

## Sampling Long Time Scale Protein Motions: OSRW Simulation of Active Site Loop Conformational Free Energies in Formyl-CoA:Oxalate CoA Transferase

Sangbae Lee,<sup>†</sup> Mengen Chen,<sup>‡</sup> Wei Yang,<sup>\*,‡</sup> and Nigel G. J. Richards<sup>\*,†</sup>

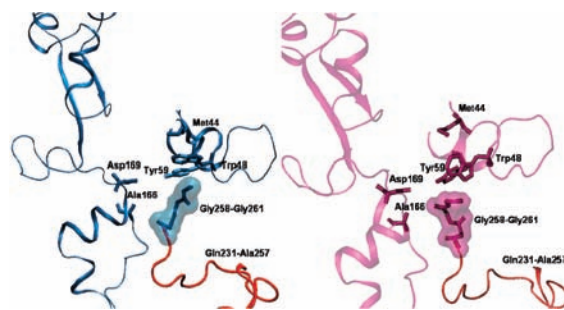
Department of Chemistry, University of Florida, Gainesville, Florida 32611 and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306

Received February 18, 2010; E-mail: yang@sb.fsu.edu; richards@qtp.ufl.edu

The importance of understanding the conformational energetics and dynamical properties of active site loops has driven the development of a number of sophisticated experimental biophysical methods to monitor the conformational transitions of these regions and their role in catalysis and/or binding.<sup>1,2</sup> Computational methods offer an alternate approach to evaluating these motions, but the determination of free energies using molecular dynamics is greatly complicated by the “lagging Hamiltonian” problem, which arises from the relatively long times that are needed for the protein environment to relax about a given loop conformation.<sup>3</sup> As a result, substantial effort is required to ensure adequate sampling for calculations employing perturbation<sup>4</sup> or thermodynamic integration<sup>5</sup> methods, limiting their application in studies of large molecular systems. In an effort to overcome this problem, Yang and co-workers have introduced an “Orthogonal Space Random Walk” (OSRW) sampling algorithm,<sup>6,7</sup> which allows synchronous acceleration of the sampling of order parameter moves and their coupled protein reorganizations. We now report the application of this method to compute the conformational energetics of the functionally important tetraglycine loop in the active site of *Oxalobacter formigenes* formyl-CoA:oxalate transferase (FRC).<sup>8</sup> Thus, OSRW-derived free energy profiles were computed for the interconversion of “open” and “closed” loop structures in FRC and correlated with the steady-state kinetic properties of the enzyme. In addition, the energetic and structural effects of introducing an alanine residue at each position of the tetraglycine loop were calculated. Taken overall, this comparison study demonstrates the capability of the OSRW strategy in predicting the free energy surfaces associated with active site loop motions.

FRC catalyzes the formation of oxalyl-CoA and formate from formyl-CoA and oxalate, which is an essential step for ATP synthesis in the bacterium.<sup>9</sup> As a member of a new family of CoA-transferases,<sup>10</sup> FRC is active as a homodimer in which the two monomers adopt a complex, interlocked fold.<sup>11</sup> Using a variety of conditions, our group recently reported an extensive set of crystal structures for FRC corresponding to various “snapshots” of the catalytic cycle of the enzyme.<sup>12</sup> These studies revealed conformational changes in a tetraglycine loop (Figure 1), which not only participates in stabilizing intermediates but also probably controls substrate access and product release.<sup>13</sup>

OSRW simulations were performed on the wild type FRC dimer and the G258A, G259A, G260A, and G261A FRC variants. Each calculation employed an all-atom representation of the homodimer embedded in a truncated octahedral box of explicit TIP3P water molecules, and the systems were described by the CHARMM27 force field.<sup>14</sup> The whole system was dynamically simulated, and the enhanced-sampling treatment was applied to one of the two



**Figure 1.** Cartoon representations of the FRC active site tetraglycine loop (Gly258Gly259Gly260Gly261) in its “open” (left) and “closed” (right) conformations. Coordinates are taken from the crystal structures of free FRC (1P5H) and the FRC/CoA complex (1P5R), respectively. Catalytic residues are also indicated although crystallographic waters are not shown for clarity.

active site loops in the homodimer. The order parameter  $\lambda$  was defined as  $(\text{rmsd}_0 - \Delta\text{rmsd})/2\text{rmsd}_0$ , where  $\text{rmsd}_0$  is the after-superimposition rms deviation between the loop structures in the “open” and “closed” conformations (Figure 1) and  $\Delta\text{rmsd}$  denotes the difference of the after-superimposition rms deviations of the dynamically evolving loop from the ones in the reference (“open” and “closed”) structures. Each OSRW simulation was begun from an equilibrated homodimer in which the loop was in its “open” conformation and continued until convergence of the free energy profile by sampling over the range  $\lambda = 0 \rightarrow \lambda = 1$  (Figures S1 and S2, Supporting Information). In general, converged energy profiles (Figure S3, Supporting Information) were generally achieved in  $\sim 2.0$  ns.

All of the free energy profiles show two minima for the tetrapeptide loops, with the exception of the G259A FRC variant, for which there are three stable conformations of similar energy. For WT FRC, the calculations showed a small free energy preference for the “open” loop conformation (consistent with the need to bind substrate). The barrier to interconversion was computed to be 13 kcal/mol (Table 1), which is lower than the experimental energy of 16.3 kcal/mol at 303.15 K based on the turnover number<sup>5</sup> and assuming a value of  $5 \times 10^{12} \text{ s}^{-1}$  for the pre-exponential factor.<sup>15</sup> We are aware, however, that the use of a  $\Delta\text{rmsd}$ -type order parameter may imperfectly describe the target events; therefore the use of this parameter may underestimate free energy barrier heights.<sup>16</sup> OSRW simulations of the loop conformational interconversion for the G261A FRC mutant gave a similar free energy profile (Table 1). This similarity in the two free energy profiles is consistent with the experimental observation that the experimental kinetics of the G261A FRC variant are essentially unchanged from those of the wild type enzyme.

OSRW free energy profiles calculated for the tetrapeptide loop conformations in the G258A, G259A, and G260A FRC mutants

<sup>†</sup> University of Florida.

<sup>‡</sup> Florida State University.

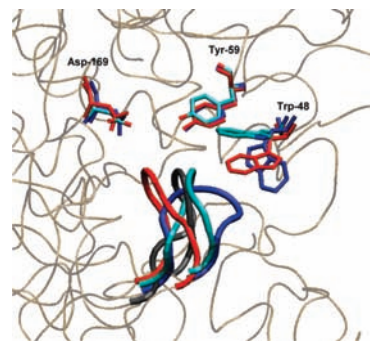
**Table 1.** Calculated Free Energy Values (kcal/mol) for Conformational Interconversion of the Tetrapeptide Loop and Steady-State Kinetic Parameters for Free FRC and Alanine-Containing FRC Mutants

Enzyme	$\Delta\Delta G^a$	$\Delta\Delta G^f$	$k_{\text{cat}}$ (s <sup>-1</sup> ) <sup>b</sup>	$k_{\text{cat}}/K_m$ (M <sup>-1</sup> s <sup>-1</sup> ) <sup>c</sup>
FRC	4.8	13.0	5.3 ± 0.1	1.4 × 10 <sup>3</sup>
G258A	5.3	23.0	— <sup>d</sup>	— <sup>d</sup>
G259A	2.2	9.0	1.9 ± 0.1	160
G260A	2.5	9.3	0.23 ± 0.02	12
G261A	3.7	10.5	1.65 ± 0.01	3.5 × 10 <sup>3</sup>

<sup>a</sup> All calculated free energies are in kcal/mol.  $\Delta\Delta G = \Delta G(\text{“closed”}) - \Delta G(\text{“open”})$ , and  $\Delta\Delta G^f$  is the barrier relative to the “open” conformation. <sup>b</sup> Experimental steady-state kinetic parameters at 30 °C have been published previously<sup>12</sup> and are included here for ease of comparison. <sup>c</sup> These  $k_{\text{cat}}/K_m$  values are for oxalate. <sup>d</sup> Technical problems in expressing and purifying the G258A FRC mutant have precluded its characterization.

were different from those determined for the wild type enzyme and G261A. Most strikingly, these simulations identified additional conformational minima that were intermediate in shape between the “open” and “closed” loop conformations and comparable in free energy (Figure S3, Supporting Information). The barriers to conformational interconversion in the G259A and G260A FRC mutants, however, were slightly lower than FRC. Experimentally, both the G259A and G260A FRC mutants exhibit reduced catalytic activity relative to wild type FRC (Table 1),<sup>12</sup> consistent with the idea that catalytically important loop conformations are no longer energetically preferred due to the alanine residue at either position. Perhaps because Gly-259 adopts backbone angles that are disallowed for alanine in both loop conformations seen in free FRC (Table S1, Supporting Information), the tetrapeptide loop segments in the G259A and G260A variants are computed to have a substantially lower preference for the “open” structure, lower barriers to conformational interconversion, and adopt an additional loop structure. These energetic changes may underlie the observed reductions in catalytic efficiency for these FRC mutants ( $k_{\text{cat}}/K_m$ ) (Table 1) because the tetraglycine loop stabilizes anhydride intermediates formed during catalytic turnover.<sup>8,12</sup> In an effort to calibrate the predictive power of our OSRW calculations on a firm structural basis, we superimposed the “open”, “closed”, and “intermediate” loop conformations computed for the G260A FRC variant (Figure 2) with the loop structure observed by X-ray crystallography (2VJN).<sup>12</sup> We were gratified to find extremely good agreement between the “intermediate” conformation and that seen in the G260A structure (Table S2, Supporting Information) *even though no information about this new loop conformation had been included in our OSRW simulations*. In addition, the noninterpolating reorientation of the Trp-48 side chain that takes place as the loop changes conformation was also correctly reproduced in these simulations. Both of these observations engender confidence in the computed thermodynamic properties. Efforts to correlate the OSRW free energy profile computed for G258A with experiment have been precluded by repeated failures to express this protein. The presence of (i) an intermediate state that is lower in energy than both the “open” and “closed” loop conformations and (ii) an increased barrier to interconversion predicts, however, that this mutant should exhibit little activity.

Finally, relaxation of the protein environment about the loop conformation at a given  $\lambda$  value in only a small number of time steps permits an assessment of side chain motions that are correlated



**Figure 2.** Cartoon showing superimposed active site tetrapeptide loops for the observed (2VJN) (cyan) and the “open” (blue), “intermediate” (black), and “closed” (red) loop conformations calculated for the G260A FRC mutant, together with the associated positions of the active site residues Trp-48, Tyr-59, and Asp-169.

with changes in loop structure. In the high-resolution X-ray crystal structure of wild type FRC, the side chain of Trp-48 undergoes a 90° rotation when the tetraglycine loop is in the “open” rather than the “closed” conformation. Detailed analysis of structures generated during the OSRW simulations showed the sampling algorithm to model side chain reorientation correctly, even though these motions were not defined explicitly in the reaction coordinate or “driven” by prior knowledge of their positions in the FRC crystal structure (Figure S3, Supporting Information). Because such changes represent motions that usually occur on a microsecond to millisecond time scale, the ability to obtain such information using nanosecond simulations is a significant advantage of the OSRW sampling strategy.

**Acknowledgment.** We thank the National Institutes of Health (DK61666) and the National Science Foundation (MCB0919983) for financial support.

**Supporting Information Available:** The full citation for ref 14, details of the theoretical and computational methods, Tables S1 and S2, and additional figures relevant to the simulations are given as Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Loria, J. P.; Berlow, R. B.; Watt, E. D. *Acc. Chem. Res.* **2008**, *41*, 214–221.
- (2) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638–1642.
- (3) Kollman, P. A.; Pearlman, D. A. *J. Chem. Phys.* **1989**, *91*, 7831–7839.
- (4) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- (5) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876–5886.
- (6) Zheng, L.; Chen, M.; Yang, W. *J. Chem. Phys.* **2009**, *130*, 234105.
- (7) Zheng, L.; Chen, M.; Yang, W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20227–20232.
- (8) Jonsson, S.; Ricagno, S.; Lindqvist, Y.; Richards, N. G. J. *J. Biol. Chem.* **2004**, *279*, 36003–36012.
- (9) Anatharam, V.; Allison, M. J.; Maloney, P. C. *J. Biol. Chem.* **1989**, *264*, 7244–7250.
- (10) Heider, J. *FEBS Lett.* **2001**, *509*, 345–349.
- (11) Ricagno, S.; Jonsson, S.; Richards, N.; Lindqvist, Y. *EMBO J.* **2003**, *22*, 3210–3219.
- (12) Berthold, C. L.; Toyota, C. G.; Richards, N. G. J.; Lindqvist, Y. *J. Biol. Chem.* **2008**, *283*, 6519–6529.
- (13) Toyota, C. G.; Berthold, C. L.; Gruez, A.; Jonsson, S.; Lindqvist, Y.; Cambillau, C.; Richards, N. G. J. *J. Bacteriol.* **2008**, *190*, 2556–2564.
- (14) MacKerell, A.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (15) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232–240.
- (16) Rosta, E.; Woodcock, H. L.; Brooks, B. R.; Hummer, G. *J. Comput. Chem.* **2009**, *30*, 1634–1641.

JA101446U