flipped conformation, S5/R, toward the original conformation S5/P, as identified in the supplementary material,[64] was 1.9 kcal/mol.

From C_α-C(O)OH *path* 1, we conclude that the orientation of the carboxyl –C(O)OH group determines if the H_α proton transfer will result in geminal diol or oxiran-2-ol formation.

c. C_α-C(O)OH path 2. Because the formation of oxiran-2-ol from alanine had a higher barrier than that of the geminal diol, we constructed *path* 2, which bypasses oxiran-2-ol, as schematically shown in Fig. 7. *Path* 2 does not involve any other intermediates besides the geminal diol but requires a rotation of the –C(OH)$_2$H group. The rotation moves the proton

TABLE II. Summary of C_α-C(O)OH path 1 for alanine inversion.[a]

Path segment	E_R (kcal/mol)	E_P (kcal/mol)	E_{TS} (kcal/mol)	$\Delta E_{HFB\text{-}TS}$ (kcal/mol)	*Imag Freq*$_{TS}$ (Δ*Freq*$_{HFB\text{-}TS}$) (cm^{-1})	*TS-HFB RMSD*$_{All}$ (*RMSD*$_{\text{-}HMethyl}$) (Å)	*k* (kcal/mol/Å^2)	*M*(*K*)
S1	0.0	19.1	50.5	−0.10	606 (29)	0.070 (0.014)	1400	31 (16)
S2	19.1	18.9	19.5	0.00	205 (−9)	0.011 (0.004)	50	10 (6)
S3	18.9	32.0	52.0	−0.03	177 (−3)	0.026 (0.012)	1400	55 (28)
S4	32.0	0.6	53.9	−0.11	344 (−64)	0.103 (0.013)	1000	49 (25)
S5	0.6	0.0	2.5	0.00	28 (0)	0.000 (0.000)	70	48 (24)

[a]E_R, E_P, and E_{TS} are B3LYP/6-31G(d,p) energy values relative to the original conformation of alanine (−323.754250 Hartree) of the exact reactant, product, and TS, respectively. $\Delta E_{HFB\text{-}TS}$ is the energy difference between the TS extracted from the CG-HFB-optimized path and the exact TS. The force constant (*k*) is different for different segments. The *M* (*K*) column shows the total number of beads (*M*), with the number of Fourier terms in the series (*K*) used in the final optimized path in parentheses. All segments were optimized without projecting the forces. The *Imag Freq*$_{TS}$ column shows the exact imaginary frequency values for the TSs and their differences with CG-HFB values in parentheses. The *TS-HFB RMSD*$_{All}$ column shows all-atom root mean square values, with *RMSD*$_{\text{-}HMethyl}$ values excluding the hydrogen atoms of the methyl groups in parentheses.

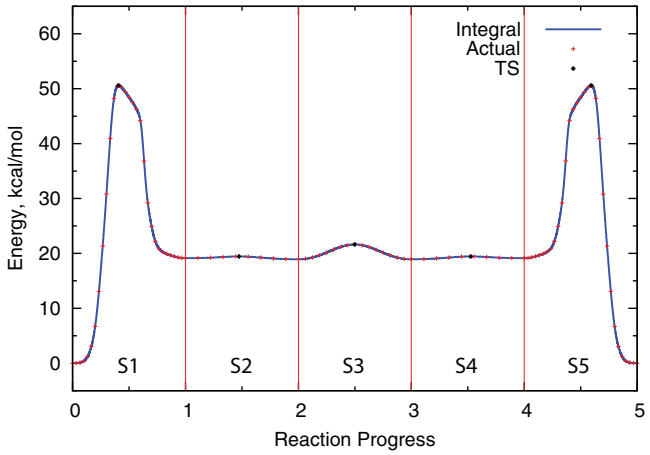

FIG. 8. Energy profile for full C_α-C(O)OH *path* 2. The integral energy profile reconstructed from the forces is shown in solid blue. The actual energy points offset to match the integral profile are shown as red plus signs. The elementary segments are separated by vertical lines for clarity. Each segment contains a single TS marked by a black "+" and covers one unit of the progress variable.

to the other side of the prochiral plane of C_α from which the proton transfers back to C_α, concluding the inversion.

The energy profile along *path* 2 is shown in Fig. 8. The first two segments of *path* 2 are the same as in *path* 1. Segments S4 and S2 as well as S5 and S1 are identical after inversion. Just like the carboxyl group, the $-C(OH)_2H$ group can rotate in two opposing directions. The lowest energy rotation of the geminal $-C(OH)_2H$ group was identified from the exploratory 360° rotation path. The exploratory path contained eight elementary segments, as shown in the supplementary material.[64] The lowest energy path for the desired rotation was identified as a single segment with a barrier of 2.7 kcal/mol and was used as Seg3 of C_α-C(O)OH *path* 2. *Path* 2 was symmetric and involved five segments. The barrier S5/TS to forming the D-alanine from the geminal diol was 31.3 kcal/mol. This transition state was 52.3 kcal/mol above the lowest energy conformation of alanine.

3. The C_α-CH_3 path

It is worth noting that the C_α-CH_3 path, unlike the C_α-NH_2 and C_α-C(O)OH paths, lacked the ability to transfer the proton from C_α to the $-CH_3$ group because the latter could not accept additional protons. Instead, optimization of the path found hydrogen elimination that split alanine into H_2 and 2-amino-acrylic acid. These two molecules were 28.8 kcal/mol above the lowest conformation of alanine. Hence, this inversion path involved a bimolecular addition of H_2 on the other side of the prochiral acrylic plane. However, pathways involving H_2 molecule elimination and addition presented challenges on the B3LYP/6-31G(d,p) potential. In particular, unlike the restricted HF/6-31G(d,p) potential, the restricted B3LYP/6-31G(d,p) potential had a cusp at ∼102 kcal/mol instead of a transition state.

Based on the energy profiles of the analyzed paths, we concluded that C_α-C(O)OH *path* 1 and *path* 2 were more favorable than the C_α-NH_2 and C_α-CH_3 paths. Of the two C_α-C(O)OH paths, *path* 2, which bypassed oxiran-2-ol, appeared to be the most favorable. The highest transition state energy along C_α-C(O)OH *path* 2 was 52.3 kcal/mol above the lowest conformation of alanine identified in this work.

B. Acylation by peptide

The alanine inversion reaction demonstrated that the combined CG-HFB method could be used to study intramolecular reactions with rearrangements of covalent bonds. Nevertheless, the experience with the C_α-CH_3 path suggested that bimolecular reactions could be more difficult to model. The preferred relative orientation of two reacting molecules is well defined when poised for the attack at the transitions state but becomes less defined with increasing separation. In the end, two molecules reacting with each other and/or their products may separate considerably to minimize their interactions.

In this section, we studied a bimolecular reaction while testing the limits of the combined CG-HFB method. Specifically, we studied alcohol acylation by a peptide, as shown in Fig. 2. This reaction is biologically important as a model of peptide bond cleavage. When catalyzed by enzymes, the reaction proceeds through a tetrahedral intermediate.[55,56] However, here, we studied a single-step acylation in the gas phase where oxygen and hydrogen atoms of the alcohol added to carbon and nitrogen atoms of the peptide bond, respectively. Figure 2 shows the two bonds that broke and two bonds that formed in this reaction.

We studied the acylation reaction with the combined CG-HFB method on both QM and ONIOM QM/MM potential energy surfaces. In particular, we compared reaction paths for the "small" (methanol + formamide) and "large" (methanol + *N*-methylacetamide) systems shown in Fig. 2.

In Secs. III B 1–III B 2, we will first discuss the QM results on the small and large systems. In Sec. III B 3, we will discuss the ONIOM QM/MM results on the large system and compare them to the QM results. In the final Sec. III B 4, we discuss the relevance and transferability of the ONIOM QM/MM reaction path optimization with CG-HFB to studies of much larger systems, such as enzymes.

1. Small system QM results

The combined CG-HFB path optimization performed exceptionally well in locating acylation transition states for the small system. The reaction began with the OH bond of the alcohol aligning with the C–N bond of the amide. First, the proton from the OH group transferred to the amide nitrogen of the C–N bond. Next, the C–N bond broke and the O–C bond formed.

Overall, the reaction was 2.5 kcal/mol exothermic at the HF/6-31G(d,p) potential. The energy difference between the exact and extracted acylation transition states in the small system was only 0.02 kcal/mol, as shown in Table III. Structural comparison of the exact and extracted transition states for the small system yielded RMSDs of 0.059 and 0.008 Å, including all atoms and excluding the hydrogen atoms of the methyl groups, respectively.

TABLE III. Summary of optimized reaction paths for alcohol acylation by peptide mimics.[a]

Potential	System	E_{TS} (Hartree)	$\Delta E_{HFB\text{-}TS}$ (kcal/mol)	E_R (kcal/mol)	E_P (kcal/mol)	*Imag Freq*$_{TS}$ ($\Delta Freq_{HFB\text{-}TS}$) ($cm^{-1}$)	*TS-HFB RMSD*$_{All}$ (*RMSD*$_{\text{-}HMethyl}$) (Å)	$M(K)$
QM^1	Large	−364.203028	0.13	−39.9	−39.4	913 (−12)	0.089 (0.019)	65 (33)
QM^2	Large	−361.973508	0.01	−61.7	−61.7	701 (10)	0.024 (0.014)	193 (97)
QM^2	Small	−283.891124	0.02	−63.0	−65.5	1343 (−3)	0.059 (0.008)	193 (97)
QM^2/MM, R-Inc	Large	−283.894698	0.01	−63.8	−65.5	1332 (−1)	0.033 (0.007)	193 (97)
QM^2/MM, R-Exc	Large	−283.888147	0.01	−62.6	−65.5	1294 (3)	0.025 (0.005)	193 (97)
QM^2/MM, P-Inc	Large	−283.888020	0.03	−61.9	−65.5	1435 (5)	0.057 (0.015)	193 (97)
QM^2/MM, P-Exc	Large	−283.888244	0.06	−63.3	−65.3	1295 (59)	0.069 (0.009)	193 (97)

[a]E_{TS} and $\Delta E_{HFB\text{-}TS}$ values show the absolute energy of the exact TS and its difference with the CG-HFB TS, respectively. E_R and E_P values are the energies of the reactant and product relative to the exact TS, respectively. QM^1 and QM^2 are B3LYP and HF with the 6-31G(d,p) basis set, respectively. QM^2/MM is ONIOM[QM^2:AMBER]. ONIOM calculations were done using MM parameters of the reactant with the transforming bonds included in (R-Inc) and excluded (R-Exc) from the connectivity table and, similarly, for the product parameters (P-Inc and P-Exc). The MM partial charges of all the atoms were updated on the fly using the charge equilibration method. *TS-HFB RMSD*$_{All}$ values compare structures of the exact and CG-HFB TSs. *RMSD*$_{\text{-}HMethyl}$ excludes the hydrogen atoms of the methyl groups of the peptide model. The force constant (k) for all the paths was 1000 kcal/mol/Å^2. Forces were projected during optimization for all the potentials except QM^1. The *Imag Freq*$_{TS}$ column shows the exact imaginary frequency values for the TSs and their differences with CG-HFB values in parentheses. The M (K) column shows the total number of beads (M), with the number of Fourier terms in the series (K) used in the final optimized path in parentheses.

2. Large system QM results

Figure 9 shows the B3LYP/6-31G(d,p) energy profile for methanol acylation in the large system. The reaction mechanism was the same as in the small system. The barrier height for the forward reaction was 39.9 kcal/mol.

Full QM CG-HFB path optimization also gave accurate transition state structures for the large system. As shown in Table III, deviations from the exact transition state at the B3LYP/6-31G(d,p) potential was 0.089 and 0.019 Å, considering all atom RMSD and excluding methyl hydrogen atoms, respectively. Correspondingly, at the HF/6-31G(d,p) potential, the deviations were 0.024 and 0.014 Å RMSDs. CG-HFB overestimated the energy of the transition state by 0.1 kcal/mol for both QM potentials.

In contrast to the small system, the reaction was energy neutral at the HF/6-31G(d,p) potential and 0.4 kcal/mol endothermic at the B3LYP/6-31G(d,p) potential. The HF/6-31G(d,p) value of the imaginary frequency of the transition state in the large system was almost two times smaller than in the small system ($701i$ cm^{-1} vs $1{,}343i$ cm^{-1}), which is in line with the curvature of the peaks in the corresponding energy profiles (see supplementary material[64]). The B3LYP/6-31G(d,p) potential gave a value of $913i$ cm^{-1} for the imaginary frequency in the large system.

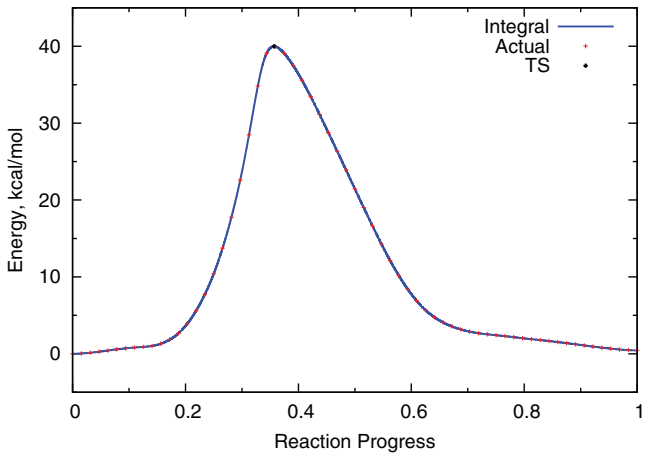

FIG. 9. B3LYP/6-31G(d,p) energy profile for alcohol acylation by peptide mimic in large system. The integral energy profile is shown in solid blue. The actual energy points offset to match the integral profile are shown as red plus signs. The TS is marked by a black "+".

3. Large system ONIOM QM/MM results

Table III shows the energetics and RMSDs of ONIOM optimized acylation reaction paths. Depending on the presence of the bonded terms associated with covalent bonds undergoing rearrangement and on which atom types and vdW parameters were used, the ONIOM transition state energy varied by, at most, 1.8 kcal/mol. With all ONIOM schemes, the reaction was exothermic by 1.3–3.6 kcal/mol, which encompasses the 2.5 kcal/mol value in the small QM case. The imaginary frequencies at the ONIOM transition states computed with all bonded terms were higher than those where the MM potential excluded the bonded terms of the affected covalent bonds. Also, the computed imaginary frequencies at the ONIOM transition states were closer to that of the small model QM result.

4. Transferability to ONIOM QM/MM studies of large systems

a. Rugged potential energy landscape. In general, large systems such as enzymes have many "soft" degrees of freedom. Consequently, their multidimensional potential energy surfaces are extremely rugged and present serious challenges for optimization.

For the chain of states path optimization methods, the ruggedness of the QM/MM potential energy surfaces can cause the beads to diverge in their soft degrees of freedom during the optimization, rendering the corresponding energy profiles discontinuous. This issue was addressed by imposing positional restraints on the soft degrees of freedom.[57] During the path optimization, the force constants of the restraints were gradually reduced while keeping their anchoring positions unchanged.

The combined CG-HFB method provides an alternative solution to the problem associated with the ruggedness of

potential energy landscapes. As was mentioned earlier, the adaptive positional restraints in CG-HFB transform the rugged QM/MM potential energy surface into a funnel-like surface that is simpler to optimize.[36,37] Thus, CG-HFB updates the anchoring positions but not the force constants of the restraints during the path optimization. With sufficiently stiff restraints CG-HFB completely prevents discontinuities in the soft degrees of freedom along the path.

Although the "large" system in the present study was much smaller than that in a typical QM/MM study, it did contain soft degrees of freedom that are characteristic of the latter and could cause problems during optimization. In particular, rotations of the three methyl groups shown in Fig. 2 could disrupt continuity of the path and of the energy profile.[57,62,63] The combined CG-HFB method performed well with respect to the soft degrees of freedom. The only problem that we found due to the soft degrees of freedom was the occasional excursions of the path to sample the methyl rotations in the early stages of path optimization (see supplementary material[64]). However, the segments involving the methyl rotations could be trivially discarded producing clean energy profiles like the one depicted in Fig. 9. Therefore, we anticipate the combined CG-HFB method will provide high quality QM/MM reaction paths and potential energy profiles for large systems.

b. Conformational sampling. Free energy profiles or potentials of mean force (PMFs) along the path are preferred to the potential energy profiles because they can be directly compared with experiment.[58–61] This is particularly true of the large systems with rugged energy landscapes. However, computing a valid reaction PMF requires conformational sampling and a proper choice of reaction coordinate. A poor choice of reaction coordinate will likely yield an invalid PMF. Although the HFB method was developed to compute PMFs in large systems, it is impractical to compute QM/MM PMFs directly.

Indirect, more affordable approaches have been developed to estimate QM/MM PMFs.[58–61] Importantly, the indirect QM/MM PMF calculations critically depend on reaction paths optimized on QM/MM potential energy surfaces.[59–61] For a given configuration of the system, the paths optimized with the combined CG-HFB method should give the best estimate of the reaction barrier and identify the best reaction coordinate. Subsequently, these could be used to obtain a valid QM/MM PMF. Therefore, the reaction paths optimized on QM/MM potential energy surfaces are prerequisite to successful QM/MM PMF calculations.

IV. CONCLUSIONS

In this report we applied a combined CG-HFB reaction path optimizations method to study chemical reactions on QM and QM/MM potential energy surfaces. With the help of CG-HFB, we were able to identify nontrivial, intramolecular paths for spontaneous alanine inversion in the gas phase starting from three crude intermediate states. The most favorable of the paths proceeded through previously unknown intermediates, namely, 2-iminopropane-1,1-diol and 3-amino-3-methyloxiran-2-ol. Furthermore, with CG-HFB, we were able to optimize bimolecular reaction paths for acylation of an alcohol by peptide models despite the poorly defined relative orientation of the reacting molecules and their products.

In all cases, CG-HFB delivered high-quality transition states that were in excellent agreement with the exact transition states. In the worst case, CG-HFB overestimated a transition state energy by 0.1 kcal/mol. All identified transition states were automatically validated by CG-HFB instead of using IRC calculations. The results demonstrate that the combined CG-HFB approach is a practical tool for finding accurate transition states with little prior knowledge of reaction mechanisms.

We anticipate that the combined CG-HFB method will be directly applicable to QM/MM studies of chemical reactions in much larger systems, such as enzymes. Our choices of CG instead of quasi-Newton optimizer and of the Cartesian instead of nonlinear coordinates were made keeping these large systems in mind. Future work will be concerned with applications of the combined CG-HFB method to QM/MM studies of enzyme catalysis and inhibition.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Defense High Performance Computing Modernization Program under the High Performance Computing Software Applications Institutes initiative. Computational time was provided by the U.S. Army Research Laboratory Department of Defense Supercomputing Resource Center. This work was sponsored by the Defense Threat Reduction Agency Grant No. CBM.THERB.02.11.RD.012. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

[1] D. G. Truhlar, B. C. Garrett, and S. J. Klippenstein, J. Phys. Chem. **100**(31), 12771 (1996).
[2] G. Henkelman, B. P. Uberuaga, and H. Jonsson, J. Chem. Phys. **113**(22), 9901 (2000).
[3] P. Y. Ayala and H. B. Schlegel, J. Chem. Phys. **107**(2), 375 (1997).
[4] A. Behn, P. M. Zimmerman, A. T. Bell, and M. Head-Gordon, J. Chem. Phys. **135**(22), 224108 (2011).
[5] A. Komornicki, K. Ishida, and K. Morokuma, Chem. Phys. Lett. **45**(3), 595 (1977).
[6] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**(6), 2800 (1981).
[7] H. B. Schlegel, J. Comput. Chem. **3**(2), 214 (1982).
[8] S. Bell and J. S. Crighton, J. Chem. Phys. **80**(6), 2464 (1984).
[9] C. Y. Peng, P. Y. Ayala, and H. B. Schlegel, J. Comput. Chem. **17**(1), 49 (1996).
[10] A. Behn, P. M. Zimmerman, A. T. Bell, and M. Head-Gordon, J. Chem. Theory Comput. **7**(12), 4019 (2011).
[11] K. Fukui, J. Phys. Chem. **74**(23), 4161 (1970).
[12] K. Fukui, Acc. Chem. Res. **14**(12), 363 (1981).
[13] K. Fukui, S. Kato, and H. Fujimoto, J. Am. Chem. Soc. **97**(1), 1 (1975).
[14] E. Weinan, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[15] W. E, W. Ren, and E. Vanden-Eijnden, J. Chem. Phys. **126**(16), 164103 (2007).
[16] A. Goodrow, A. T. Bell, and M. Head-Gordon, Chem. Phys. Lett. **484**(4-6), 392 (2010).

[17]S. Mallikarjun Sharada, P. M. Zimmerman, A. T. Bell, and M. Head-Gordon, J. Chem. Theory Comput. **8**(12), 5166 (2012).
[18]P. Tao, M. Hodoscek, J. D. Larkin, Y. H. Shao, and B. R. Brooks, J. Chem. Theory Comput. **8**(12), 5035 (2012).
[19]B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, J. Chem. Phys. **120**(17), 7877 (2004).
[20]S. Maeda, E. Abe, M. Hatanaka, T. Taketsugu, and K. Morokuma, J. Chem. Theory Comput. **8**(12), 5058 (2012).
[21]S. Maeda and K. Morokuma, J. Chem. Theory Comput. **7**(8), 2335 (2011).
[22]S. Maeda and K. Morokuma, J. Chem. Theory Comput. **8**(2), 380 (2012).
[23]G. Henkelman and H. Jonsson, J. Chem. Phys. **113**(22), 9978 (2000).
[24]E. F. Koslover and D. J. Wales, J. Chem. Phys. **127**(13), 134102 (2007).
[25]I. V. Khavrutskii, K. Arora, and C. L. Brooks III, J. Chem. Phys. **125**(17), 174108 (2006).
[26]I. V. Khavrutskii, M. Fajer, and J. A. McCammon, J. Chem. Theory Comput. **4**(9), 1541 (2008).
[27]I. V. Khavrutskii and J. A. McCammon, J. Chem. Phys. **127**(12), 124901 (2007).
[28]L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, J. Chem. Phys. **125**(2), 024106 (2006).
[29]A. Goodrow, A. T. Bell, and M. Head-Gordon, J. Chem. Phys. **129**(17), 174109 (2008).
[30]A. E. Cho, J. D. Doll, and D. L. Freeman, Chem. Phys. Lett. **229**(3), 218 (1994).
[31]A. Matro, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **101**(12), 10458 (1994).
[32]I. V. Khavrutskii, J. Dzubiella, and J. A. McCammon, J. Chem. Phys. **128**(4), 044106 (2008).
[33]I. V. Khavrutskii, A. A. Gorfe, B. Lu, and J. A. McCammon, J. Am. Chem. Soc. **131**(5), 1706 (2009).
[34]I. V. Khavrutskii, B. Grant, S. S. Taylor, and J. A. McCammon, Biochemistry **48**(48), 11532 (2009).
[35]R. Fletcher and C. M. Reeves, Comput. J. **7**(2), 149 (1964).
[36]K. A. Dill and H. S. Chan, Nat. Struct. Mol. Biol. **4**(1), 10 (1997).
[37]P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**(5204), 1619 (1995).
[38]C. C. J. Roothaan, Rev. Mod. Phys. **23**(2), 69 (1951).
[39]A. D. Becke, Phys. Rev. A **38**(6), 3098 (1988).
[40]C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**(2), 785 (1988).
[41]A. D. Becke, J. Chem. Phys. **98**(7), 5648 (1993).
[42]P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**(3), 213 (1973).
[43]I. V. Khavrutskii, R. H. Byrd, and C. L. Brooks III, J. Chem. Phys. **124**(19), 194903 (2006).
[44]R. Elber and M. Karplus, Chem. Phys. Lett. **139**(5), 375 (1987).
[45]L. Maragliano and E. Vanden-Eijnden, Chem. Phys. Lett. **446**(1-3), 182 (2007).
[46]E. Weinan, W. Ren, and E. Vanden-Eijnden, J. Phys. Chem. B **109**(14), 6688 (2005).
[47]M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 09, Gaussian, Inc., Wallingford, CT, 2009.
[48]T. Vreven, K. S. Byun, I. Komaromi, S. Dapprich, J. A. Montgomery, K. Morokuma, and M. J. Frisch, J. Chem. Theory Comput. **2**(3), 815 (2006).
[49]A. K. Rappe and W. A. Goddard III, J. Phys. Chem. **95**(8), 3358 (1991).
[50]A. G. Csaszar, J. Phys. Chem. **100**(9), 3541 (1996).
[51]S. Gronert and R. A. J. O'Hair, J. Am. Chem. Soc. **117**(7), 2071 (1995).
[52]K. M. Westerberg and C. A. Floudas, J. Chem. Phys. **110**(18), 9259 (1999).
[53]S. Blanco, A. Lesarri, J. C. Lopez, and J. L. Alonso, J. Am. Chem. Soc. **126**(37), 11675 (2004).
[54]S. G. Stepanian, I. D. Reva, E. D. Radchenko, and L. Adamowicz, J. Phys. Chem. A **102**(24), 4623 (1998).
[55]M. Perakyla and J. Rouvinen, Chem. Eur. J. **2**(12), 1548 (1996).
[56]M. Perakyla and P. A. Kollman, J. Am. Chem. Soc. **119**(6), 1189 (1997).
[57]L. Xie, H. Liu, and W. Yang, J. Chem. Phys. **120**(17), 8039 (2004).
[58]H. Hu, Z. Lu, J. M. Parks, S. K. Burger, and W. Yang, J. Chem. Phys. **128**(3), 034105 (2008).
[59]M. Valiev, E. J. Bylaska, M. Dupuis, and P. G. Tratnyek, J. Phys. Chem. A **112**(12), 2713 (2008).
[60]M. Valiev, B. C. Garrett, M.-K. Tsai, K. Kowalski, S. M. Kathmann, G. K. Schenter, and M. Dupuis, J. Chem. Phys. **127**(5), 051102 (2007).
[61]T. H. Rod and U. Ryde, Phys. Rev. Lett. **94**(13), 138302 (2005).
[62]G. A. Cisneros, H. Liu, Z. Lu, and W. Yang, J. Chem. Phys. **122**(11), 114502 (2005).
[63]H. Liu, Z. Lu, G. A. Cisneros, and W. Yang, J. Chem. Phys. **121**(2), 697 (2004).
[64]See supplementary material http://dx.doi.org/10.1063/1.4826470 for the details of the combined CG-HFB method, its latest stand-alone implementation, and practical aspects of applying the method to study reaction mechanisms. The results of CG-HFB path optimizations for alanine inversion on the HF/6-31G(d,p) potential are provided. In addition, the energy profile for the full 360° rotation of the geminal diol $-C(OH)_2H$ group at the B3LYP/6-31G(d,p) potential is provided. HF/6-31G(d,p) energy profiles for methanol acylation in both large and small systems are also provided.