THE JOURNAL OF PHYSICAL CHEMISTRY B

Subscriber access provided by University of East Anglia Library

B: Biophysics; Physical Chemistry of Biological Systems and Biomolecules

Free Energy Profile of Domain Movement in Ligand-Free Citrate Synthase

Danilo Roccatano, and Steven Hayward

J. Phys. Chem. B, **Just Accepted Manuscript** • DOI: 10.1021/acs.jpcb.8b12001 • Publication Date (Web): 11 Feb 2019

Free Energy Profile of Domain Movement in Ligand-Free Citrate Synthase

Danilo Roccatano[a]* *and Steven Hayward*[b]

[a]School of Mathematics and Physics, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS, United Kingdom. [b]Computational Biology Laboratory, School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom.

ABSTRACT: Citrate synthase plays a fundamental role in the metabolic cycle of the cell. Its catalytic mechanism is complex involving the binding of two substrates that cause a domain movement. In this paper, we used classical molecular dynamics simulations and umbrella sampling simulations to determine the potential of mean force along a reaction coordinate for the domain movement in ligand-free citrate synthase from pig (*Sus Scrofa*). The results show that at 293 K, the closed-domain conformation has a ~4 k_bT higher energy than the open-domain conformation. In a simple two-state model, this difference means that the enzyme spends 98% of the time in the open-domain conformation ready to receive the substrate, oxaloacetate, rather than the closed-domain conformation where the binding site would be inaccessible to the substrate. Given that

experimental evidence indicates that the binding of oxaloacetate induces at least partial closure, this would imply an induced-fit mechanism which we argue is applicable to all enzymes with a functional domain movement for reasons of catalytic efficiency. A barrier of 4 k_bT gives an estimation of the mean first passage time in the range 1-10 μs.

INTRODUCTION

The relative movement of domains often occurs as part of the functional process in proteins.[1] The understanding of the mechanism governing their movements is therefore important for understanding function in these complex molecules, and, eventually, for exploiting them for medical and biotechnological applications[2]. Recent progress in experimental techniques based on neutron scattering methods[3], nuclear magnetic resonance[4,5], X-ray diffraction[6] and molecular dynamics (MD) simulations of biomolecules[7-8] are providing new insights into the details of the dynamics of these processes. In particular, the advance in parallel computing is filling the time-scale gap between theoretical simulations and experimental measurements, suggesting that MD simulations can act as a powerful *in-silico microscope*[9] for the study of these systems.

Citrate synthase (CS) is an essential enzyme in the cellular metabolism of both eukaryotic and prokaryotic organisms since it catalyzes the Claisen condensation reaction between the acetyl-coenzyme A and oxaloacetate yielding citrate and coenzyme A.[10] As one of the first enzymes to exhibit a significant, functionally related,

domain movement as determined from crystallographic structures, it has been used as an early protein paradigm for domain motions studies.[11-12] Several crystallographic structures of CS from different organisms have been solved. The homodimeric structures of pig and chicken heart CS are particularly interesting since they have been determined in the "open" and "closed" domain conformation, the former substrate free with the domains separated, the latter with both coenzyme A and citrate bound in the inter-domain cleft.[10] These structural studies and enzymatic kinetics measurements[13] have shown that the binding of the substrate (oxaloacetate) to the active site probably causes partial domain closure. [10] The binding of the acetyl-coenzyme A completes domain closure creating the favorable conditions in the active site for the catalysis of the reaction, and the formation of the citrate product. Upon completion of the enzymatic reaction, coenzyme A detaches from the enzyme surface and the product a molecule of citrate is released.

Theoretical studies of the protein[14] using the crystallographic open and closed structures of the protein have described the functional motion of the protein as a 19° rotation of the small domain with respect the large domain, around a hinge axis defined by two well-separated regions that act as mechanical hinges[15]. The first region is a β-hairpin (residues 56-65) which acts as a hinged loop,[15] the other being situated near the C-terminal end of the α-helix formed from residues 375-383. There are other inter-domain bending regions situated at some distance from the hinge axis, most notably the region 274-281 where a large change in the ψ-angle of His274 occurs. Being a comparison of two states (open and closed) of the enzyme, analysis of the X-ray

structures does not provide details about the pathway between them. Biochemical studies of the enzyme as well as the NMR studies[16] of domain motions suggest that the time scale of these movements can be on the order of microseconds, indicating that MD simulation studies can be used to address this. For CS several MD studies have been reported. In an early study, where only a few nanoseconds could be simulated, the opening process was seen to occur from the closed structure of the enzyme with both citrate and coenzyme A removed[17]. In this way, a non-equilibrium state, unable to maintain the catalytic competent closed conformation, underwent rapid spontaneous opening. Furthermore, the unbound enzyme simulated in the open state was not able to approach the closed state conformation within a few nanoseconds. These initial results suggested a barrier between the open- and closed-domain conformations and indeed this was confirmed in longer (50 ns) MD simulations performed more recently be Wells et al. [18] where again the unliganded, closed conformation opened spontaneously and did not return to the closed state. However, the limited sampling of the simulations prevented a thermodynamic estimation of the height of the barrier between the open and closed states.

In order to quantify the height of the barrier between the open and closed states, umbrella-sampling (US) simulations can be used to calculate the potential of mean force (PMF) along an appropriate reaction coordinate. Umbrella sampling has been applied to determine the PMF for a domain movement in Ribose Binding Protein (RBP)[19], in Lysine/Arginine/Ornithine-Binding Protein (LAOBP)[20], Maltose Binding Protein (MBP)[20 21] and in Biotin Carboxylase [22]. Here, more than 15 years after our early study,

we have extended our previous MD investigation by performing US simulations on the domain movement in CS to give an estimate of the energy barrier between the open and closed state in the absence of a ligand. The height of the barrier was used to estimate the mean first-passage time of the closure process. The results of this study provide new interesting insights on the thermodynamics and kinetics of the domain motion in the CS.

THEORETICAL CALCULATIONS

Starting coordinates. All the simulations have been conducted starting from crystallographic structures of CS from pig (*Sus Scrofa*) heart mitochondria obtained from the Protein Data Bank (PDB). For the unloaded open state, a more recent structure was used than previously (PDB entry 3ENJ[23]), and for the closed structure, one in complex with acetyl CoA and citrate was used (PDB entry 2CTS[24]). In the closed structure, acetyl CoA and citrate were removed and Ala32 was modeled to the native Val32 as in the open structure. For all simulations, the biological homodimeric molecule was used. The starting conformations had an N- terminal NH_3 group, and a C-terminal COO^- group. The protonation state of the residues in the protein was assumed to be the same as of the isolated amino acids in solution at pH 7.

Molecular dynamic simulations. The GROMOS96 43a1 force field[25] – also applied in our earlier studies of the same enzyme[26]- was used for all simulations. Protein structures were centered in an octahedral periodic box the size of which was set such

that the minimum distance between the protein and any side of the box was 0.80 nm. They were solvated by stacking equilibrated boxes of water molecules to completely fill the simulation box. Water molecules within a distance of 0.15 nm from protein atoms were removed. The simulations were performed using the SPC water model[27]. Sodium counter ions were added by replacing water molecules that had the most negative electrostatic potential to obtain a neutral simulation box. All bond lengths were constrained by the LINCS[28] algorithm and the SETTLE[29] algorithm was used for solvent molecules. Electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method.[30] For the long-range interactions, a grid spacing of 0.12 nm combined with a fourth-order B-spline interpolation was used to compute the potential and forces between grid points. A pair-list for non-bonded interactions within a cutoff of 1.3 nm was used and updated every 5 time-steps.

The simulated systems were first energy minimized, using the steepest descent algorithm, for at least 2000 steps in order to remove clashes between atoms. After energy minimization, initial velocities obtained from a Maxwell-Boltzmann velocity distribution at 293 K were assigned to all atoms. All systems were initially equilibrated for 100 ps with position restraint on the heavy atoms of the solute in order to relax the solvent molecules. The v-rescale thermostat[31] with a time constant of 0.1 ps was used to keep the temperature constant at 293 K. The pressure of the system was kept at 1 bar by using the Berendsen's barostat[32] with a time constant of 1 ps. A time step of 2 fs was used to integrate the equations of motions. Position restraints were then removed and systems were gradually heated from 50 K to 293 K in 200 ps. Two 1.2 μs simulations

(named OP1 and OP2, respectively) starting from the open conformation using different initial velocities at 293 K were performed. In addition, ten 25 ns long simulations were performed on systems starting from the closed conformation. All the simulations and analysis of the trajectories were performed by using the GROMACS (versions 5.2) software package.[33]

Rigid-body domain analysis. Rigid-body domain analysis was performed using the procedure previously described[17] where internal deformation within each domain was removed by superposing a reference structure on to it. This results in a trajectory of one rigid body moving relative to another rigid body. One can use this approach to calculate the Rigid-Body RMSD (RB RMSD) which quantifies extent of opening or closing relative to the experimentally determined open- or closed-domain conformation.

Umbrella sampling simulations. US was used to calculate the free energy profiles along a coordinate that joins the centers of mass of two regions, and which quantifies the extent of domain closure in subunit B. The first region is the small domain and the second region was determined from the crystallographic structures as follows. The open and closed dimer structures were superposed on their large domains and, considering one subunit, the line joining the centers of mass of the small domain in the open and closed structures was calculated. This line has the same direction as the displacement vector for the movement of the small domain and is oriented at 90.4° to the hinge axis direction indicating that movement along the line is a closure motion. Moving along this line one is able to find a group of residues in the large domain of the other subunit

that has a radius of gyration (1.27 nm calculated from a point on line) approximately equal to the radius of gyration of the small domain (1.35 nm calculated from the center of mass), comprises a similar number of residues (101 residues compared to 95 in the small domain), and has a center of mass close to the line (0.57 nm). These residues (1-15, 93-98, 109-119, 148-182, 255-262, 402-416) define the second region. Note that the two regions are in different subunits. The distance between the centers of mass of these two regions (see Figure 1) will be referred to as the inter-domain (ID) distance, *d*. In the open crystallographic structure, the ID distance is equal to 3.74 nm whereas, in the closed one, is equal to 3.21 nm.

A harmonic restraint with a force constant 3000 kJ/mol nm^2 was applied to the ID distance of subunit B from OP2 only. Twenty-six starting US configurations were extracted from the first 500 ns trajectory of simulations OP2. The ID distance ranged from 3.0 to 3.8 nm with a difference between two consecutive conformations less than 0.1 nm to ensure proper overlap of the histograms needed for correct calculation of the PMF profile. Each US configuration was simulated for 50 ns for a total simulation time of 1.3 μs. The ID's were sampled from the US trajectory every 0.2 ps. The weighted histogram analysis method (WHAM)[34] was used to calculate the PMF profile from histograms of the ID distributions.

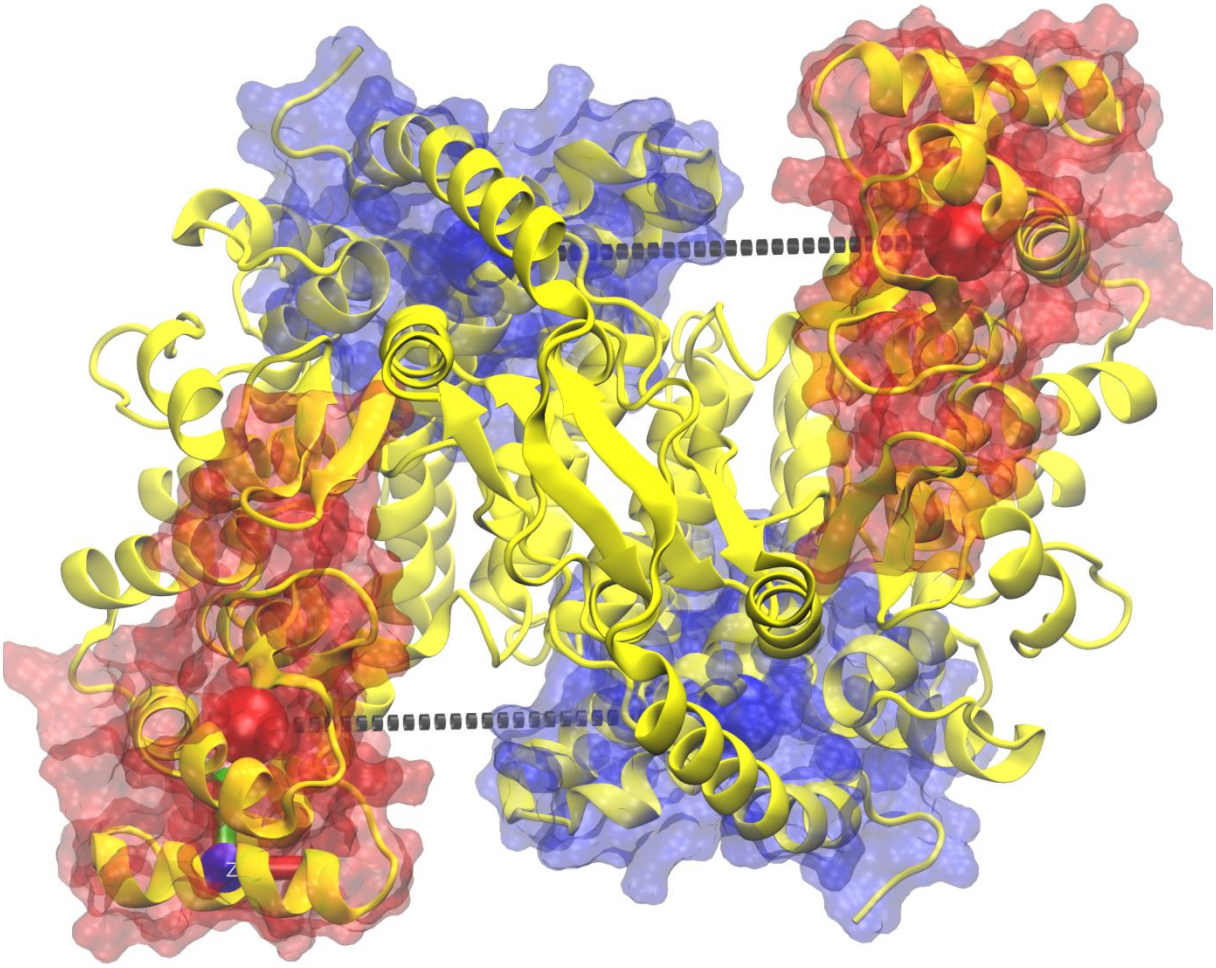


Figure 1: Crystallographic structure of CS in open conformation (3ENJ) showing the two regions in each unit (1st region in blue, 2nd in red) used to calculate the ID distance (dashed line in black). The colored spheres indicate the position of CoM of the regions represented as transparent surface.

Mean First Passage Time. The PMF obtained from the US calculations was used to make an estimation of the time scale for spontaneous transition from the open to closed state. The behavior of the domain motion along the reaction coordinate is approximated by the one-dimensional Fokker-Planck equation in the high-friction limit (viz. the

Smoluchowski equation).[35] In this limit which we know is valid for global collective motions,[36] and as such, domain motions , the mean first passage time (MFPT) is given by the relation[37]

$$\tau(\xi) = \frac{1}{D_0}\int_{\xi^o}^{\xi} dx\, e^{\beta U(x)} \int_{x}^{\xi^c} dx'\, e^{-\beta U(x')} \quad (1)$$

with $\beta = 1/k_B T$, D_0 is the self-diffusion coefficient[37] that is considered, as an approximation, constant along the reaction coordinate, and $U(x)$ the PMF obtained from the US simulation. The integrals are evaluated using the relative distance variable $\xi = -(d - d^o)$ defined over the region between the minimum of $U(x)$ in the open state $(d^o = 3.7\ nm)$ and the minimum of $U(x)$ in the closed state $(d^c = 3.0\ nm)$; thus integrals are calculated between $\xi^o(\xi = 0)$ and $\xi^c(\xi = 0.7\ nm)$.

We use the one-dimensional Einstein relation[38]

$$msd(t) = < [d(t) - d(0)]^2 > = 2D_0 t, \quad (2)$$

to estimate D_0 from the ID trajectories obtained from the free simulations OP1 and OP2. The average value indicated by the angular brackets was calculated using starting points separated by 1 ps along ID trajectory.[38-39]

RESULTS

The results presented in this section are organized as follows. First the results of two simulations, OP1 and OP2 that start from the open crystallographic structure are presented, revealing that subunit B in OP2 undergoes domain closure. The PMF profile along the ID reaction coordinate is presented which was calculated from US

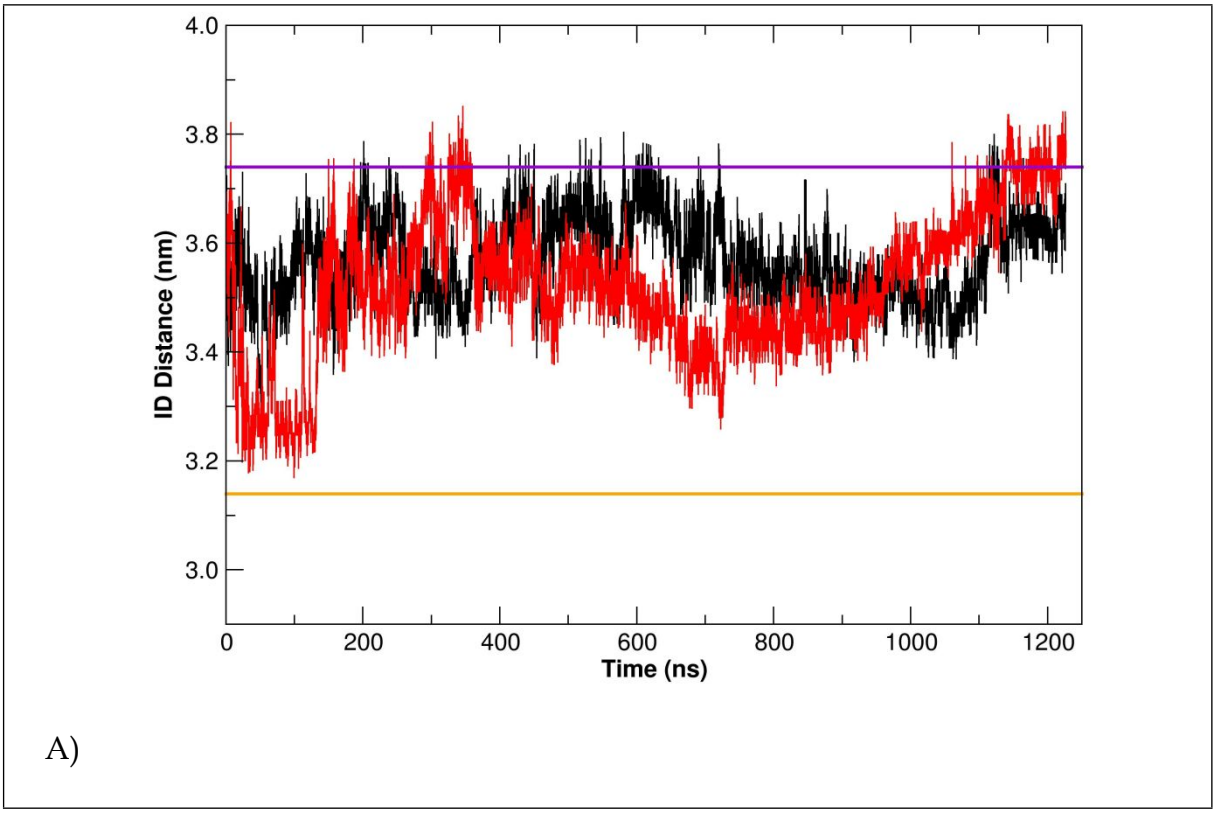
simulations performed using starting structures from the OP2 subunit B trajectory. The MFPT is then estimated using Eqn. 1 from the PMF profile and values of the D_0 estimated from OP1 and OP2 trajectories. Finally, results from MD simulations starting from the closed crystallographic structure with the ligands removed are presented.

MD simulations from the open structure. In Table S1, average values of the backbone root mean square deviation (RMSD with respect to the starting crystallographic structure of each subunit), the radius of gyration, α-helix and β-sheet content, and the solvent accessible surface area during the final 100 ns of the simulations are presented. The values of the backbone RMSD are similar in both simulations with a slightly higher value for subunit A in OP2. Other properties have fluctuations in value that are no more than ~7% of their value in the corresponding crystallographic structure. It is evident from the average values of the radius of gyration (Rg) reported in Table S1 that there is a systematic tendency of the subunits to become more compact with a maximum in the average deviation from the crystallographic structure of ~4%.

In Figure 2A and B, ID trajectories are shown. For OP1, subunit B tends to close up during the first 100 ns of the simulation with its ID almost reaching that of the closed crystallographic structure. It then opens again and fluctuates about an average distance of 3.5± 0.1 nm. For subunit A the ID fluctuates for most of the simulation time within a distance of 3.6± 0.1 nm. In both cases the equilibrium conformation is closer to the open crystallographic configuration (3.7 nm) by approximately 0.2 nm. In the OP2 simulation, after an increase in the ID beyond the crystallographic open structure, subunit B begins to close after 100 ns, reaching the closed crystallographic structure ID

value within 300 ns. It remains closed for the rest of the simulation with it ID fluctuating around the crystallographic closed value. Closure of subunit B in OP2 is also supported by the trajectory of the Rg of subunit B (see Figure S1(B)) which decreases to a value comparable to that of the crystallographic closed structure, and the trajectory of the RB RMSD (see Figure S2).

Subunit A undergoes large fluctuations within the first 400 ns reaching the closed crystallographic structure twice, but it then opens again stabilizing around 3.4 ± 0.1 nm.



A)

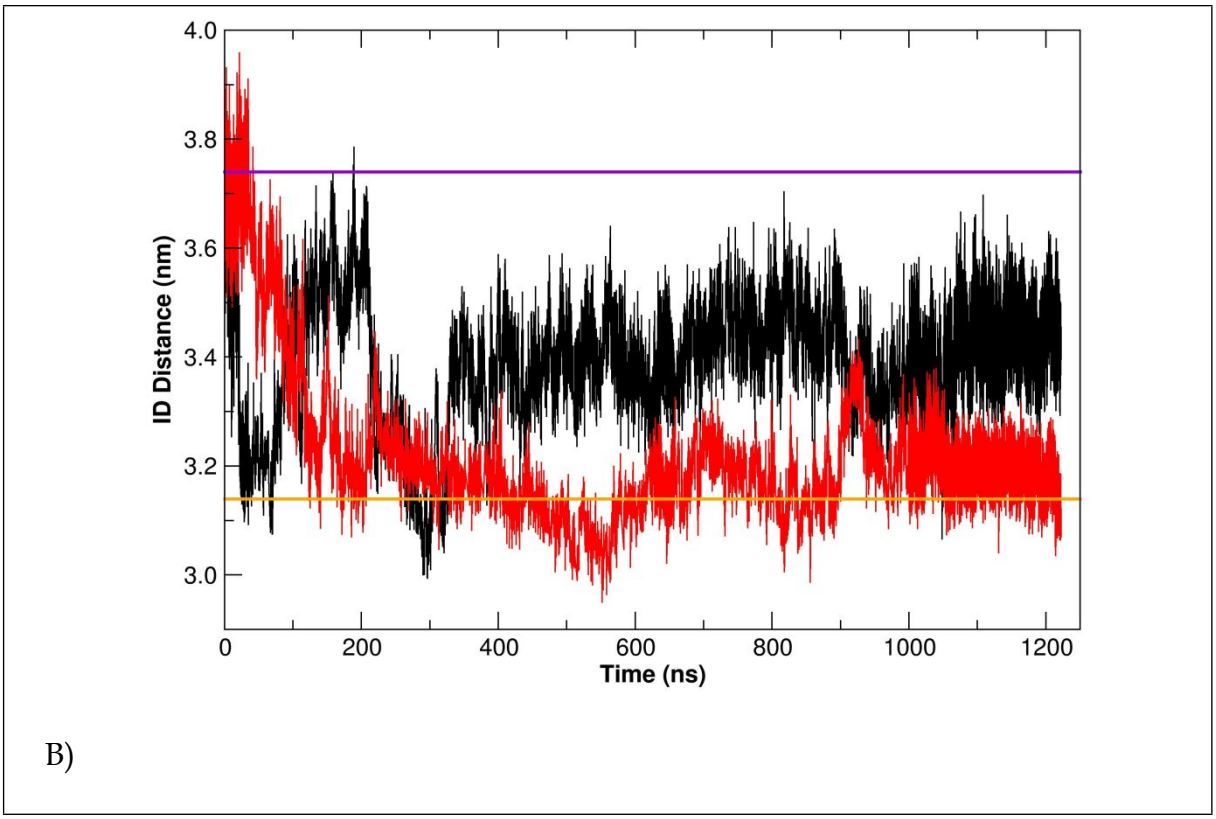


Figure 2: Trajectories of IDs from simulation OP1 (A) and OP2 (B). The two colors refer to subunit A (black) and B (red). The purple and orange lines indicate the values for the open and closed crystallographic structures, respectively.

In both simulations, neither of the subunit IDs stabilized around the crystallographic open state although this state is clearly reached as shown in the case of the subunit B in OP1 simulation where this region is explored mainly at the end of the simulation.

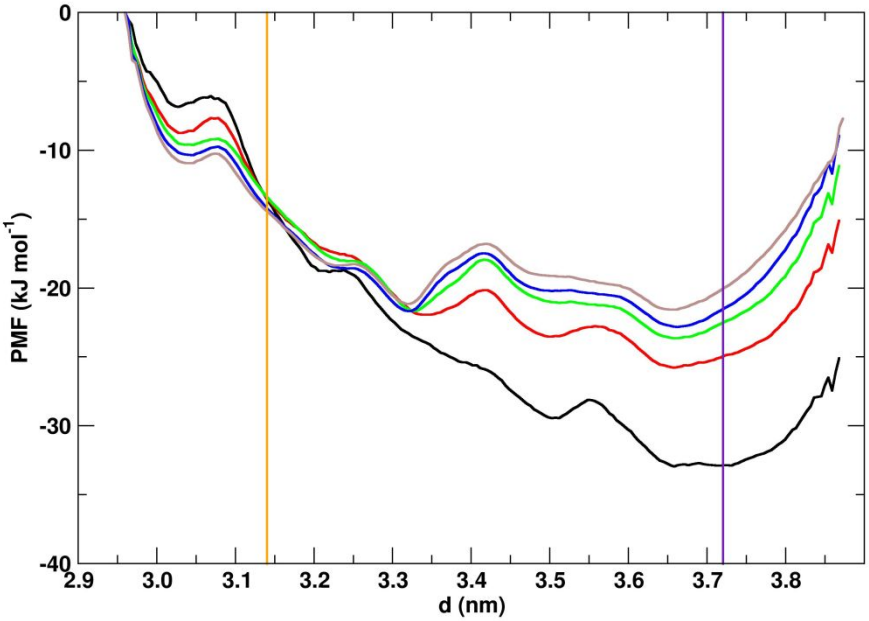
Umbrella sampling simulation. To obtain an accurate estimation of the free energy barrier involved in a spontaneous closing process, a US simulation was performed as described in the Method section, that is, using starting structures from along the closing trajectory of subunit B in OP2. The starting structures were selected based on their ID

values and in Supporting Information evidence is presented that confirms that these structures have domain conformations that span the open- and closed-domain conformations as defined by the open and closed X-ray structures.

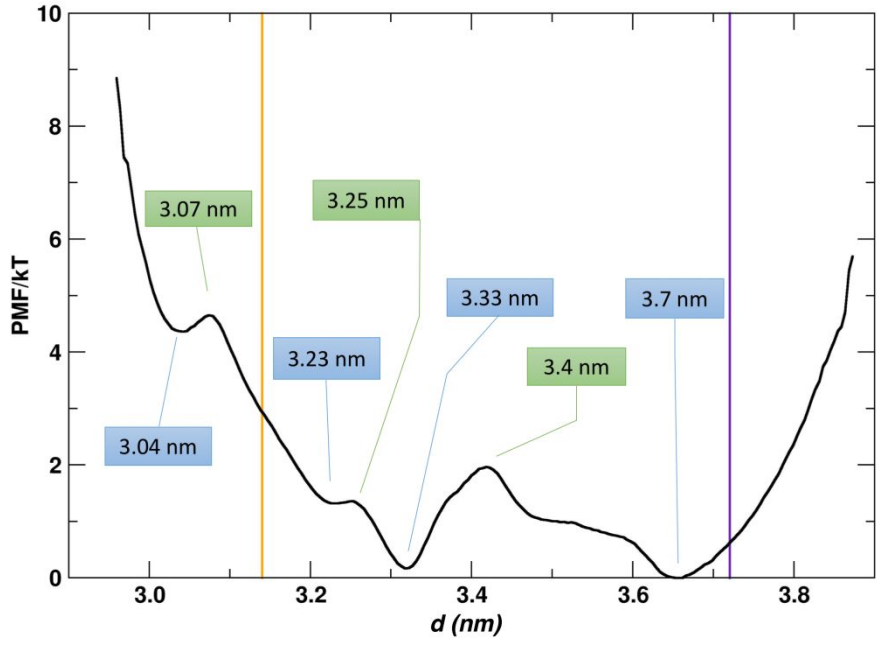
In Figure 3A, the PMF curves obtained at sampling times of 10, 20, 30, 40 and 50 ns are shown to confirm convergence of the PMF. In Figure 3B, the PMF curve obtained using the full 50 ns of sampling is shown. The value of the energy has been divided by k_bT=2.436 kJ/mol (k_b the Boltzmann constant and T=293K) for easier comparison. The PMF shows local minima slightly shifted from those observed in the open and closed crystal structures (purple and orange vertical lines). The PMF shows two barriers between the open and closed conformations. The energy difference between the open- and closed-domain conformation is estimated to be ~4 k_bT

.

Estimation of the MFPT. Using the calculated PMF, an estimation of the time scale of spontaneous closure was obtained by calculating the value of the MFPT along the reaction coordinate by the numerical integration of the Eqn. 1. The value of D_0 in (1) was estimated as described in the Method section from each of the ID trajectories in Figure 2. In Figure S3 we show the MSD curves and the straight lines obtained from linear regression, the gradient of which give D_0 (see Table 1).


PMF (kJ mol^{-1})
0
-10
-20
-30
-40
2.9
3.0
3.1
3.2
3.3
3.4
3.5
3.6
3.7
3.8
d (nm)

A)


PMF/kT
10
8
6
4
2
0
3.0
3.2
3.4
3.6
3.8
d (nm)
3.07 nm
3.25 nm
3.33 nm
3.7 nm
3.23 nm
3.4 nm
3.04 nm

B)

Figure 3. A) Convergence of the PMF curves with sampling times of 10 (black), 20(red), 30(green), 40(blue) and 50(brown) ns. B) The PMF curve obtained with 50 ns sampling was translated to have the zero-reference energy value at the minimum of the open structure. Energy values have been rescaled to units of k_bT. The blue and green backgrounds indicate the positions of the minima and maxima, respectively.

Table 1: Diffusion coefficients obtained from the ID trajectories, and the corresponding MFPT values at $\tau(\xi = 0.7)$.

Simulations(subunit)	$\mathbf{D_0}$ (nm^2/ns)	$\boldsymbol{\tau(\xi = 0.7)}$ (ns)
OP1 (subunit A)	0.22×10^{-4}	6645
OP1 (subunit B)	1.10×10^{-4}	1330
OP2 (subunit A)	0.14×10^{-4}	10440
OP2 (subunit B)	1.90×10^{-4}	770

The PMF from the US simulation was translated and mirrored along the x-axis to have the origin in the open-domain conformation with an energy equal to zero (see Figure 4). The boundary conditions at $\xi < 0$ and $\xi > 0.7$ were assumed to be repulsive. In Figure S4, MFTP curves calculated using Eqn. 1 with the four estimated values of D_0 (see Table 1) are shown. In Table 1, the values of $\tau(\xi = 0.7)$ are also given for comparison.

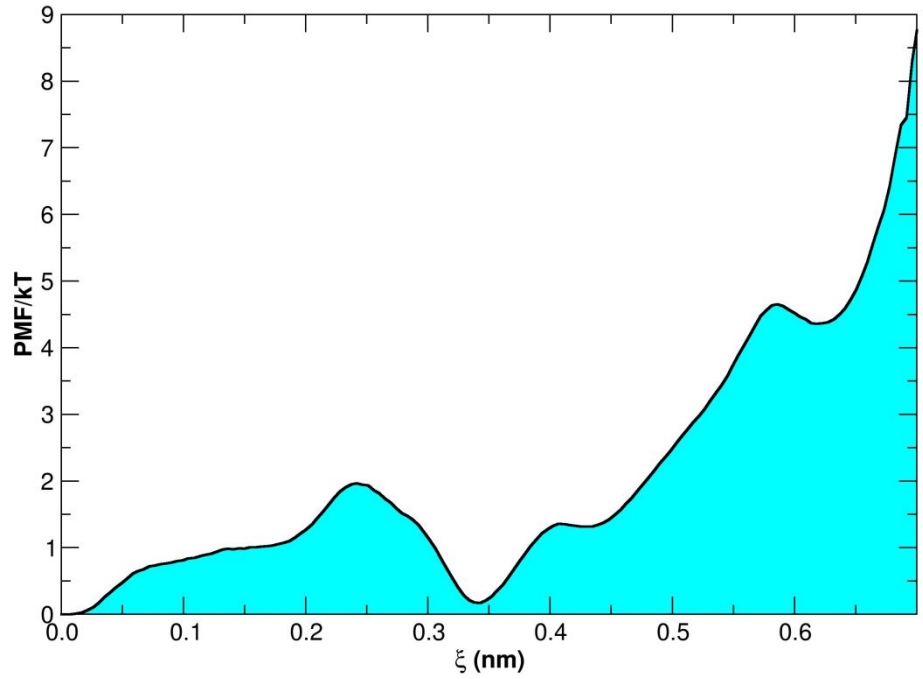


Figure 4: Calculation of the MFPT. The original PMF profile was first translated to the minimum located at 3.66 nm in the open-domain conformation region and then mirrored with respect the ordinate axis. The regions for $\xi < 0$ and $\xi > 0.70$ are considered both repulsive.

MD simulations from the closed structure. In Figure S5, the ID trajectories of both subunits are presented for all simulations starting from the closed-domain structure (CL) where all ligands have been removed. The results indicate that spontaneous opening can occur on the nanosecond timescale as seen in our previous work. In four simulations (CL3, CL5, CL7, CL8, CL9), the process occurs within a few nanoseconds. Interestingly, opening usually occurs in one of the two subunits, the other remaining closed (examples are simulations CL5, CL7 or CL8) indicating possible anticooperativity. A similar result was obtained in our previous work where we tracked

the trajectory of the RB RMSD of the domains[17] and anti-cooperative behavior in the opening and closing of the domains in the same enzyme (pig heart CS) was also observed by Wells et al. [18] . In some simulations one of the subunits shows no tendency to open as, for example, subunit B in simulation CL10; or both subunits remain predominantly closed within 25 ns as in CL1, CL2 and CL4. In some simulations, one subunit tends to close further than the observed crystallographic structure (e.g. subunit B in simulation CL5) increasing the compactness of the protein.

DISCUSSION

The use of microsecond time-scale MD simulations of biological molecules is expanding the capability of the MD simulations to study quantitatively domain motions in proteins. The result of two 1.2 μs unconstrained simulations of the CS reported in this paper revealed, in the limit of accuracy of the force field model adopted, the unexpected capability of the protein to exhibit domain closure even in the absence of the substrate. This enabled us to perform US from which the calculated energy barrier for the spontaneous closure at 293 K was estimated to be ~4 k_bT. This gives a MFPT of at least 10 μs. These values seem quite plausible for the model used although we expect that it is a lower limit since as observed for other proteins the mobility of domains and loops is faster compared to those found using NMR techniques.

To the best of our knowledge, experimental measurements on the timescale of domain motions in CS are not yet available. However, we think that the use of spectroscopic

FRET measurements could be a direction for a direct comparison of our simulation data[40-41].

In our previous MD study of CS[12, 17], we performed six 2ns simulations, three starting from the free open X-ray structure and three starting from the closed structure with all ligands removed. In that study we did not see a transition from open to closed, but for one simulation starting from the closed we did see a transition from closed to open. From those results we proposed a PMF profile that did resemble, in its general form, that seen in Fig3B, that is the closed structure being at a higher free energy than the open and a comparatively small barrier to cross in going from the closed to the open. MD simulations by Wells et al.[18] appeared to support this picture. The quantitative results presented here confirm this general profile. The barrier to overcome in going from open to closed is about 4 k_bT with no appreciable barrier between closed to open meaning the closed conformation is about 4 k_bT higher in energy than the open. For a simple two-state model where the enzyme is either in an open or closed state, a 4 k_bT difference means that CS will be in the open state 98% of the time, spending just 2% of the time in the closed state. The results can be compared with US simulations on binding proteins such as RBP [19], MBP and LAOBP [20] where for the free (Apo) protein there is a free-energy barrier to overcome in going from the open to closed and/or the closed conformation is at a higher free energy than the open. For liver alcohol dehydrogenase, MD simulations studies suggested that the well-known loop in that enzyme acts as a block to domain closure only moving out of the inter-domain cleft to enable domain closure when NAD^+ binds [42]. Although the result of umbrella sampling

on Biotin Carboxylase, [22] would appear to be a counter-example as the most stable state for the unliganded enzyme was found to be a closed state, for purpose of efficiency in enzymes with a domain movement, it makes sense to maintain the ligand-free enzyme in an open state ready to receive a substrate or coenzyme rather than being closed a significant proportion of time where the binding site in the inter-domain cleft would be inaccessible. If the ligand-free enzyme spends the majority of time in the open state and the ligand-bound enzyme is in the closed state, this would imply an induced-fit mechanism rather than pre-existing populations/conformational selection mechanism [43] for enzymes that have a functional domain movement. Kondo et al., [44] where for MBP, a second minimum situated towards the crystallographic closed structure was named the "semi-open" state, suggested that the binding of the ligand to this semi-open state was a possible indication of conformational selection within an induced-fit "framework". One might suggest that here we have a similar scenario with a semi-open state situated at the d = 3.33 nm minimum (see Figure 3B). Perhaps this is a matter of degree and interpretation but for an enzyme with a functional domain movement the induced fit mechanism makes sense in the context of catalytic efficiency. The pertinent question for enzyme mechanism is how the binding of a substrate or coenzyme removes the barrier to closure and tips the balance in favor of the closed-domain conformation.

The experimental barrier for the catalytic process involving deprotonation of acetyl-coenzyme A has been estimated to be about 37.7 kJ/mol[45] from experiment and accurate QM calculations. This is a 3.9 times larger barrier than the one for the

spontaneous closure of the domains. From the biochemical data, it is also evident from the measured turnover of the enzyme, that the product is processed at the millisecond time scale.[46] Therefore, we expect that once the substrate and the Acetyl-CoA are in place, the chemical reaction and release of the product will be rate determining.

CONCLUSIONS

In this work, we have calculated the free energy barrier for the domain movement in ligand-free citrate synthase from *Sus Scrofa.* at 293 K. The calculations indicate a difference of 9.7 kJ/mol (~4 k_bT) between the high-energy closed-domain and low-energy open-domain conformational state. The energy barrier allows giving an estimation of the mean first passage time in the range 1-10 μs. In a simple two-state model, this difference means that the enzyme spends 98% of the time in the open-domain conformation state more accessible to the substrate binding than the closed-domain conformation. Given that experimental evidence indicates that the binding of substrate oxaloacetate induces at least partial closure, this would imply an induced-fit mechanism which we argue is applicable to all enzymes with a functional domain movement for reasons of catalytic efficiency.

ASSOCIATED CONTENT

Supporting Information

US starting structures span crystallographic open- and closed-domain structures. Average structural properties of the OP1 and OP2 simulations; time series of the radius of gyration for the OP1 and OP2 simulations; time series of the RB RMSD of subunit B from the OP2 simulation; mean square displacement along the ID distances for the two subunits in the OP1 and OP2 simulations; MFPT curves calculated using the different values of diffusion coefficients (D_0); time series of ID distances for the simulations starting from the closed crystallographic structure.

AUTHOR INFORMATION

Corresponding Author

* E-mail: droccatano@lincoln.ac.uk

ORCID

Danilo Roccatano: 0000-0002-8495-3815

Steven Hayward: 0000-0001-6959-2604

ACKNOWLEDGMENT

The authors thank the ARCHER supercomputer centre at the University of Edinburgh for granting the access of the facilities within the ARCHER Instant Access Project "Thermodynamics analysis of domain motion".

REFERENCES

1. Williamson, M. P., *How proteins work*. Garland Science: New York, **2012**.
2. Murphy, W. L., Emerging area: biomaterials that mimic and exploit protein motion. *Soft Matter* **2011,** *7* (8), 3679-3688.
3. Smith, J. C.; Tan, P.; Petridis, L.; Hong, L., Dynamic neutron scattering by biological systems. *Annu. Rev. Biophys.* **2018,** *47*, 335-354.
4. Boehr, D. D.; Dyson, H. J.; Wright, P. E., An NMR perspective on enzyme dynamics. *Chem. Rev.* **2006,** *106* (8), 3055-3079.
5. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.
6. Kupitz, C.; Olmos, J. L., Jr.; Holl, M.; Tremblay, L.; Pande, K.; Pandey, S.; Oberthur, D.; Hunter, M.; Liang, M.; Aquila, A., et al., Structural enzymology using X-ray free electron lasers. *Struct . Dyn.* **2017,** *4* (4), 044003.
7. van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glattli, A.; Hunenberger, P. H., et al., Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Edit.* **2006,** *45* (25), 4064-4092.
8. Adcock, S. A.; McCammon, J. A., Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006,** *106*, 1589-1615.
9. Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E., Biomolecular simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.* **2012,** *41*, 429-52.
10. Remington, S. J., Structure and mechanism of citrate synthase. *Curr. Top. Cell. Regul.* **1992,** *33*, 209-229.
11. Lesk, A. M.; Chothia, C., Mechanisms of domain closure in proteins. *J. Mol. Biol.* **1984,** *174*, 175-191.

12. Hayward, S.; Berendsen, H. J. C., Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct., Funct., Genet.* **1998,** *30*, 144-154.
13. Wiegand, G.; Remington, S. J., Citrate synthase, structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **1986,** *15*, 97-117.
14. Hayward, S.; Berendsen, H. J. C., Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins-Structure Function and Genetics* **1998,** *30* (2), 144-154.
15. Hayward, S., Structural principles governing domain motions in proteins. *Proteins: Struct., Funct., Genet.* **1999,** *36*, 425-435.
16. Tang, C.; Schwieters, C. D.; Clore, G. M., Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **2007,** *449* (7165), 1078-U12.
17. Roccatano, D.; Mark, A. E.; Hayward, S., Investigation of the mechanism of domain closure in citrate synthase by molecular dynamics simulation. *J. Mol. Biol.* **2001,** *310* (5), 1039-1053.
18. Wells, S. A.; van der Kamp, M. W.; McGeagh, J. D.; Mulholland, A. J., Structure and function in homodimeric enzymes: simulations of cooperative and independent functional motions. *Plos One* **2015,** *10* (8): e0133372.
19. Ravindranathan, K. P.; Gallicchio, E.; Levy, R. M., Conformational equilibria and free energy profiles for the allosteric transition of the ribose-binding protein. *J. Mol. Biol.* **2005,** *353* (1), 196-210.
20. Kondo, H. X.; Okimoto, N.; Morimoto, G.; Taiji, M., Free-energy landscapes of protein domain movements upon ligand binding. *J. Phys. Chem. B* **2011,** *115* (23), 7629-7636.
21. Mascarenhas, N. M.; Kastner, J., How maltose influences structural changes to bind to maltose-binding protein: Results from umbrella sampling simulation. *Proteins: Struct., Funct., Genet.* **2013,** *81* (2), 185-198.

22. Novak, B. R.; Moldovan, D.; Waldrop, G. L.; de Queiroz, M. S., Umbrella sampling simulations of biotin carboxylase: is a structure with an open ATP grasp domain stable in solution? *J. Phys. Chem. B* **2009,** *113* (30), 10097-10103.
23. Larson, S. B.; Day, J. S.; Nguyen, C.; Cudney, R.; McPherson, A., Structure of pig heart citrate synthase at 1.78 Å resolution. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **2009,** *65* (5), 430-434.
24. Remington, S.; Wiegand, G.; Huber, R., Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. *J. Mol. Biol.* **1982,** *158* (1), 111-152.
25. van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hunenberger, P. H.; Kruger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G., *Biomolecular Simulation: the GROMOS96 manual and user guide.* **1996**; Vol. 1.
26. Roccatano, D.; Mark, A. E.; Hayward, S., Investigation of the mechanism of domain closure in citrate synthase by molecular dynamics simulation. *J. Mol. Biol.* **2001,** *310*, 1039-1053.
27. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., Interaction models for water in relation to protein hydration. *Intermolecular Forces* **1981**, 331-342.
28. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.
29. Miyamoto, S.; Kollman, P. A., Settle - an analytical version of the shake and rattle algorithm for rigid water models. *J. Comput. Chem.* **1992,** *13* (8), 952-962.
30. Darden, T.; York, D.; Pedersen, L., Particle mesh ewald - an Nlog(N) method for ewald sums in large systems. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
31. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126* (1), 014101.

32. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Di Nola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.
33. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008,** *4* (3), 435-447.
34. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted Histogram analysis method for free-energy calculations on biomolecules .1. The method. *J. Comput. Chem.* **1992,** *13* (8), 1011-1021.
35. Zwanzig, R., *Nonequilibrium statistical mechanics*. Oxford University Press: Oxford; New York, **2001**.
36. Hayward, S.; Kitao, A.; Hirata, F.; Go, N., Effect of solvent on collective motions in globular protein *J. Mol. Biol.* **1993,** *234* (4), 1207-1217.
37. Weiss, G. H., First passage time problems in chemical physics. In *Advances in Chemical Physics*, **1967**; pp 1-18.
38. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*. Oxford University Press: **1989**.
39. Haile, J. M., *Molecular dynamics simulation : elementary methods*. Wiley: New York, **1992**.
40. Okamoto, K.; Sako, Y., Recent advances in FRET for the study of protein interactions and dynamics. *Curr. Opin. Struct. Biol.* **2017,** *46*, 16-23.
41. Aviram, H. Y.; Pirchi, M.; Mazal, H.; Barak, Y.; Riven, I.; Haran, G., Direct observation of ultrafast large-scale dynamics of an enzyme under turnover conditions. *Proc. Natl. Acad. Sci. U S A* **2018,** *115* (13), 3243-3248.
42. Hayward, S.; Kitao, A., Molecular dynamics simulations of NAD(+)-induced domain closure in horse liver alcohol dehydrogenase. *Biophys. J.* **2006,** *91* (5), 1823-1831.

43. Ma, B. Y.; Shatsky, M.; Wolfson, H. J.; Nussinov, R., Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* **2002,** *11* (2), 184-197.

44. Kondo, H. X.; Okimoto, N.; Morimoto, G.; Taiji, M., Free-energy landscapes of protein domain movements upon ligand binding. *J. Phys. Chem. B* **2011,** *115* (23), 7629-7636.

45. van der Kamp, M. W.; Żurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J., Testing high-level QM/MM methods for modeling enzyme reactions: acetyl-CoA deprotonation in citrate synthase. *J. Phys. Chem. B* **2010,** *114* (34), 11303-11314.

46. Kurz, L. C.; Shah, S.; Frieden, C.; Nakra, T.; Stein, R. E.; Drysdale, G. R.; Evans, C. T.; Srere, P. A., Catalytic strategy of citrate synthase: subunit interactions revealed as a consequence of a single amino acid change in the oxaloacetate binding site. *Biochemistry* **1995,** *34* (41), 13278-13288.

TOC Graphic