JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Comparison of Three Chain-of-States Methods: Nudged Elastic Band and Replica Path with Restraints or Constraints

Peng Tao,*[,†] Milan Hodošček,[‡] Joseph D. Larkin,[†] Yihan Shao,[†,§] and Bernard R. Brooks[†]

[†]Laboratory of Computational Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

[‡]Center for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia

[§]Q-Chem Inc., 5001 Baum Boulevard, Suite 690, Pittsburgh, Pennsylvania 15213, United States

Supporting Information

ABSTRACT: Chain-of-state methods are becoming important tools in studying the chemical reaction mechanisms, especially for biomacromolecules. In this article, three chain-of-state methods, the nudged elastic band (NEB) method and the replica path method with restraints or constraints, were tested and compared using three model systems with various sizes and at different levels of theory: alanine dipeptide isomerization, β-alanine intramolecular condensation, and the matrix metalloproteinase 2 inhibition mechanism. The levels of theory used to describe the three model systems include molecular mechanics (MM), quantum mechanics (QM), and combined quantum mechanics and molecular mechanics (QM/MM). All three methods could correctly determine a reaction path with reasonable estimation of reaction barriers in most cases. The RMSD measurement with additional weighting schemes provides practically infinite choices of reaction coordinates to describe the reaction progress. These findings demonstrate that the chain-of-state methods are powerful tools when being used carefully to generate a plausible reaction mechanism with full pathway for complex systems at an affordable computational cost.

1. INTRODUCTION

Reaction mechanisms are important concepts in the understanding of the transition from reactants to products in chemistry.[1−4] In computational chemistry, the reaction mechanism can be represented as a reaction pathway on the potential energy surface (PES) of the system of interest through construction of a potential energy function of the nuclear coordinates.[5−9] A reaction pathway can be identified as the minimum energy path (MEP) connecting two local minima through one or more first-order saddle points that correspond to transition states (TS) on the PES. The MEP can be calculated in various methods. Walking on the PES either downhill from TS through the steepest descent pathway toward energy minima,[10−16] or uphill from energy minimum toward TS,[17−19] can produce the MEP. However, the need of TS *a priori* or *posteriori* makes the construction of MEP a difficult task for many systems.

To bypass the calculation of TS before the construction of MEP, there has been rapid development of so-called chain-of-states methods, in which a number of images (i.e., states) of a system are used to connect two end points and are subject to minimization simultaneously. Restraints or constraints between images are added to maintain the distance between adjacent images to ensure the even distribution of the target reaction path. Some chain-of-state methods include the following: conjugate peak refinement,[20] nudged elastic band (NEB),[21−28] replica path (RPATH),[29−32] line-integral,[33−37] combined reaction path and stationary structures optimization,[38] zero temperature string (ZTS) methods,[39−42] finite temperature string (FTS) methods,[43] quadratic string method,[44] and growing string methods.[45−50]

In the NEB method,[21] the images are held together by harmonic spring forces. The orthogonal forces are projected out and do not affect the minimization of each image. Therefore, the NEB calculations can principally produce the MEP when fully converged. Chu et al.[51] developed the first superlinear minimizer for the NEB method. Their development was based on expanding the adopted basis Newton-Raphson (ABNR) method[52] and is available in the CHARMM program suite.[53]

Restraints have been implemented within the replica path (RPATH) method in CHARMM.[31] Spring forces are introduced to harmonically restrain distances between adjacent images along the reaction pathway. Additional forces can also be added to maintain path smoothness. The study of the chorismate mutase mechanism using this approach demonstrated that this method can be effectively applied on macromolecules.[31] Holonomic constraints have also been implemented within the replica path facility in CHARMM.[32] The distances between adjacent images are maintained equal to each other up to convergence for each round of optimization. This can provide an even distribution of images to better represent the reaction process in the TS region.

It should be noted that the MEP is not sufficient to determine a mechanism in that entropic effects are ignored. Free energy of barrier crossing can be calculated or estimated by a variety of techniques starting from an MEP. The focus of this paper is the determination of MEP or approximate MEPs

Special Issue: Berny Schlegel Festschrift

Received: July 20, 2012
Published: September 27, 2012

that are suitable for further investigation using free energy simulation techniques or entropy estimation such as those involving harmonic analysis approaches.[54]

Unlike single geometry optimization strategies, the chain-of-states methods have not been widely applied, especially in QM/MM enzymatic mechanism studies. A benchmark study with multiple test cases other than the original development work is necessary to demonstrate both the strengths and weaknesses of available methods and to promote their application and development. In this study, three chain-based optimization methods from CHARMM——NEB and RPATH with restraint and constraint——are tested using three reactions as test cases: alanine dipeptide isomerization, β-alanine (3-aminopropanoic acid) intramolecular condensation, and the matrix metalloproteinase 2 (MMP2) inhibition mechanism. The alanine dipeptide isomerization is a typical test case in computational methodology developments.[55−60] The β-lactam, the intramolecular condensation product of β-alanine, is part of the basic structure of widely used β-lactam antibiotics.[61,62] The MMP2 is a proteolytic enzyme that digests type IV collagens.[63] The structure and the catalytic mechanisms of MMP2 were under comprehensive studies.[64−69] All three systems together provide adequate assessment and discrimination of the efficiency of the methods under this study.

It has been shown that the reweighting of the atoms involved in the reaction path was crucial to obtain a reaction path in NEB calculations.[27] In addition to mass-weighting scheme, user defined weighting schemes are also applied in some of the path calculations in this study for better results.

2. METHOD AND MATERIALS

2.1. NEB Method. The superlinear NEB minimizer[51] is implemented on the basis of ABNR method[52] in CHARMM.[53] The forces on each replica in NEB framework are projected using a tangent vector (τ_i) along the path

$$\begin{aligned}\mathbf{F}_i &= \mathbf{F}_i^{\perp} + \mathbf{F}_i^{\parallel}\\ \mathbf{F}_i^{\perp} &= -\nabla V(\mathbf{r}_i)\cdot(1-\tau_i\tau_i)\\ \mathbf{F}_i^{\parallel} &= -\nabla_i\left(\frac{1}{2}k\sum_{j=1}^{N}(\Delta l^i-\overline{\Delta l})^2\right)\cdot(\tau_i\tau_i)\end{aligned} \tag{1}$$

where $\mathbf{F}_i^{\perp}$ and $\mathbf{F}_i^{\parallel}$ are force components perpendicular (off-path direction) and parallel (elastic band with force k) to the tangent vector of replica i (τ_i), respectively. V is the potential energy function, and Δl is the distance between adjacent replicas. To improve the computational efficiency, both steepest-descent (SD) and Newton−Raphson (NR) in ABNR are extended to the NEB framework.

Various choices of tangent vector have been proposed and could lead to a different behavior of NEB calculations in terms of computational efficiency and smoothness of calculated reaction path.[22,25,26,28] The tangent vector used in the NEB method implemented in CHARMM is defined as

$$\tau_i = \mathrm{NORM}[\mathbf{w}\cdot(\mathbf{r}_{i+1,j}^{\mathrm{RMS}\to i}-\mathbf{r}_{i-1,j}^{\mathrm{RMS}\to i})] \tag{2}$$

where NORM is the normalization operator, $\mathbf{w}$ is the weighting vector, and RMS→i superscript indicates that the neighboring replica is best fitted to replica i.[51]

2.2. RPATH with Restraints. To optimize the reaction path represented by a series of replicas and their environment, penalty functions are needed to maintain the distance between adjacent replicas. In a restraint framework,[31] each replica is restrained using best-fit root-mean-square distances (RMSD) to the adjacent replicas. The RMS restraint forces are defined by the following equation

$$E_{rms} = \sum_{i=1}^{N}\frac{1}{2}k_{rms}\omega_i(r_i-\overline{r})^2 \tag{3}$$

where N is the number of replicas, k_{rms} is the force constant used to restrain distances between adjacent replicas along the reaction pathway, r_i is the best-fit RMSD between replica i and i+1, and $\overline{r}$ is the average distance between adjacent replicas. An atomic weight factor ω_i is used to select atoms and to determine their strength in the fitting procedure.

An additional force can be added to restrain the angle between replicas along the pathway through an angle energetic penalty term

$$\begin{aligned}E_{angle} &= \sum_{i=1}^{N}\frac{1}{2}k_{ang}(\mathrm{COSMAX}-\cos(\theta)_i)^2,\\ &\quad \text{if COSMAX} > \cos(\theta)_i\\ E_{angle} &= 0, \text{if COSMAX} \le \cos(\theta)_i\end{aligned} \tag{4}$$

The angle θ, illustrated in Figure 1, defines the deviation of the pathway from linearity. The force constant k_{ang} controls the rigidity of the pathway. The constant, COSMAX, determines the value of $\cos(\theta)$ subject to the angle forces. Angle term forces are converted to best-fit RMSD radial forces using the definition of cosines. Best-fit RMSD forces are computed analytically.[51]

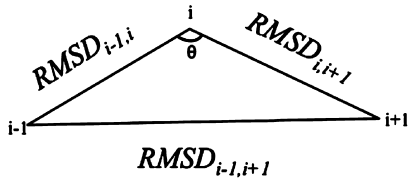

Figure 1. Illustration of angle θ for replica i in RPATH calculation. $\mathrm{RMSD}_{i\text{-}1,i}$ is the distance between replica i-1 and i. It is similar to $\mathrm{RMSD}_{i,i+1}$ and $\mathrm{RMSD}_{i\text{-}1,i+1}$.

2.3. RPATH with Constraint. Recently, the equal distance holonomic constraint[32] method has been implemented in the RPATH framework in CHARMM. Given two states of a molecular system with N atoms, $\mathbf{r}^0$ and $\mathbf{r}^k$, a chain of K+1 replicas can be constructed to connect these two states. The distance between each pair of adjacent replicas is set to be equal to each other

$$\Delta l^0 = \bullet\bullet\bullet = \Delta l^i = \bullet\bullet\bullet = \Delta l^{K-1} = \overline{\Delta l} \tag{5}$$

Here, Δl^i is the distance between replica i and i+1 and can be in any form, including best-fit RMS distance. $\overline{\Delta l}$ is the average distance between adjacent replicas. The following scheme is used to propagate the reaction path, which satisfies eq 5.

(i) Set up and calculate initial average distance, $\overline{\Delta l}$, for replicas $r^{0(0)}$ through $r^{k(0)}$. The superscript "(0)" indicates the optimization iteration step.

(ii) To maintain the equal distance, a set of K coefficients, $(\lambda^i)^{(n)}(i = 0,K-1)$, are used to update the coordinates of each replica i:

$$(\mathbf{r}^i)^{(n+1)} = (\mathbf{r}^i)^{(n)} + (\lambda^{i-1})^{(n)}\left(\frac{\partial \Delta l^{i-1}}{\partial \mathbf{r}^i}\right)^{(n)} + (\lambda^i)^{(n)}\left(\frac{\partial \Delta l^i}{\partial \mathbf{r}^i}\right)^{(n)} \quad (6)$$

(iii) Solve $(\lambda^i)^{(n)}$ by setting the first-order Taylor expansion of each of $((\lambda^i)^{(n)} - \overline{\Delta l})$ $(i = 0,K-1)$ with respect to $(\lambda^j)^{(n)}$ to zero:

$$-((\Delta l^i)^{(n)} - \overline{\Delta l}) = \sum_{j=i-1}^{i+1}\left(\frac{\partial \Delta l^i}{\partial \lambda^j}\right)_{\lambda^j}^{(n)} (\lambda^j)^{(n)} \quad (7)$$

(iv) If any of the values of $|(\Delta l^i)^{(n+1)} - \overline{\Delta l}|(i = 0,K-1)$ is greater than a selected tolerance, then repeat steps (ii) and (iii).

(v) After convergence, the RPATH calculation leads to a reaction path composed by $K+1$ equal distance replicas connecting states $\mathbf{r}^0$ and $\mathbf{r}^k$.

When using constraints with RPATH, a kinetic energy potential can be added to the potential energy making the overall objective function to be minimized, a Hamiltonian.[32] Therefore the optimized path is a so-called minimum Hamiltonian path (MHP) instead of an MEP. The kinetic energy component in the potential helps to prevent kinks and therefore helps to maintain the smoothness of the path. However, this smoothness comes at the cost of deviation from the MEP, resulting in higher reaction barriers.

2.4. Replication Schemes. Two replication schemes are available in CHARMM for chain-of-states calculations. In one scheme, all the replicas are contained in a single CHARMM protein structure file (PSF). Either the full system or specific "important" parts (e.g., active site of a catalytic enzyme) can be chosen and replicated. In the other scheme, the parallel distributed replicas (REPD) framework, a series of replicas with independent setup, are generated and run on different processors. Each replica has its own setup including PSF file, which contains the complete system information. With REPD, users have more flexibility to treat each replica differently with separate PSFs without affecting other replicas.

3. RESULTS

3.1. Isomerization of Alanine Dipeptide. The first test case is the isomerization of the alanine dipeptide (N-acetylalanyl-N-methyl-amide). Two backbone dihedral angles (ϕ and ψ) are used to describe the isomerization process of this molecule (Figure 2). The reaction path in the current study connects two conformers C_{7eq} and C_{ax} corresponding to two minima on the PES (Figure 3). The CHARMM 22 force field[70] with CMAP backbone dihedral angle corrections[71] was used for the calculation. No solvent molecules were present in the model system. All the RPATH calculations are carried out using 25 replicas. The initial guess of reaction path for optimization was constructed through linear interpolation (see Figure S1 in the Supporting Information). The TS structure of isomerization was optimized in CHARMM with the barrier as 8.74 kcal/mol with reference to conformer C_{ax}.

3.1.1. NEB Results. Six NEB calculations of alanine dipeptide were carried out as follows: In calculations 1−3, all atoms are used to calculate the mass-weighted RMSD between each adjacent replicas with a spring constant as 10, 100, and 1000 kcal•mol^{-1}•Å^{-2}, respectively. Calculations 4−6 repeat calculations 1−3 but have hydrogen atoms excluded from mass-weighted RMSD calculation. All six calculations lead to almost identical reaction pathways, which are presented on a contour plot using ϕ and ψ as reaction coordinates (Figure 3). This observation demonstrates that the MEP obtained from NEB does not depend on the value of the spring constant k. The reaction barriers for isomerization with reference to C_{7eq} conformer obtained from these six calculations are within a very narrow range, which is between 8.74 and 8.81 kcal/mol (Table 1). For all six calculations, the ninth replica starting from the C_{ax} conformer represents an approximate transition state (TS) of the isomerization. From this point forward, the C_{ax} conformer always serves as the first replica for the replica numbering.

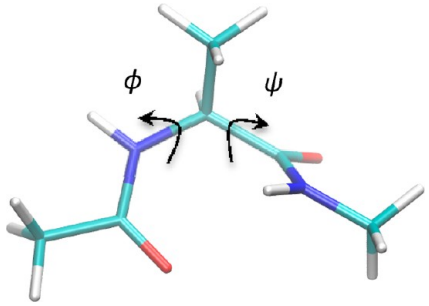

Figure 2. Structure of alanine dipeptide and two dihedral angles as reaction coordinate of isomerization. H, C, N, and O shown in white, cyan, blue, and red, respectively.

Both RMS forces perpendicular (off-path) and parallel to the tangent vector of the reaction pathway are plotted for calculations 1−6 (Figures 4 and 5). The RMS forces fluctuate during optimization due to the fact that the tangent vector at each replica used for force projection is defined by a discrete reaction path and changes from one optimization step to the next. The plots of tangent forces in Figure 4 start at different levels for the calculations with different spring force constants but converge toward zero, showing that a MEP can be obtained independently from spring constants (Figure 4). The off-path RMS forces are independent from added spring forces, resulting from all six plots starting at the same level (Figure 5). The RMS forces decrease smoothly but slowly for about ten steps initially before decreasing more rapidly. This is due to the ABNR optimizer implemented in CHARMM being a combination of SD and NR methods with SD dominating the initial steps of the optimization. The convergence rates are not significantly different in these NEB calculations, except for calculation 5 with the force constant set to 100 kcal•mol^{-1}•Å^{-2} and hydrogens excluded from the RMSD measurement.

3.1.2. RPATH/Restraint Results. Six RPATH/restraint calculations of alanine dipeptide were carried out as the following. In calculations 1−3, all atoms are used to calculate the mass-weighted RMSD between each adjacent replica with a spring constant as 1000, 10000, and 100000 kcal•mol^{-1}•Å^{-2}, respectively. Calculations 4−6 repeat calculations 1−3 but have hydrogens excluded from the mass-weighted RMSD calculation. Larger spring constants are used in this setup, because smaller forces are not sufficient to maintain the close distance between adjacent replicas around a TS region (see Figure S2 in the Supporting Information). All six calculations have k_{ang} as 100 kcal•mol^{-1}•Å^{-2}, and COSMAX = 0.95. The deviation

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

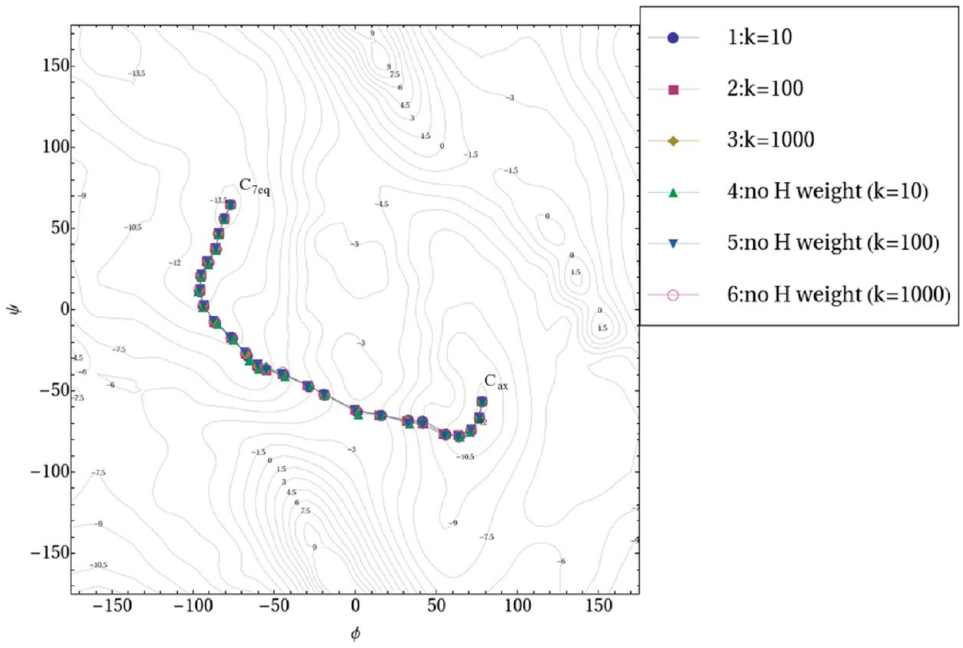

Figure 3. Reaction pathways of alanine dipeptide isomerization using the NEB method. In pathways 1−3, the elastic force constant k is 10, 100, and 1000 kcal•mol^{-1}•Å^{-2}, respectively. The mass-weighted RMSD of all the atoms is used to measure the distance between replicas. Pathways 4−6 repeat the calculations of 1−3 with hydrogens excluded from mass-weighted RMSD calculation.

Table 1. Calculations for Alanine Dipeptide Isomerization

methods	parameters[a]		barrier (kcal/mol)	replica ID of approximate TS
NEB	1 k = 10		8.78	9
	2 k = 10^2		8.81	9
	3 k = 10^3		8.81	9
	4 k = 10, no H weight		8.74	9
	5 k = 10^2, no H weight		8.81	9
	6 k = 10^3, no H weight		8.78	9
restraint	k_{ang} = 10^2, COSMAX = 0.95	1 k_{rms} = 10^3	8.52	8
		2 k_{rms} = 10^4	8.64	8
		3 k_{rms} = 10^5	8.63	7
		4 k_{rms} = 10^3, no H weight	8.54	8
		5 k_{rms} = 10^4, no H weight	8.77	8
		6 k_{rms} = 10^5, no H weight	8.76	8
	k_{rms} = 10^5, no H weight COSMAX=1.00	1 k_{ang} = 10^2	8.64	8
		2 k_{ang} = 10^3	8.78	9
		3 k_{ang} = 10^4	8.87	10
		4 k_{ang} = 2×10^4	10.37	10
constraint	1 k_{pki}= 0, all atom mass-weighted		8.52	8
	2 k_{pki}= 0, no hydrogen weight		8.73	9
	3 k_{pki} = 100, no hydrogen weight		9.63	15
	4 k_{pki} = 10, no hydrogen weight		9.59	11
	5 k_{pki} = 1, no hydrogen weight		8.78	9

[a]Unit for force constant is kcal•mol^{-1}•Å^{-2}.

among six pathways is noticeable around the region connecting C_{7eq} and the TS (Figure 6). The barriers with reference to the C_{7eq} conformer range narrowly from 8.52 to 8.77 kcal/mol for all six pathways (Table 1), with barrier from pathway 6 (8.76 kcal/mol) as the closest to the one from the TS calculation in CHARMM (8.74 kcal/mol). Because no force projection is involved in RPATH with restraint calculations, the RMS forces actually converge toward zero for all six calculations (Figure 7). Calculations 4 and 5 have the similar and the fastest convergence rates, while calculations 3 and 6 with largest k_{rms} display the slower convergence rates than other calculations.

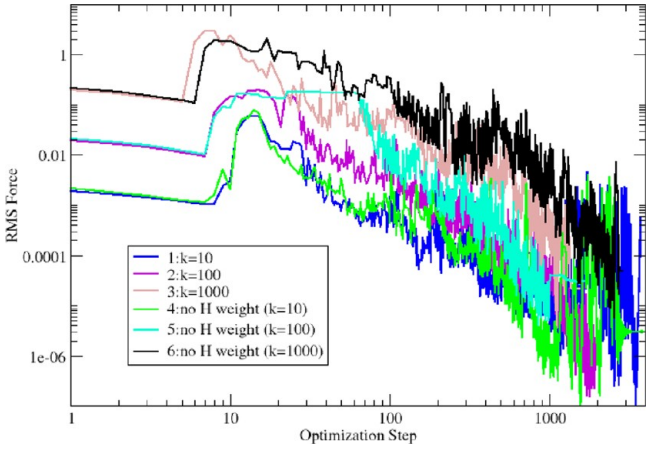

Figure 4. The tangent RMS force (kcal•mol^{-1}•Å^{-1}) in NEB optimization of alanine dipeptide isomerization.

Different k_{ang} values were applied with k_{rms} = 100000 kcal•mol^{-1}•Å^{-2} with hydrogen atoms excluding from mass-weighted RMSD measurement (Figure 8). The COSMAX value was set to 1.00 for these calculations to increase the smoothness of the pathways further. With larger k_{ang} values, the pathways become more rigid and smoother and deviate significantly from the MEP. Interestingly, pathways 2 and 3 in Figure 8 with k_{ang} as 1000 and 10000 kcal•mol^{-1}•Å^{-2}, respectively, still go through the TS region with barriers rather close to the one from TS calculation.

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

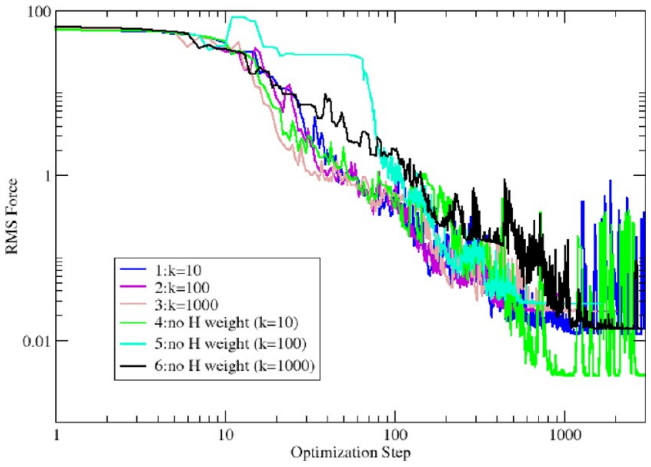

Figure 5. The off-path RMS force (kcal•mol^{-1}•Å^{-1}) in NEB optimization of alanine dipeptide isomerization.

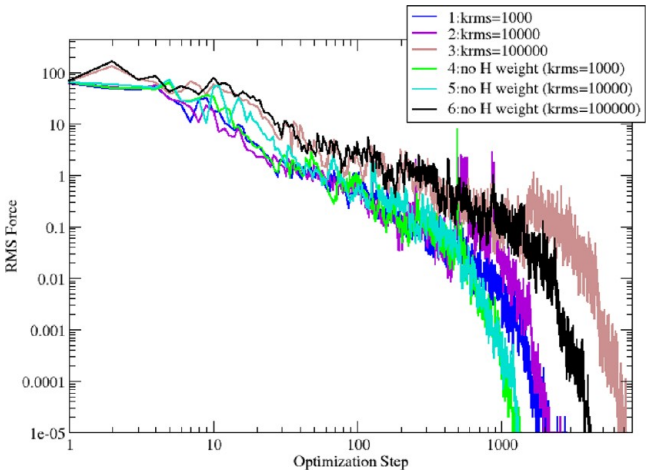

Figure 7. The RMS force (kcal•mol^{-1}•Å^{-1}) in the RPATH/restraint optimization of alanine dipeptide isomerization.

Using k_{ang} values less than 10 kcal•mol^{-1}•Å^{-2} leads to unusable pathways that tend to "hover" in a minimal basin (see Figure S3 in the Supporting Information). In this example, there is a two order-of-magnitude range of values for k_{ang} that provides a pathway that is very similar to the MEP. The optimal value will vary from system to system and also will depend on the number of replicas. Thus some care and preliminary investigation is required to use this method well.

3.1.3. RPATH/Constraint Results. Five RPATH/constraint calculations of alanine dipeptide were carried out and plotted in Figure 9. The reaction barrier obtained from calculation 1, with hydrogen included in the mass-weighted RMSD measurement, is 8.52 kcal/mol for replica 8. In calculation 2, with hydrogen excluded from RMSD measurement, replica 9 has a reaction barrier of 8.73 kcal/mol. The force constants of the kinetic energy potential (k_{pki}) as 100, 10, and 1 kcal•mol^{-1}•Å^{-2} were applied in another three calculations, respectively. For calculation 3 with largest k_{pki}, the MHP as the reaction pathway resembles the straight-line interpolation between two end points. When reducing the kinetic energy force constant by an order of magnitude, the corresponding MHP is roughly in the middle between the MEP and the straight line connecting the two end points. The corresponding MHP for the smallest k_{pki} closely resembles the MEP with correct reaction barrier (8.78 kcal/mol) given by replica 9.

The RMS forces along the pathways during the path optimization are plotted in Figure 10. Calculation 1 has difficulty to converge (1 in Figure 10). This is due to the sensitivity of holonomic constraint iterations to the rotation of methyl groups. After excluding hydrogen from the RMSD measurement, the RPATH calculation converges rapidly (2 in Figure 10). When including kinetic energy components, the convergence rate accelerates further.

3.2. β-Alanine Intramolecular Condensation. The intramolecular condensation of β-alanine is a one-step reaction (Figure 11) with a barrier of 56.79 kcal/mol at the B3PW91/6-31g(d,p) level of theory[72−74] calculated in Q-Chem.[75] It should be noted that this barrier is based on the internal energy from

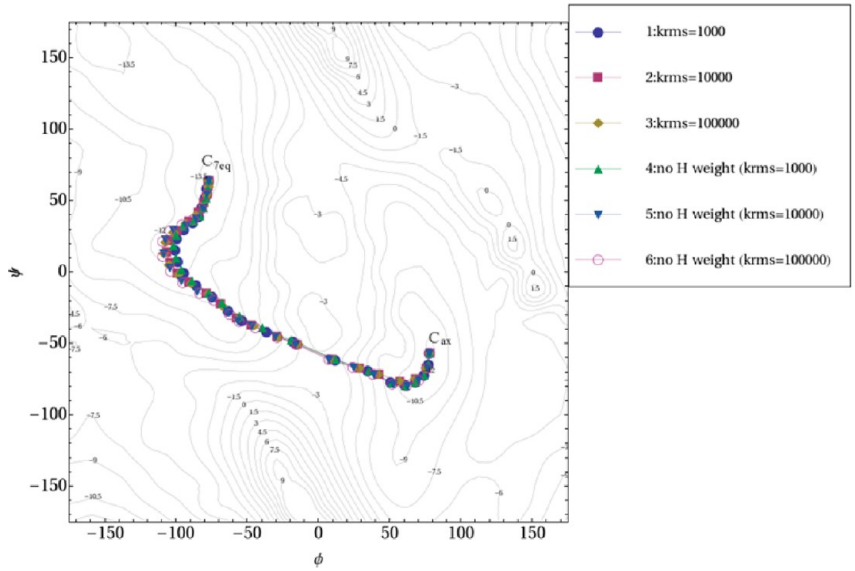

Figure 6. Reaction pathways of alanine dipeptide isomerization using the RPATH/restraint method. In pathways 1−3, the elastic force constant k is 1000, 10000, and 100000 kcal•mol^{-1}•Å^{-2}, respectively. The mass-weighted RMSD of all the atoms is used to measure the distance between replicas. Pathways 4−6 repeat the calculations of 1−3 with hydrogens excluded from the mass-weighted RMSD calculation.

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

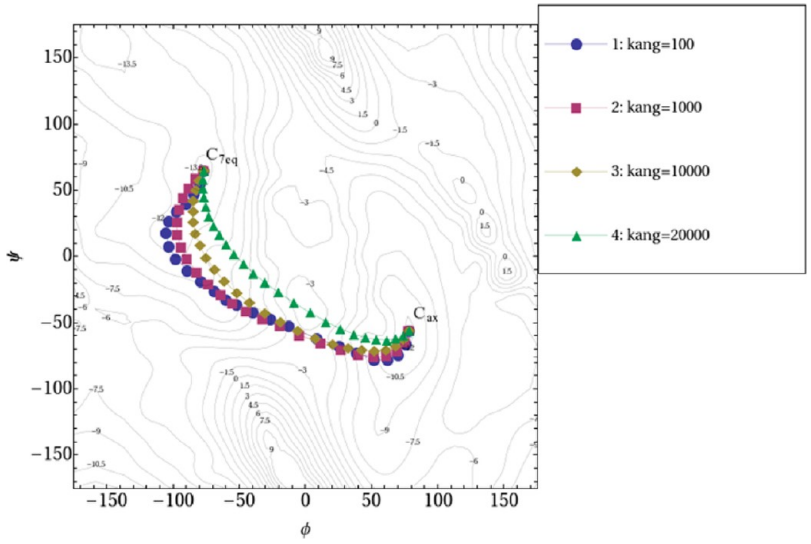

Figure 8. Reaction pathways of alanine dipeptide isomerization using the RPATH/restraint method. The k_{ang} for the pathway curvature controlling is 100, 1000, 10000, and 20000, for pathways 1−4, respectively. For all the calculations, k_{rms} is 100000, and hydrogens are excluded from mass-weighted RMSD measurement (force constant unit: kcal•mol^{-1}•Å^{-2}).

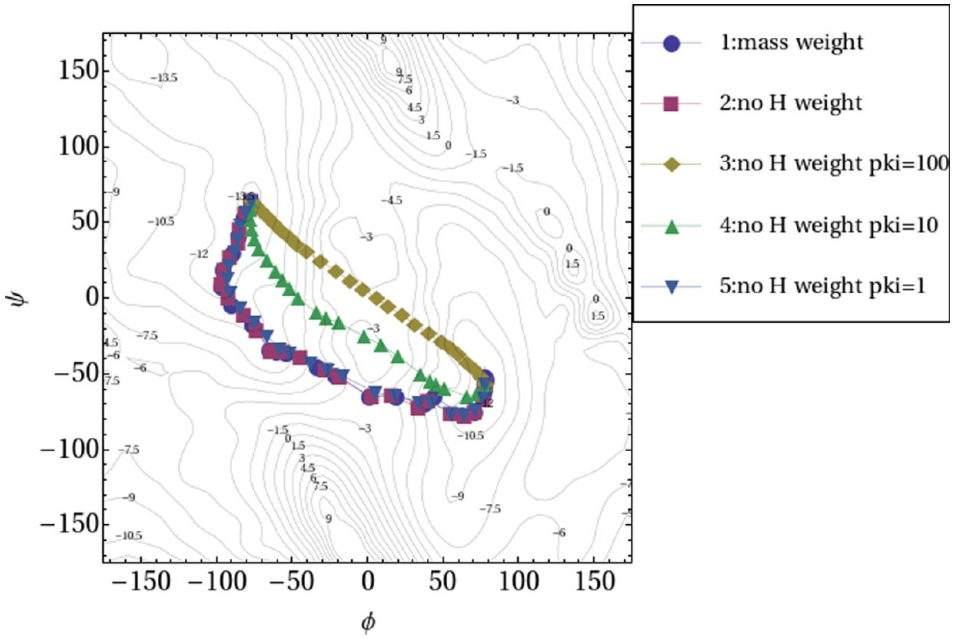

Figure 9. Reaction pathways of alanine dipeptide isomerization using the RPATH/constraint method. The hydrogen is included for mass-weighted RMSD for pathway 1 but excluded in all other pathways. The kinetic energy potential is included in pathways 3−5 with force constants as 100, 10, and 1 kcal•mol^{-1}•Å^{-2}, respectively.

the QM calculations that do not include the zero point vibrational energy. Both the reaction barrier and TS structure from these QM calculations serve as benchmarks for the RPATH calculations. Due to the computational cost of the RPATH QM calculations, they were considered converged when the RMS force of the path is below 0.1 kcal•mol^{-1}•Å^{-1} and the change of the total reaction path energy was less than 0.01 kcal•mol^{-1} for the last optimization step of RPATH calculations.

3.2.1. NEB Results. Four NEB calculations were carried out and plotted in Figure 12. In calculations 1 and 2, all atoms are used to calculate the mass-weighted RMSD between adjacent replicas with a spring constant as 10 and 100 kcal•mol^{-1}•Å^{-2}, respectively. Calculations 3 and 4 repeat the first two calculations but with an additional weighting factor of 16 on the migrating hydrogen added to the mass-weighting scheme. For each NEB calculation, the structure with the highest energy is referred as the approximate TS structure. The approximate TS structures from these pathways are superimposed with the QM TS in Figure 13.

The barrier of pathway 1 is 67.83 kcal/mol, which is about 10 kcal/mol higher than the QM barrier. The barriers of pathways 2, 3, and 4 (55.83, 58.83, and 54.10 kcal/mol, respectively) are rather close to the QM barrier. The approximate TS structures from pathways 1, 2, and 4 (yellow, green, and red, respectively) in Figure 13 show that the positions of migration hydrogen are significantly different from that in the QM TS structure. The position of migration hydrogen in the approximate TS structure

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

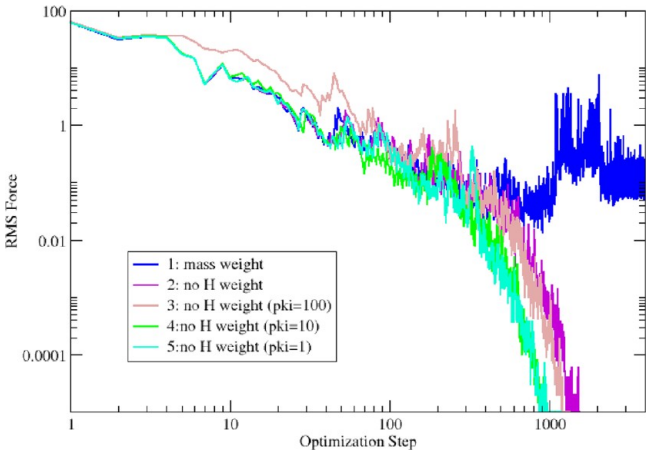

Figure 10. The RMS force (kcal•mol^{-1}•Å^{-1}) in the RPATH/constraint optimization of alanine dipeptide isomerization.

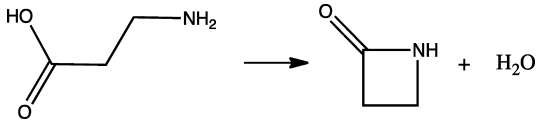

Figure 11. Intramolecular condensation of β-alanine.

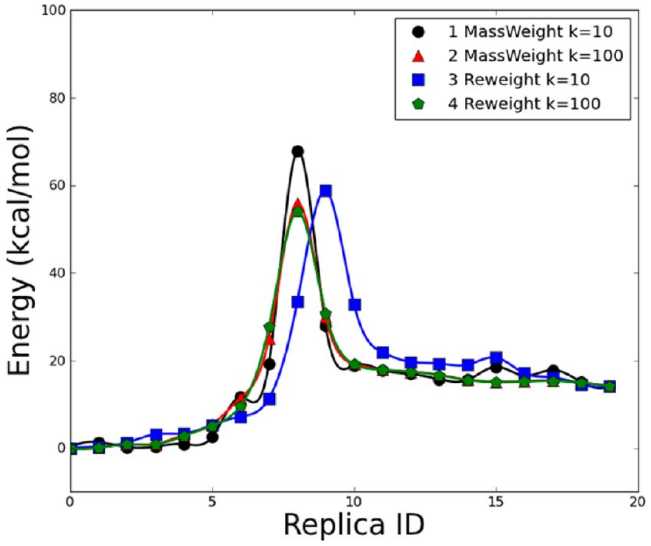

Figure 12. Energetic profile of β-alanine intramolecular condensation reaction using the NEB method with 20 replicas. Four calculations with different force constants k (kcal•mol^{-1}•Å^{-2}) and weighting schemes: 1, k = 10, mass-weight; 2, k = 100, mass-weight; 3, k = 10, additional weight on migration hydrogen; 4, k = 100, additional weight on migration hydrogen. All the calculations were carried out at the B3PW91/6-31g(d,p) level of theory, which was applied for all other β-alanine calculations.

from pathway 3 (blue in Figure 13) closely resembles the QM TS structure, but the overall structural difference between the approximate TS from pathway 3 and the QM TS is rather significant.

3.2.2. RPATH/Restraint Results. Four RPATH/restraint calculations are presented in Figure 14. The mass-weighted RMSD was used as a measurement of the distance between replicas. The approximate TS structures from these pathways are superimposed with the QM TS in Figure 15. The barriers of pathways 1 and 2 are around 40 kcal/mol, which is about 16

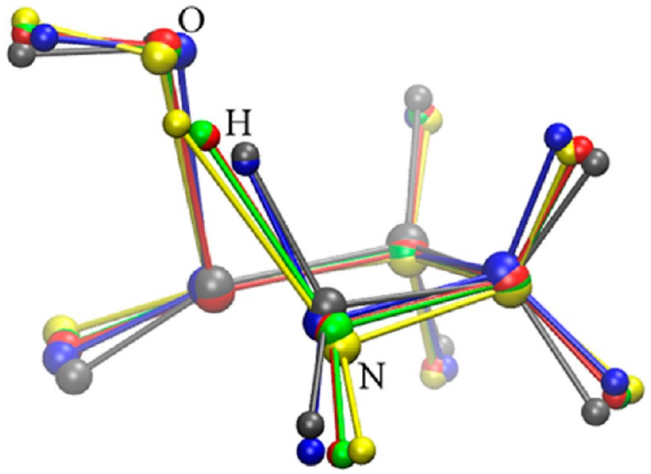

Figure 13. β-Alanine intramolecular condensation reaction approximate transition states from NEB calculations using 20 replicas. Four calculations with different force constants k (kcal•mol^{-1}•Å^{-2}): yellow, pathway 1, k = 10, mass-weight; green, pathway 2, k = 100, mass-weight; blue, pathway 3, k = 10, additional weight on migration hydrogen; red, pathway 4, k = 100, additional weight on migration hydrogen; gray: the TS obtained from QM calculation.

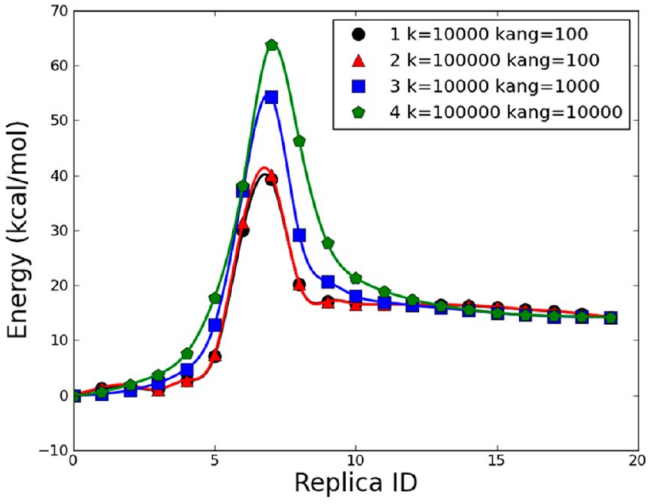

Figure 14. Energetic profile of β-alanine intramolecular condensation reaction using the RPATH/restraint method with 20 replicas. Four calculations with different force constants k and angle force k_{ang} in kcal•mol^{-1}•Å^{-2}.

kcal/mol lower than the QM barrier. The approximate TS structures from pathways 1 and 2 (yellow and green, respectively) in Figure 15 show that the positions of migration hydrogen are significantly different from that in the QM TS structure. Both energy and structure differences indicate that pathways 1 and 2 do not capture the TS accurately. Both calculations have large force constants for RMSD distance (k = 10000 and 100000 kcal•mol^{-1}•Å^{-2} for 1 and 2, respectively) and relatively small force constants for the angle term (k_{ang} = 100 kcal•mol^{-1}•Å^{-2}). Pathways 3 and 4 repeat the calculations of pathway 1 and with larger k_{ang} as 1000 and 10000 kcal•mol^{-1}•Å^{-2}, respectively. The barriers of pathways 3 and 4 are 54.32 and 63.78 kcal/mol, respectively, and are closer to the barrier from the QM calculation than pathways 1 and 2. Both approximate TSs from pathways 3 and 4 (blue and red) have the same nonmass-weighted RMSD distance (0.10 Å) to

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

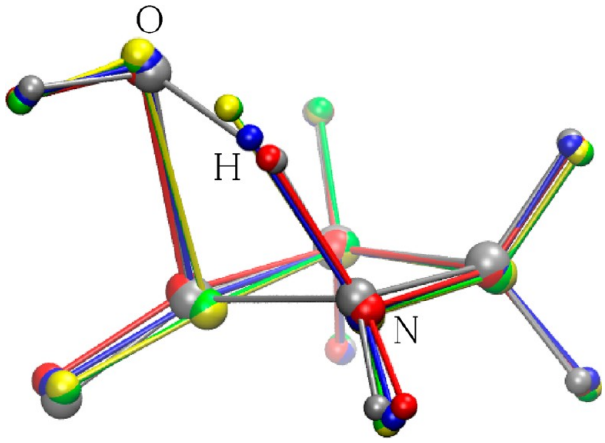

Figure 15. β-Alanine intramolecular condensation reaction approximate transition states from RPATH/restraint calculations using 20 replicas. Four calculations with different force constants k and angle force k_{ang} in kcal•mol^{-1}•Å^{-2}: yellow, pathway 1, k = 10000, k_{ang} = 100; green, pathway 2, k = 100000, k_{ang} = 100; blue, pathway 3, k = 10000, k_{ang} = 1000; red, pathway 4, k = 100000, k_{ang} = 10000; gray: the TS obtained from QM calculation.

the QM TS (gray), while the approximate TSs from pathways 1 and 2 also have the same value (0.18 Å) for such distance to the QM TS.

These four RPATH/restraint calculations were repeated with an additional weighting factor of 16 on the migrating hydrogen added to the mass-weighting scheme. However, the added weighting factor on the migrating hydrogen did not improve the pathways in terms of smoothness and estimated reaction barriers (see Figure S4 in the Supporting Information).

3.2.3. RPATH/Constraint Results. Two sets of RPATH/constraint calculations are presented: one with mass-weighted RMSD (set A) and the other with an additional weighting factor of 16 on the migrating hydrogen added to the mass-weighting scheme (set B). For set A, pathways 1, 2, and 3 with k_{pki} as 100, 10, and 1 kcal•mol^{-1}•Å^{-2}, respectively, are plotted in Figure 16A. Pathway 3 has the smallest k_{pki} and yields a barrier that is the closest to the QM calculation, differing by only 3 kcal/mol. All three approximate TS structures from pathways 1, 2, and 3 are superimposed with the QM TS in Figure 17A. The migrating hydrogen in all three approximate TSs shows a significant difference from QM TS. The RMSD of these approximate TSs in reference to the QM TS ranges from 0.13 to 0.45 Å, with the approximate TS of pathway 1 displaying the largest value.

In set B, pathways 1, 2, 3, and 4 with k_{pki} as 100, 10, 1, and 0.1 kcal•mol^{-1}•Å^{-2}, respectively, are plotted in Figure 16B. Pathway 1 has an overestimated barrier, 65.22 kcal/mol (Table 2). Pathways 2, 3, and 4 have reaction barriers very close to the QM barrier with less than 0.5 kcal/mol difference. The superimposed approximate TS structures in Figure 17B show the extreme similarities between the approximate TSs from pathways 2, 3, and 4 (green, blue, and red, respectively) and the QM TS (gray), especially the position of the migrating hydrogen. All three approximate TSs have very small RMSDs, which are equal to or are less than 0.03 Å with reference to the QM TS (Table 2). By emphasizing the movement of the migrating hydrogen with a large weighting factor, set B showed significant improvement compared with set A.

3.3. Inhibition Mechanism of MMP2. The inhibition mechanism of MMP2 by its potent inhibitor SB-3CT is a coupled deprotonation of the methylene group juxtaposed between the sulfone and the thiirane that opens the thiirane ring (Figure 18). This reaction creates a thiolate anion that strongly coordinates with a zinc atom in the active site. This reaction has been previously studied by Tao et al. using the ONIOM method, which is a QM/MM method.[76−78] In the present work, this reaction mechanism was employed as a test case for NEB and RPATH methods in CHARMM. The calculations were carried out using QM/MM methods through an interface of Q-Chem and CHARMM developed in our lab.[79] The CHARMM 22 force field[70] with CMAP backbone dihedral angle corrections[71] is used for protein and the CHARMM general force field (CGenFF) for the inhibitor.[80] The B3LYP/

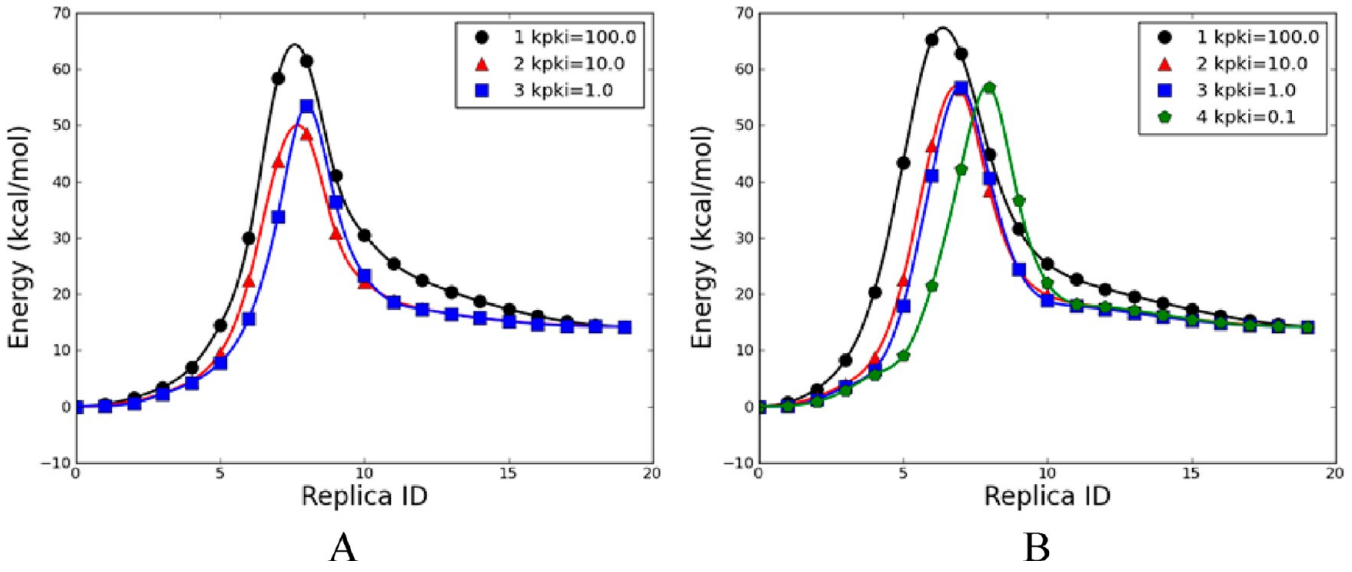

Figure 16. Energetic profile of β-alanine intramolecular condensation reaction using the RPATH/constraint method with 20 replicas. A. Three calculations using mass-weighted RMSD with different k_{pki} (kcal•mol^{-1}•Å^{-2}) values. B. Four calculations using mass-weighted RMSD and an additional weighting factor on migration hydrogen with different k_{pki} values.

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

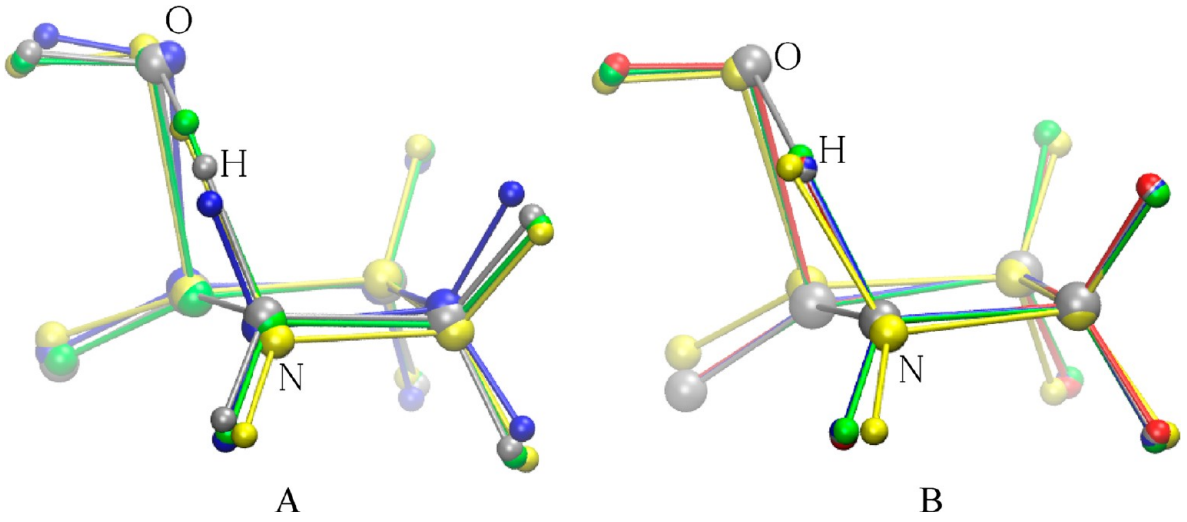

Figure 17. β-Alanine intramolecular condensation reaction approximate transition states from RPATH/constraint calculations with 20 replicas. A. Three calculations using mass-weighted RMSD with different kinetic k_{pki} (kcal•mol^{-1}•Å^{-2}) values: yellow, k_{pki} = 100.0; green, k_{pki} = 10.0; blue, k_{pki} = 1.0; gray: the TS obtained from QM calculation. B. Four calculations using mass-weighted RMSD and an additional weighting factor on migration hydrogen with different kinetic k_{pki} values: yellow, k_{pki} =100.0; green, k_{pki} = 10.0; blue, k_{pki} =1.0; red, k_{pki} = 0.1; gray: the TS obtained from QM calculation.

Table 2. Calculations for β-Alanine Intramolecular Condensation Reaction[a]

methods	parameters[b]		barrier (kcal/mol)	RMSD with QM TS (Å)	replica ID of approximate TS
NEB	mass-weight	1 k = 10	67.83	0.28	9
		2 k = 10^2	55.88	0.19	9
	mass-weight with additional weight on migrating hydrogen	1 k = 10	58.83	0.18	10
		2 k = 10^2	54.10	0.18	8
restraint	1 k_{rms} = 10^4, k_{ang} = 10^2		39.19	0.18	8
	2 k_{rms} = 10^5, k_{ang} = 10^2		40.16	0.18	8
	3 k_{rms} = 10^4, k_{ang} = 10^3		54.32	0.10	8
	4 k_{rms} = 10^5, k_{ang} = 10^4		63.78	0.10	8
constraint	mass-weight	1 k_{pki}= 10^2	61.52	0.45	9
		2 k_{pki}= 10	48.52	0.13	9
		3 k_{pki} = 1	53.45	0.18	9
	mass-weight with additional weight on migrating hydrogen	1 k_{pki}= 10^2	65.22	0.18	7
		2 k_{pki}= 10	56.43	0.03	8
		3 k_{pki} 1	56.79	0.01	8
		4 k_{pki} = 0.1	56.78	0.02	9

[a]All calculations were carried with the whole system described in QM at the B3PW91/6-31g(d,p) level of theory. [b]Unit for force constant is kcal•mol^{-1}•Å^{-2}.

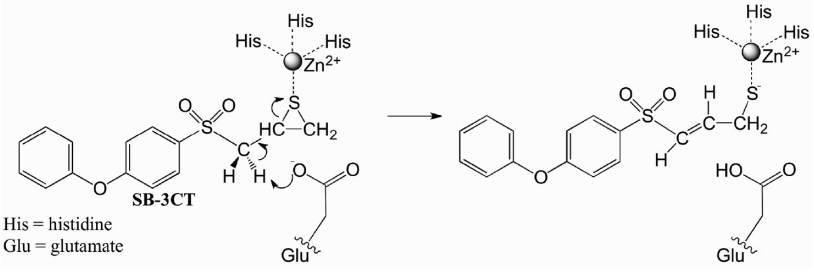

Figure 18. Inhibition mechanism of MMP2 by its inhibitor SB-3CT is coupled deprotonation of the methylene group juxtaposed between the sulfone and the thiirane and the opening of the thiirane ring.

6-31G(d) level of theory[73,74,81,82] was employed for all QM calculations. It should be noted that the QM/MM implementation for RPATH calculations through the CHARMM/Q-Chem interface[79] is an additive scheme using the electrostatic embedding method.[83] Unlike the ONIOM method,[84] the QM/MM geometry optimization of individual replica in this study is not an iterative procedure, i.e. no microiteration optimization was carried out for the QM subsystem.

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

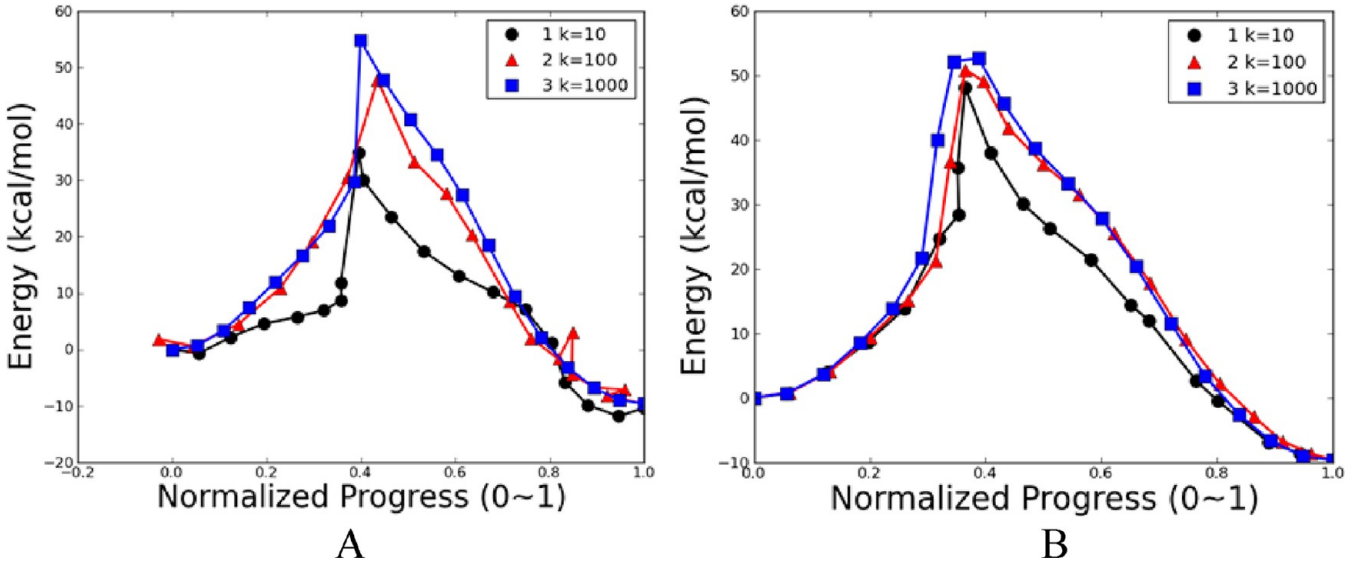

Figure 19. Energetic profiles of MMP2 inhibition mechanism by SB-3CT using the NEB method with 20 replicas and different force constants k in kcal•mol^{-1}•Å^{-2}. A. Three calculations using mass-weighted RMSD. B. Three calculations using a combination of mass and an additional weighting factor (see text for details).

Due to the high QM/MM computational cost, the RPATH calculations are considered as converged when the RMS force is less than 0.1 kcal•mol^{-1}•Å^{-2}, and the total energy change is less than 0.01 kcal/mol. A total of 20 replicas are employed in the RPATH optimization calculation for MMP2. For some of the calculations, additional replicas were inserted between each adjacent replica pair after optimization. The generated reaction path with 39 replicas in total was also subject to RPATH optimization to confirm the reaction barriers obtained in the calculations with 20 replicas. For the reaction pathways obtained from RPATH calculations, the approximate TS refer to the replica with the highest energy along the path.

An estimated TS of this reaction was obtained from restrained scan as benchmark. The restrained scan with 21 steps was performed between the two adjacent replicas for the replica with the highest energy from pathway 4 using RPATH/constraint. The breaking C−H and C−S bonds and forming O−H bond were restrained simultaneously, while all other degrees of freedom are fully optimized for each calculation. This restrained scan generated a quadratic energetic profile (see Figure S5 in the Supporting Information) with the highest energy structure as estimated TS, which leads to a barrier as 33.85 kcal/mol. In a previous study of this enzyme,[76] the reaction barrier of this inhibiting mechanism by SB-3CT was estimated as 19.9 kcal/mol using a subtractive QM/MM method, ONIOM.[84] The difference between the reaction barriers estimated in this and previous studies may originate from the fact that the initial MMP2 inhibitor complex structure used in this study with the CHARMM force field[70,71] was directly taken from the previous study, in which the AMBER force field[85] was used in both molecular dynamics and QM/MM calculations. The ONIOM reactant and TS from the previous study were also subjected to a single point QM/MM calculation using the CHARMM/Q-Chem interface at the same level of theory in this study. However, due to the difference between the force fields applied in this and previous studies, the reaction barrier calculated in this way is too high to be meaningful (data not shown).

3.3.1. NEB Results. In NEB calculations, two sets of weighting schemes were applied for RMSD distance measurements between adjacent replicas. In scheme A, mass-weighted RMSD is calculated using only QM atoms including hydrogen. In an attempt to better describe the reaction progress, additional arbitrary weighting factors were added in addition to an atomic mass-weighting scheme emphasizing different atoms of the QM region. In scheme B, additional weighting factors were added to give a different emphasis on different parts of the QM region. A factor of 50.0 is given to the migrating hydrogen, a factor of 3.0 is given to the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from SB-3CT, and a factor of 1.0 for all other QM atoms. Three harmonic spring constants, 10, 10^2, and 10^3 kcal•mol^{-1}•Å^{-2}, were applied using weighting schemes A and B and are plotted in Figure 19A and B, respectively. The reaction pathway energetics are plotted against the progression parameter d, which is illustrated in Figure 20.

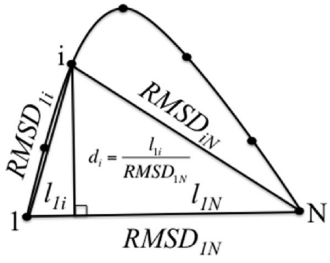

Figure 20. Illustration of progression parameter d_i for replica i.

For pathways using weighting scheme A, a single reaction barrier is present in all three calculations as 34.90, 47.76, and 54.80 kcal/mol, respectively, with the barrier from pathway 1 closest to the benchmark 33.85 kcal/mol. It is noticeable that the parameter d does not progress smoothly in pathways 1 or 2 (Figure 19A). In comparison, the pathways using weighting scheme B display smoother progressing of parameter d (Figure 19B). However, all the barriers of pathways using weighting

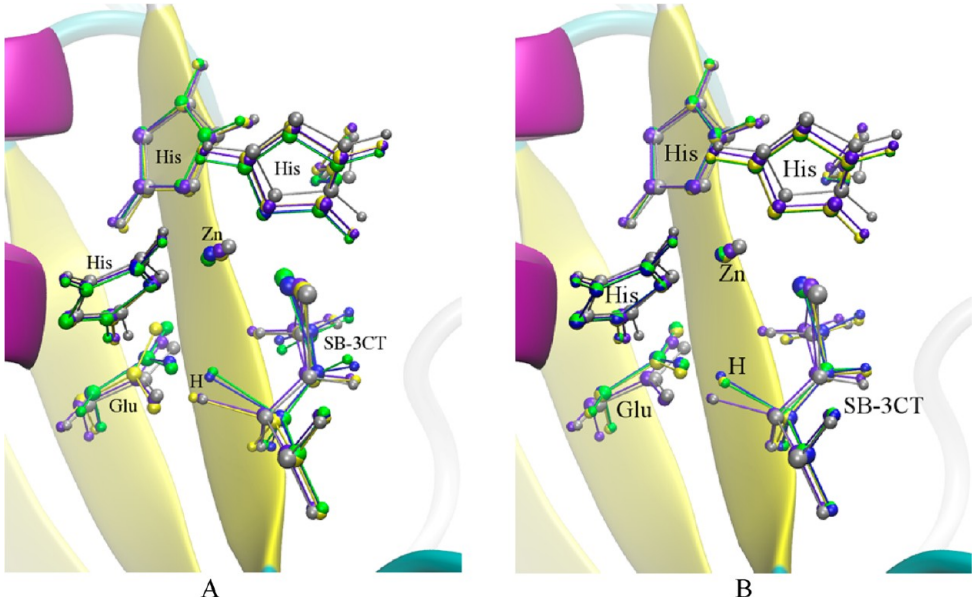

Figure 21. SB-3CT ring-opening approximate transition states from NEB calculations and different force constant k in kcal•mol^{-1}•Å^{-2}. A. Three calculations using mass-weighted RMSD. B. Three calculations using a combination of mass and an additional weighting factor (see text for details). In both figures: yellow: k = 10; green: k = 100; blue: k = 1000; violet: the estimated TS; gray: the TS obtained from ONIOM calculation in a previous study.[76]

Table 3. Calculations for Matrix Metalloproteinase 2 (MMP2) Inhibition Mechanism[a]

methods	parameters[b] (weighting scheme)	total replicas	barrier (kcal/mol)	RMSD with QM/MM TS (Å)	replica ID of approximate TS
NEB	1: k = 10, (A)[c]	20	34.90	0.63	9
	2: k = 100, (A)	20	47.76	0.75	8
	3: k = 1000, (A)	20	54.80	0.74	9
	4: k = 10, (B)[d]	20	48.08	0.72	9
	5: k = 100, (B)	20	50.86	0.73	8
	6: k = 1000, (B)	20	52.66	0.74	9
restraint	1: k = 1000, (A)	20	35.88	0.73	10
	2: k = 1000, (A)	39	40.31	0.71	17
	3: k = 1000, (C)[e]	20	43.33	0.71	9
	4: k = 1000, (C)	39	47.47	0.71	16
constraint	1: (A)	20	34.46	0.64	9
	2: (A)	39	34.73	0.64	17
	3: (B)	20	33.20	0.64	10
	4: (B)	39	34.74	0.63	18
	5: k_{pki} = 1, (A)	20	52.90	0.74	9
constraint/REPD	1: (A)	20	30.57	0.64	9
	2: (A)	39	32.00	0.63	17
	3: (D)[f]	20	33.76	0.62	9
	4: (D)	39	33.77	0.62	17

[a]The RPATH calculations were carried with the whole system described in QM/MM at the CHARMM22:B3LYP/6-31g(d) level of theory. [b]Unit for force constant is kcal•mol^{-1}•Å^{-2}. [c]Weighting scheme A: mass-weighted RMSD is calculated using only QM atoms including hydrogen. [d]Weighting scheme B: In combination with weighting scheme A, a factor of 50.0 is given to the migrating hydrogen, a factor of 3.0 is given to the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from SB-3CT, and a factor of 1.0 for all other QM atoms. [e]Weighting scheme C: In combination with weighting scheme A, a factor of 36 is given for the migrating hydrogen, a factor of 3 is given for the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from SB-3CT, and a factor of 0.5 for all other QM atoms. [f]Weighting scheme D: In combination with weighting scheme A, a factor of 50.0 is given for the migrating hydrogen, a factor of 3.0 is given for the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from SB-3CT, and a factor of 0.5 for all other QM atoms.

scheme B (48.08, 50.86, and 52.66 kcal/mol) are much higher than the benchmark value. The approximate TS structures of all pathways using weighting schemes A and B are illustrated in Figure 21 with the estimated TS and TS obtained from previous QM/MM study.[76] Pathway 1 with the weighting scheme A is more consistent with the estimated TS than the other five pathways in terms of reaction barrier and the nonmass-weighted RMSD of the atoms in QM region with reference to the estimated TS (Table 3). The position of migrating hydrogen in the estimated TS structures from the five pathways (Figure 21 A and B) except for pathway 1 with the

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

weighting scheme A is significantly different from those in the benchmark TS structures.

3.3.2. RPATH/Restraint Results. Force constants k_{rms} = 1000 kcal•mol^{-1}•Å^{-2}, k_{ang} = 100 kcal•mol^{-1}•Å^{-2}, COSMAX = 0.95 were applied for RPATH/restraint calculations for the MMP2/SB-3CT system. First, the weighting scheme A was used as distance measurements between adjacent replicas. The single-point QM/MM energy of each replica is plotted as pathway 1 in Figure 22 against the normalized reaction progress

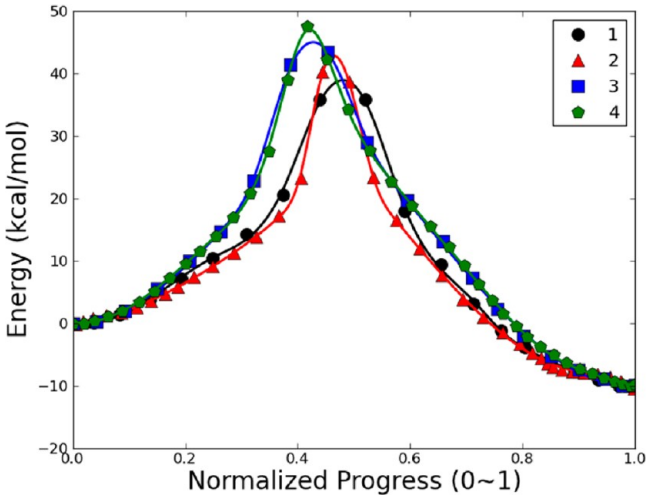

Figure 22. Energetic profile of MMP2 inhibition mechanism by SB-3CT using RPATH/restraint. 1: mass-weighted RMSD with 20 replicas; 2: mass-weight with 39 replicas; 3: combination of mass and additional weighting factors RMSD and with 20 replicas; 4: combination of mass and additional weighting factors RMSD and with 39 replicas.

parameter. A single reaction barrier 35.88 kcal/mol is obtained from this pathway (Table 3). To obtain a better estimation of the reaction barrier, and confirm that no other reaction barrier along the obtained reaction pathway exists, an additional replica was inserted between each adjacent replica pair using linear interpolation. The new reaction pathway with 39 replicas was optimized (pathway 2 in Figure 22) and shows an increased reaction barrier of 40.31 kcal/mol (Table 3), which is much higher than 33.85 kcal/mol, the estimated barrier.

In an attempt to better describe the reaction progress using RPATH/restraint, a weighting scheme C was added in addition to atomic mass-weighting scheme A. In this case, a factor of 36 is given for the migrating hydrogen, a factor of 3 is given for the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from substrate SB-3CT, and a factor of 0.5 for all other QM atoms. The reaction pathway shows a single barrier as 43.33 kcal/mol. Further optimization with an additional 19 replicas increases the barrier to 47.47 kcal/mol. Both barriers are much higher than 33.85 kcal/mol, the estimated barrier.

These plots demonstrate that the reaction proceeds from reactant to product smoothly. The inter-replica distance is larger in the vicinity around the TS region compared to the end point regions. When applying additional weighting factors, the distribution of replicas along reaction pathway is somewhat more even than just merely mass-weighting. All four pathways clearly show a single barrier in agreement with a concerted reaction mechanism for proton migration and thiirane ring-opening.[76]

The approximate TS structures of all four pathways are illustrated in Figure 23 with the estimated TS and the TS

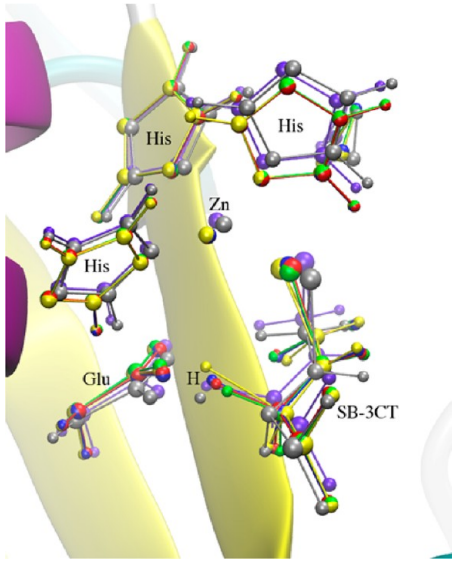

Figure 23. SB-3CT ring-opening approximate transition states from RPATH/restraint calculations: yellow: calculation with 20 replicas using mass-weighted RMSD, green: calculation with 39 replicas using mass-weighted RMSD, blue: calculation with 20 replicas using mass-weighted RMSD with additional weighting factors, red: calculation with 39 replicas using mass-weighted RMSD with additional weighting factors, gray: the TS obtained from ONIOM calculation in a previous study.[76]

obtained from the previous QM/MM study.[76] All four structures have nonmass-weighted RMSD of the QM region close to 0.3 Å with reference to the estimated TS (Table 3).

3.3.3. RPATH/Constraint Results. The RPATH calculation with 20 replicas using a constraint on the weighting scheme A shows a single barrier of 34.46 kcal/mol (pathway 1 in Figure 24). After inserting additional replicas between adjacent

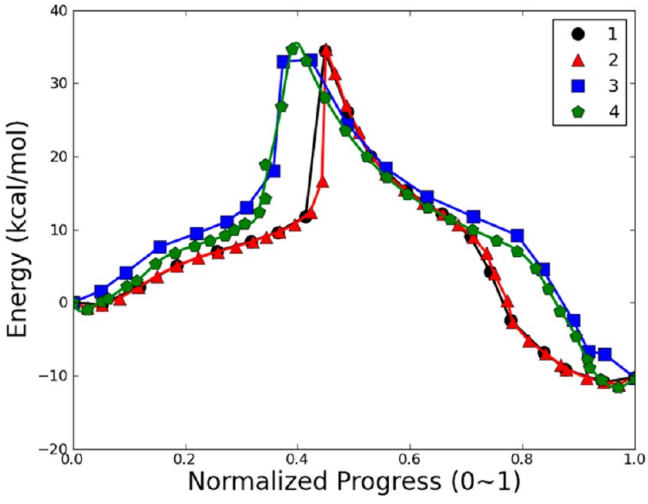

Figure 24. Energetic profiles of MMP2 inhibition mechanism by SB-3CT using RPATH/constraint. 1: mass-weighted RMSD with 20 replicas; 2: mass-weight with 39 replicas; 3: combination of mass and additional weighting factor RMSD and with 20 replicas; 4: combination of mass and additional weighting factor RMSD and with 39 replicas.

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

replicas, the new reaction pathway with 39 total replicas was optimized showing a single reaction barrier of 34.73 kcal/mol (pathway 2 in Figure 24). The energetic profiles of these two pathways are consistent with each other and with differences less than 1 kcal/mol to the estimated barrier, 33.85 kcal/mol. The two approximate TSs have almost identical progression parameters.

The weighting scheme B used in section 3.3.1 was used in addition to atomic mass-weighting scheme A. The reaction pathway has a single barrier of 33.20 kcal/mol (pathway 3 in Figure 24). A further optimized pathway with 19 additional replicas (pathway 4 in Figure 24) has two replicas close in energy around the TS region with a barrier of 34.74 kcal/mol. Both barriers have differences less than 1 kcal/mol to the estimated barrier, 33.85 kcal/mol. A kinetic energy force constant of 1 kcal•mol^{-1}•Å^{-2} was applied to obtain a MHP of this reaction. The reaction pathway is rather smooth and evenly distributed (Figure 26). However, the reaction barrier of this pathway is 52.90 kcal/mol, and is much higher than the estimated barrier.

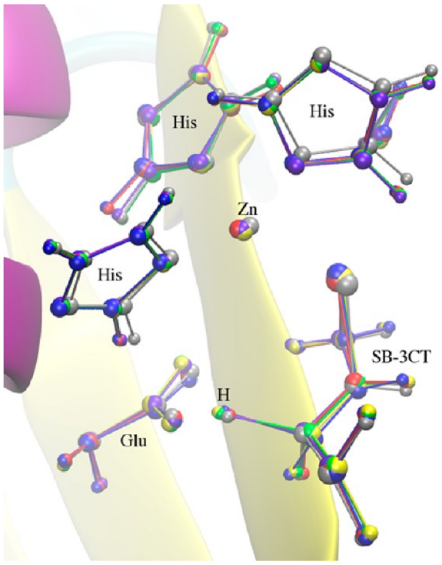

Figure 25. SB-3CT ring-opening approximate transition states from RPATH/constraint calculations: yellow: calculation with 20 replicas using mass-weighted RMSD, green: calculation with 39 replicas using mass-weighted RMSD, blue: calculation with 20 replicas using mass-weighted RMSD with additional weighting factors, red: calculation with 39 replicas using mass-weighted RMSD with additional weighting factors, gray: the TS obtained from ONIOM calculation in a previous study.[76]

The approximate TS structures of all four pathways from Figure 24 are illustrated in Figure 25 with the estimated TS and the TS obtained from the previous QM/MM study.[76] All four approximate TSs obtained from RPATH calculations resemble with each other closely as well as with the two reference TSs. The migrating hydrogens in the approximate TSs from four pathways are almost on top of each other as well as the two reference TSs. The geometries of the thiirane ring-opening in all six approximate TSs are also very close. This observation shows that the RPATH/constraint is a very useful tool to produce the reaction pathway with accurate TS information.

3.3.4. RPATH/Constraint with REPD Framework Results. The reaction path calculated using RPATH with constraints

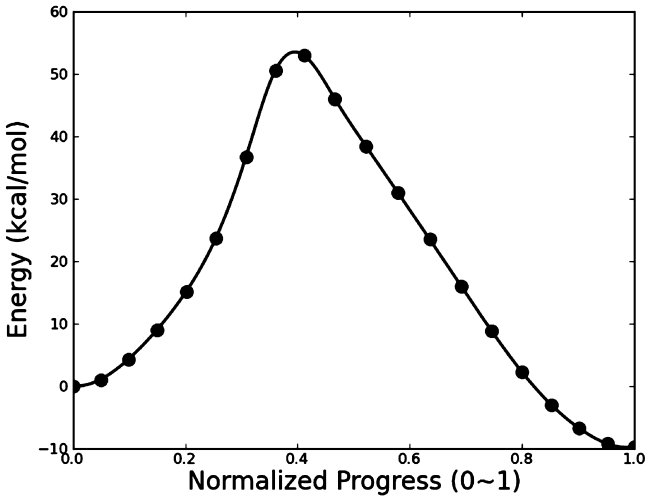

Figure 26. Energetic profile of MMP2 inhibition mechanism by SB-3CT using the RPATH/constraint method with more weight on migration hydrogen and k_{pki} = 1 kcal•mol^{-1}•Å^{-2}.

and the REPD framework computes a reaction barrier of 30.57 kcal/mol (pathway 1 in Figure 27). After inserting additional

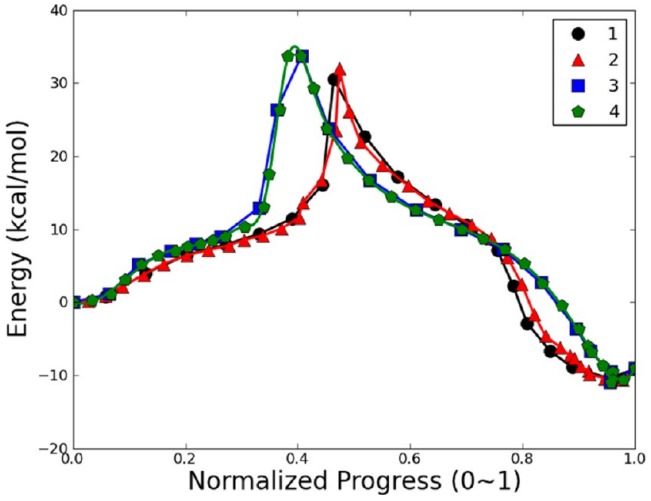

Figure 27. Energetic profiles of MMP2 inhibition mechanism by SB-3CT using the RPATH/constraint method in distributed replica (REPD) framework. 1: mass-weighted RMSD with 20 replicas; 2: mass-weighted RMSD with 39 replicas; 3: combination of mass and additional weighting factor RMSD and with 20 replicas; 4: combination of mass and additional weighting factor RMSD and with 39 replicas.

replica between adjacent replicas, the new reaction pathway with 39 replicas was optimized and showed a single reaction barrier as 32.00 kcal/mol (pathway 2 in Figure 27). Both barriers are lower than the estimated barrier, 33.85 kcal/mol. The two pathways are very consistent with each other.

A weighting scheme D was added in addition to atomic mass-weighting scheme A. In this case, a factor of 50.0 is given for the migrating hydrogen, a factor of 3.0 is given for the carboxylate group of Glu289, sulfone, thiirane ring, and the methylene group (excluding the migrating hydrogen) from SB-3CT, and a factor of 0.5 for all other QM atoms. The pathways with 20 replicas (3 in Figure 27) and 39 replicas (4 in Figure 27) have

dx.doi.org/10.1021/ct3006248 | J. Chem. Theory Comput. 2012, 8, 5035−5051

barriers as 33.76 and 33.77 kcal/mol (4 in Figure 27), respectively, very close to the estimated barrier, 33.85 kcal/mol.

The approximate TS structures of these four pathways are illustrated in Figure 28 and are compared with the estimated TS

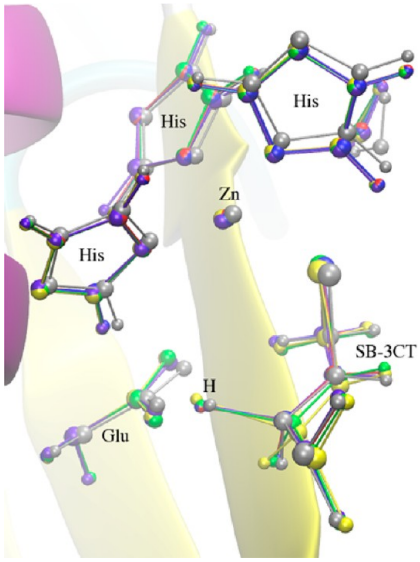

Figure 28. SB-3CT ring-opening approximate transition states from RPATH/constraint calculations in REPD framework: yellow: calculation with 20 replicas using mass-weighted RMSD, green: calculation with 39 replicas using mass-weighted RMSD, blue: calculation with 20 replicas using mass-weighted RMSD with additional weighting factors, red: calculation with 39 replicas using mass-weighted RMSD with additional weighting factors, gray: the TS obtained from ONIOM calculation in a previous study.[76]

and the TS from the previous QM/MM study.[76] Similar to the RPATH/constraint calculations, all four approximate TSs obtained from RPATH/constraint with the REPD framework are very close to the reference TSs, especially the position of the migrating hydrogen and geometry of the opening thiirane ring.

4. DISCUSSION

In the present study, we applied three chain-based methods implemented in CHARMM on three test cases: alanine dipeptide isomerization, β-alanine intermolecular condensation, and the inhibition mechanism of MMP2 by SB-3CT. The levels of theory applied for these three systems are MM, QM, and QM/MM, respectively.

4.1. Isomerization of Alanine Dipeptide. The NEB method works very well for alanine dipeptide isomerization. For all three spring constants applied in the calculations, the obtained reaction pathways are almost identical. Theoretically, the path optimization using the NEB method should provide a MEP of the target reaction. This is apparently the case for the alanine dipeptide isomerization. For mass-weighted RMSD either with or without hydrogen atoms, the same reaction pathways were obtained for six NEB calculations using spring constants that differ by 3 orders of magnitude. Because of the force projection applied in NEB calculation, the RMS fluctuates during the optimization. This fluctuation does not diminish when close to convergence. This is one undesired feature and leads to slow convergence for more complicated systems when applying the NEB method.

The force constants applied in the RPATH/restraint calculations are much stronger than those applied in the NEB calculations. When applying the same force constants as in NEB, the replicas in the TS region suffer significant sliding down, i.e. the distances between adjacent replicas around the TS are much larger than those close to the end points. For force constants, k_{rms}, larger than 1000 kcal•mol^{-1}•Å^{-2}, the RPATH/restraint method calculations gave reaction pathways close to the MEP. The angle force constant, k_{ang}, is a very useful tool to control the smoothness of the pathway. By increasing k_{ang}, the optimized pathway can vary between the MEP and a straight-line interpolation between the two end points. This feature can be important when studying a reaction with a rough PES to prevent kinks along the optimized pathway. Smooth pathways can also serve as the reference for off-path dynamic simulation to obtain the free energy of the reaction.

The RPATH/constraint calculations also give an MEP. The fast convergence of the RMS force shows that this method is promising for reaction path optimizations. However, the convergence failure when including hydrogen in the RMSD measurement reminds the user to be careful when choosing a reaction coordinate. It is noticeable that the replicas on the ϕ and ψ contour plot for calculations 1 and 2 in Figure 9 are not evenly distributed as those obtained in NEB calculations. Including a kinetic potential energy in the objective function is an effective way to increase the smoothness of the reaction pathway but with the price of deviating from the MEP. By adjusting the kinetic energy force constant, one could obtain multiple reaction pathways that vary between the MEP and the straight line connecting the two end points (see calculations 3–5 in Figure 9). Similar to the k_{ang} for the restraint method, this could be a convenient tool when studying pathways on rough PESs or generating references for an off-path simulation.

4.2. β-Alanine Intermolecular Condensation. The NEB method provides rather smooth energetic profiles in all four calculations. The estimated reaction barriers using NEB from three pathways are also close to the barrier obtained from benchmark calculation. However, the estimated TS structures show significant deviation from the QM TS structure, even with an additional weighting factor on the migrating hydrogen. This observation suggests that caution needs to be used when applying the NEB method on more complex systems.

In calculations using the RPATH/restraint method, the force to control the smoothness of the path (k_{ang}) played an important role in the minimization. Only with a large k_{ang}, the RPATH calculations give a barrier and approximate TS that are close to the results from QM calculations. This is also likely due to the flatness of the PES around the product with two separate molecules.

The RPATH/constraint calculations with mass-weighted RMSD measurements do not show significant improvement compared with the restraint results. However, with an additional weighting factor on the migrating hydrogen, the constraint calculations reproduce the reaction barrier and the approximate TS structures of this reaction very accurately. It should be pointed out that the kinetic energy potential (with nonzero k_{pki}) is necessary in constraint calculations for convergence.

As a summary, special caution needs to be taken when applying the RPATH method on small organic systems using the QM methods, especially when separate molecules are present in either reactant or product or both. In such cases, additional forces or terms to maintain the pathway smoothness

and rigidity are needed to ensure the convergence of the calculations and the accuracy of the reaction barrier and the approximate TS structure.

4.3. Inhibition Mechanism of MMP2. The barriers of two NEB calculations with larger force constants are higher than those with a force constant of 10 kcal•mol^{-1}•Å^{-2}. This suggests that further relaxation of the reaction path is needed for the NEB calculations with large force constants. The progression of the reaction pathways with two smaller force constants is not smooth. The pathway with the largest force constant has a smoother progression but with higher reaction barrier than the estimated value. The calculation with additional weighting factors led to smoother energetic profiles, but higher reaction barrier, especially with large force constants. When using the NEB method to study large systems at a high level of theory, one needs to be very careful to choose the appropriate force constants to balance between the smoothness of pathway and the converging rate of the calculation.

The RPATH/restraint generated rather smooth pathways. The consistency of the four pathways shows the reliability of the RPATH/restraint method to capture the reaction mechanism. All the reaction barriers obtained from these calculations but one are much higher than the estimated barrier. The only one that is close to the estimated barrier, however, has the approximate TS structure with the largest RMSD from the estimated TS. The significant variation of the approximate TS structures and high reaction barriers indicates the difficulty of consistently converging to MEP for large systems when using various setups in the RPATH/restraint calculations. This inconsistency of RPATH/restraint calculations is also shown from β-alanine results.

The four RPATH/constraint calculations generate very consistent barrier heights, all between 33.20 and 34.74 kcal/mol, very close to the estimated barrier. The replicas are evenly distributed along the pathway, providing good coverage of the TS regions (Figure 24). All four approximate TS structures obtained from the RPATH/constraint calculations are very similar to the reference TS (Figure 25). The RPATH/constraint method seems to be a robust tool to study the reaction mechanisms of protein reactions. By including a kinetic energy potential, the RPATH/constraint calculation could produce a very smooth pathway, but the calculated reaction barrier may be significantly higher than the real barrier.

It should be pointed out that the different reaction progression parameters of approximate TSs shown in the energetic plots do not indicate significant difference among the approximate TS structures of the reaction, because this parameter depends on the definition of the RMSD and any additional weighting factors applied in the distance measurement. It is obvious that the approximate TS structures obtained from different RPATH calculations are very similar to each other as well as to the reference TSs. It should be emphasized that no universal setup of the RMSD distance will work well for all the RPATH calculations, especially for complicated protein systems. The users are suggested to try different RMSD schemes initially to find the best way for certain RPATH methods.

5. CONCLUDING REMARKS

In conclusion, replica path (RPATH) methods implemented in CHARMM are powerful tools to elucidate the reaction mechanism of systems with various sizes and complexity. Starting from reactant and product, a RPATH calculation could generate a reaction pathway represented by multiple replicas providing reaction barriers that can be used in comparisons with experimental results. There is no single option that will prevail in every case. Each method has its own strengths and weaknesses. The best choice for reaction pathway calculation clearly depends on the nature of the system of interest itself. In general, the NEB method works well on the system at low level of theory and low computational cost, i.e. molecular mechanics. After convergence, the NEB method could generate rather accurate MEP. The RPATH with restraints or constraints works well with large systems at higher levels of theory due to their computational efficiency and fast convergence rates.

For small organic reactions, most with separate reactants or products, special caution needs to be taken when applying RPATH methods to study reaction mechanisms. The separation of the molecules could bring difficulty to the optimization convergence. Options are available to control the rigidity of the pathway and therefore accelerate the convergence of the optimization. For small systems, it is always recommended to carry out a standalone TS search from the approximate TS structure obtained from the RPATH calculations. For many large systems, the choice of reaction coordinates may not be obvious. The RMSD measurement with additional weighting schemes provides practically infinite choices to describe the reaction progress. Our test calculations of the MMP2 system demonstrated that the choice of RMSD distance does not lead to different reaction mechanisms and has limited effects on the reaction barrier and approximate TS structures, especially with constraint methods.

The key point of this study is that the RPATH methods are powerful and useful tools to study the reaction mechanisms of macromolecular systems, such as enzymes. With these tools, plausible reaction mechanisms as a full pathway could be generated without TS information *a priori*. Currently, the RPATH minimization only provides MEPs and estimations of the reaction barrier without free energy information. In our future studies, the reaction paths will be subject to dynamics simulation to estimate the reaction free energy profile, which is directly connected to experimental observation.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Initial path for alanine dipeptide isomerization, reaction pathways of alanine dipeptide isomerization using PATH/restraint method with small force constants, the energetic profile of β-alanine internal condensation reaction using the RPATH/restraint method with an additional weighting factor on migration hydrogen, the energetic profile of the MMP2 system for the restrained scan between replicas 17 and 19 from pathway 4 using RPATH/constraint. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: peng.tao@nih.gov.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research was supported by the Intramural Research Program of the NIH, NHLBI. Computational resources and

services used in this work were provided by the LoBoS cluster of the National Institutes of Health.

REFERENCES

(1) Tolman, R. C. *J. Am. Chem. Soc.* **1925**, *47*, 1524−1553.
(2) Hänggi, P.; Borkovec, M. *Rev. Mod. Phys.* **1990**, *62*, 251−341.
(3) Heidrich, D. *The reaction path in chemistry: current approaches and perspectives*; Kluwer Academic Publishers: Boston, MA, 1995; pp 1−308.
(4) March, J. *March's Advanced organic chemistry: Reactions, mechanisms and structures*, 5th ed.; Wiley: New York, NY, 2001; pp 389−1604.
(5) Bader, R. F. W.; Gangi, R. A. In *Theoretical Chemistry*; Dixon, R. N., Thomson, C., Eds.; Royal Society of Chemistry: Cambridge, 1975; pp 1−65.
(6) Mezey, P. *Potential energy hypersurfaces*; Elsevier: Amsterdam; New York, 1987; pp 117−180.
(7) Truhlar, D. G. In *Encyclopedia of Physical Science and Technology*, 3rd ed.; Meyers, R. A., Ed.; Academic Press: New York, 2001; pp 9−17.
(8) Wales, D. *Energy landscapes*; Cambridge University Press: Cambridge, UK; New York, 2003; pp 1−433.
(9) Lewars, E. G. In *Computational Chemistry*; Lewars, E. G., Ed.; Springer Netherlands: Dordrecht, 2011; pp 9−43.
(10) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154−2161.
(11) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523−5527.
(12) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1991**, *95*, 5853−5860.
(13) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Phys.* **2004**, *120*, 9918−9924.
(14) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Theory Comput.* **2005**, *1*, 61−69.
(15) Hratchian, H. P.; Frisch, M. J.; Schlegel, H. B. *J. Chem. Phys.* **2010**, *133*, 224101.
(16) Hratchian, H. P.; Frisch, M. J. *J. Chem. Phys.* **2011**, *134*, 204103.
(17) Taylor, H.; Simons, J. *J. Phys. Chem.* **1985**, *89*, 684−688.
(18) Simons, J.; Nichols, J. *Int. J. Quantum Chem.* **1990**, *38*, 263−276.
(19) Nichols, J.; Taylor, H.; Schmidt, P.; Simons, J. *J. Chem. Phys.* **1990**, *92*, 340−346.
(20) Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 252−261.
(21) Jónsson, H.; Mills, G.; Jacobsen, K. W. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: Singapore, 1998; pp 385−404.
(22) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978−9985.
(23) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(24) Maragakis, P.; Andreev, S. A.; Brumer, Y.; Reichman, D. R.; Kaxiras, E. *J. Chem. Phys.* **2002**, *117*, 4651−4658.
(25) Alfonso, D. R.; Jordan, K. D. *J. Comput. Chem.* **2003**, *24*, 990−996.
(26) Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *120*, 2082−2094.
(27) Xie, L.; Liu, H.; Yang, W. *J. Chem. Phys.* **2004**, *120*, 8039−8052.
(28) Galván, I. F.; Field, M. J. *J. Comput. Chem.* **2008**, *29*, 139−143.
(29) Czerminski, R.; Elber, R. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 6963−6967.
(30) Czerminski, R.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 5580−5601.
(31) Woodcock, H. L.; Hodošček, M.; Sherwood, P.; Lee, Y. S.; Schaefer Iii, H. F.; Brooks, B. R. *Theor. Chem. Acc.* **2003**, *109*, 140−148.
(32) Brokaw, J. B.; Haas, K. R.; Chu, J.-W. *J. Chem. Theory Comput.* **2009**, *5*, 2050−2061.
(33) Elber, R.; Karplus, M. *Chem. Phys. Lett.* **1987**, *139*, 375−380.
(34) Czerminski, R.; Elber, R. *Int. J. Quantum Chem.* **1990**, *38*, 167−185.
(35) Ulitsky, A.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 1510−1511.
(36) Choi, C.; Elber, R. *J. Chem. Phys.* **1991**, *94*, 751−760.
(37) Nowak, W.; Czerminski, R.; Elber, R. *J. Am. Chem. Soc.* **1991**, *113*, 5627−5637.
(38) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1997**, *107*, 375−384.
(39) Ren, W. *Comm. Math. Sci.* **2003**, *1*, 377−384.
(40) E, W.; Ren, W.; Vanden-Eijnden, E. *Phys. Rev. B* **2002**, *66*, 052301.
(41) E, W.; Ren, W.; Vanden-Eijnden, E. *J. Chem. Phys.* **2007**, *126*, 164103.
(42) Cameron, M.; Kohn, R. V.; Vanden-Eijnden, E. *J. Nonlinear Sci.* **2010**, *21*, 193−230.
(43) E, W.; Ren, W.; Vanden-Eijnden, E. *J. Phys. Chem. B* **2005**, *109*, 6688−6693.
(44) Burger, S. K.; Yang, W. *J. Chem. Phys.* **2006**, *124*, 054109.
(45) Peters, B.; Heyden, A.; Bell, A. T.; Chakraborty, A. *J. Chem. Phys.* **2004**, *120*, 7877−7886.
(46) Quapp, W. *J. Chem. Phys.* **2005**, *122*, 174106.
(47) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *129*, 174109.
(48) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *130*, 244108.
(49) Quapp, W. *J. Theor. Comput. Chem.* **2009**, *8*, 101−117.
(50) Goodrow, A.; Bell, A. T.; Head-Gordon, M. *Chem. Phys. Lett.* **2010**, *484*, 392−398.
(51) Chu, J.-W.; Trout, B. L.; Brooks, B. R. *J. Chem. Phys.* **2003**, *119*, 12708−12717.
(52) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(53) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(54) Jensen, F. In *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons: 2006; pp 421−444.
(55) Ren, W.; Vanden-Eijnden, E.; Maragakis, P.; E, W. *J. Chem. Phys.* **2005**, *123*, 134109.
(56) Khavrutskii, I. V. *J. Chem. Phys.* **2006**, *125*, 174108.
(57) Gfeller, D.; Rios, P. D. L.; Caflisch, A.; Rao, F. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 1817−1822.
(58) Quapp, W. *J. Comput. Chem.* **2007**, *28*, 1834−1847.
(59) Goodrow, A. *J. Chem. Phys.* **2009**, *130*, 244108.
(60) Velez-Vega, C. *J. Chem. Phys.* **2009**, *130*, 225101.
(61) Tipper, D. J. *Rev. Infect. Dis.* **1979**, *1*, 39−54.
(62) Mascaretti, O. A.; Boschetti, C. E.; Danelon, G. O.; Mata, E. G.; Roveri, O. A. *Curr. Med. Chem.* **1995**, *1*, 441−470.
(63) Liotta, L. A.; Tryggvason, K.; Garbisa, S.; Robey, P. G.; Abe, S. *Biochemistry* **1981**, *20*, 100−104.
(64) Alexander, D. S.; Aimes, R. T.; Quigley, J. P. *Enzyme Protein* **1996**, *49*, 38−58.
(65) Briknarova, K.; Grishaev, A.; Banyai, L.; Tordai, H.; Patthy, L.; Llinas, M. *Structure (London, U. K.)* **1999**, *7*, 1235−1245.
(66) Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.; Tryggvason, K. *Science* **1999**, *284*, 1667−1670.
(67) Briknarova, K.; Gehrmann, M.; Banyai, L.; Tordai, H.; Patthy, L.; Llinas, M. *J. Biol. Chem.* **2001**, *276*, 27613−27621.
(68) Feng, Y.; Likos, J. J.; Zhu, L.; Woodward, H.; Munie, G.; McDonald, J. J.; Stevens, A. M.; Howard, C. P.; De Crescenzo, G. A.; Welsch, D.; Shieh, H.-S.; Stallings, W. C. *Biochim. Biophys. Acta, Proteins Proteomics* **2002**, *1598*, 10−23.
(69) Diaz, N.; Suarez, D. *J. Phys. Chem. B* **2008**, *112*, 8412−8424.
(70) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe,

M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(71) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(72) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244−13249.
(73) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(74) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269.
(75) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Yeh Lin, C.; Van Voorhis, T.; Hung Chien, S.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Min Rhee, Y.; Ritchie, J.; Rosta, E.; David Sherrill, C.; Simmonett, A. C.; Subotnik, J. E.; Lee Woodcock Iii, H.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer Iii, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.
(76) Tao, P.; Fisher, J. F.; Shi, Q.; Vreven, T.; Mobashery, S.; Schlegel, H. B. *Biochemistry* **2009**, *48*, 9839−9847.
(77) Zhou, J.; Tao, P.; Fisher, J. F.; Shi, Q.; Mobashery, S.; Schlegel, H. B. *J. Chem. Theory Comput.* **2010**, *6*, 3580−3587.
(78) Tao, P.; Fisher, J. F.; Shi, Q.; Mobashery, S.; Schlegel, H. B. *J. Phys. Chem. B* **2010**, *114*, 1030−1037.
(79) Woodcock, H. L.; Hodošček, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F.; Brooks, B. R. *J. Comput. Chem.* **2007**, *28*, 1485−1502.
(80) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671−690.
(81) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(82) Becke, A. D. *Phys. Rev. A: Gen. Phys.* **1988**, *38*, 3098−3100.
(83) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.
(84) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A., Jr.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815−826.
(85) 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.