Transport cycle of *Escherichia coli* lactose permease in a nonhomogeneous random walk model

Yan B. Barreto, Matheus L. Rodrigues, and Adriano M. Alencar
Instituto de Física, Universidade de São Paulo, 05508-090 São Paulo, São Paulo, Brazil

(Received 28 August 2018; revised manuscript received 8 April 2019; published 22 May 2019)

We present Monte Carlo simulations for the transport cycle of *Escherichia coli* lactose permease (LacY), using as a starting point the model proposed by Kaback *et al.* [Nat. Rev. Mol. Cell Biol. **2**, 610 (2001)], which is based on functional properties of mutants and x-ray structures. Kaback's model suggests the existence of six states for the whole cycle of lactose-H^+ symport. However, the free-energy differences between these states have not yet been reported in the literature. Here, we analyzed the biochemical structure of each state and determined a range of possible values for each one of the five free-energy variations. Then, using the Metropolis algorithm in a nonhomogeneous random walk model, we tested all the possible combinations with these values to find the free-energy curve that best reproduces the dynamics of LacY. The agreement between our model predictions and the experimental data suggests that our free-energy curve is appropriate for describing the lactose-H^+ symport. We found not only this curve, but also the time of occupancy of the permease at each conformation. In addition, we paved the way in this work to solve an open question related to this transport mechanism, which is the importance of protonation for lactose binding.

DOI: 10.1103/PhysRevE.99.052411

I. INTRODUCTION

Escherichia coli lactose permease (LacY) is a transport protein that facilitates the passage of lactose across the cell membrane of *E. coli*. This protein utilizes the free energy stored in a proton gradient, $\Delta\mu_{H^+}$, and an electrostatic membrane potential, $\Delta\Psi$, to drive the transport of β-galactosides, such as lactose, against its concentration gradient, $\Delta\mu_L$. Conversely, in the absence of $\Delta\mu_{H^+}$, LacY can also utilize the free energy stored in a lactose concentration gradient to drive the transport of H^+ with the generation of an electrochemical proton gradient, $\Delta\tilde{\mu}_{H^+} = \Delta\mu_\Psi + \Delta\mu_{H^+}$ [1]. Since its first report in 1955 by Cohen and Rickenberg, LacY has been extensively studied to elucidate the molecular basis of its transport process [2]. One of the main results of these studies suggests that, despite lactose permease being composed of 417 amino-acid residues, only six are irreplaceable for the active transport of lactose: Glu126, Arg144, Glu269, Arg302, His322, and Glu325 [3]. In 2001, Kaback *et al.* proposed the currently accepted model for the lactose-H^+ transport mechanism [4].

This model has six steps, starting with unprotonated LacY in the outward-facing (C_{Out}) conformation (Fig. 1, State 1). In this very unstable state, LacY is protonated immediately ($1 \rightarrow 2$), sharing the H^+ between the residues His322 and Glu269 (Fig. 1, State 2). Then, one lactose molecule binds to LacY ($2 \rightarrow 3$), being recognized by the charge pair between Arg144 and Glu126 (Fig. 1, State 3). After substrate binding, a rapid transition to the inward-facing (C_{In}) conformation occurs ($3 \rightarrow 4$). This transition is associated with the hydrogen bond formed between Glu269 and Arg144 (Fig. 1, State 4). Additionally, the H^+ is transferred to Glu325. The lactose is then released into the cytoplasm ($4 \rightarrow 5$), which triggers a conformational change that allows Arg302 to approximate protonated Glu325, resulting in deprotonation ($5 \rightarrow 6$). After releasing the H^+, a transition into the C_{Out} conformation is induced ($6 \rightarrow 1$). This transport mechanism corresponds to the transport of lactose in the influx direction, clockwise. However, transport in the opposite direction, counterclockwise, can also be explained by this model, since influx and efflux are functionally symmetric processes [5].

Despite the fact that this model exists for the transport mechanism of LacY, which is consistent with the experimental observations [6–14], two questions still remain open. These questions, according to Guan and Kaback [15], are as follows: (i) What is the time of occupancy of LacY in the C_{Out} and C_{In} conformations? (ii) Why is protonation important for lactose binding? The nonhomogeneous random walk model that we proposed here for the lactose-H^+ symport answers these open questions and offers a theoretical framework to describe experimental data.

II. METHODS

Our starting point is the fact that the free-energy curve for the transport cycle represented in Fig. 1(a) is not established in the literature. The only information we have is that the State 2 is the ground state [5]. Thus, the following procedure was adopted to obtain it. First, we assume that residue-residue interactions are hydrogen bonds, which range between 2 and 5 k_BT, and proton-residue and lactose-residue interactions are Van der Waals bonds, which range between 1 and 2 k_BT. This assumption is based on a detailed description of the interactions among the elements of Kaback's model [15] and the literature data on hydrogen bonding in proteins [16]. Then, on the basis of the difference in the number of hydrogen and Van der Waals bonds between the states of LacY and the ground state, which can be easily obtained from Fig. 1(a), we determine a range $(0, \Delta E_S^{max}]$ of possible values for each one of the five free-energy variations $\Delta E_S = E_S - E_2 > 0$, where S are the states. After that, we incorporate into these

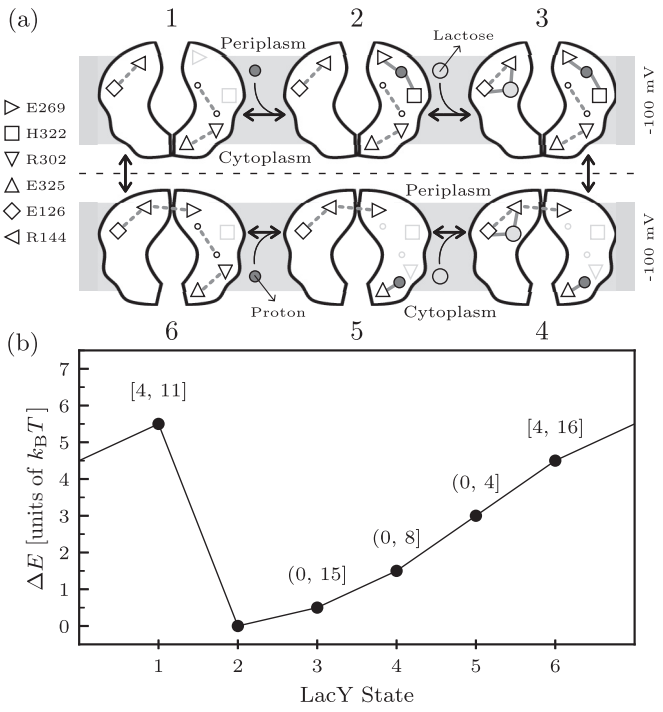

FIG. 1. Mechanism of lactose transport by LacY. (a) Key residues are identified. An electrostatic membrane potential of −100 mV is indicated. Hydrogen and Van der Waals bonds are shown as broken and solid lines, respectively. Also shown is the hydrogen bond between the residues Asp240 and Lys319, which are represented as small white circles. These residues are not essential to Kaback's model, but they do affect our free-energy curve (for details, see Ref. [4]). The lactose and the H^+ are represented as light and dark gray circles, respectively. (b) Free-energy variations between the states of LacY and the ground state. Also shown are the ranges of possible values for each $\Delta E_S = E_S - E_2$ that were used to find the best free-energy curve.

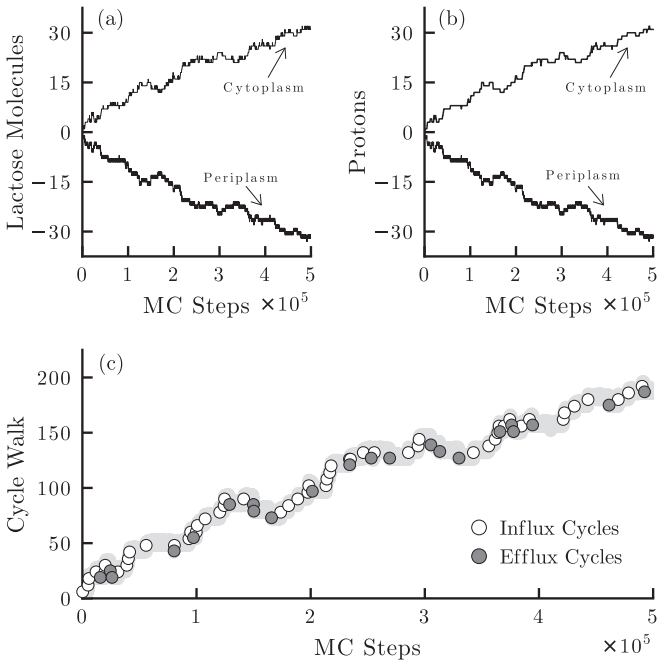

FIG. 2. Dynamics of lactose transport by LacY. In this simulation, the probabilities of finding lactose or H^+ near the entrance of LacY are non-time-dependent. (a) Number of lactose molecules in the cytoplasm and periplasm. (b) Number of protons in the cytoplasm and periplasm. These two plots look identical due to the coupling between the lactose and H^+ transport. (c) Completed influx (total of 53) and efflux (total of 22) cycles and the corresponding MC step in which each one was concluded. At the beginning of the simulation, the variable cycle walk is equal to zero, and we add +1 to each step in the influx direction, and −1 to each step in the efflux direction, where each one of these unit steps represents 1/6 of a full cycle.

ranges the effect of an electrostatic membrane potential $\Delta\Psi$, which in all living cells is typically around −100 mV [17], corresponding to an electrical gradient of $\Delta\mu_\Psi = e\,\Delta\Psi \approx -4\,k_B T$, where e is the charge of one H^+. This can be done by mapping the ranges of the protonated states ($S = \{2, 3, 4, 5\}$) onto $(-4\,k_B T, \Delta E_S^{\max} - 4\,k_B T]$, or, equivalently, by mapping the ranges of the unprotonated states ($S = \{1, 6\}$) onto $[4\,k_B T, \Delta E_S^{\max} + 4\,k_B T]$. For convenience, we chose the latter option, since it has the advantage of keeping the ground state unchanged. The resulting ranges are indicated in Fig. 1(b).

Next, we consider from each range discrete values equally spaced by $\delta\,k_B T$, where δ is a free parameter. Larger values of δ, above 1, reduce considerably the precision of our method. Small values of δ, below 0.25, generate too many free-energy curves to be tested, and are well above the precision of the experimental data available in the literature. Thus, here we use $\delta = 0.5$. We tested all the possible combinations of these values and discarded the combinations that do not follow the three boundary conditions (BCs) established in the literature that LacY must follow [15]. The first BC is when there is a high $\Delta\tilde{\mu}_{H^+}$ pointing from the cytoplasm to the periplasm. In this condition, the influx of lactose must predominate over the efflux. To implement this BC and constrain the free-energy landscape, we carried out simulations with $\Delta\tilde{\mu}_{H^+} \ll 0$ and counted, for each possible combination, the number of completed influx and efflux cycles, as shown in Fig. 2(c). After that, we discarded the combinations in which the number of efflux cycles is greater than the number of influx cycles on average. Second, the rate-limiting step of the transport cycle when $\Delta\mu_{H^+} = 0$ must be one of the deprotonations, $2 \to 1$ or $5 \to 6$, depending on the sign of $\Delta\mu_L$. This BC was implemented by performing simulations with $\Delta\mu_{H^+} = 0$ for the combinations that were not discarded in the first BC. In each simulation, the rate-limiting step was determined by the occupancy time of LacY in each one of its six states, as shown in Fig. 3(c). We discarded the combinations in which the rate-limiting step was not one of the deprotonations. Finally, the third BC is that, when $\Delta\mu_{H^+} = 0$, the flux of lactose must be controlled exclusively by $\Delta\mu_L$. This was implemented by setting equal concentrations of H^+ in both sides of the membrane, say C_{H^+}, and then testing if the lactose fractions in the equilibrium, as shown in Fig. 3(a), keep unchanged for $0.1\,C_{H^+}$, $10\,C_{H^+}$, and $100\,C_{H^+}$. We discarded the combinations in which the lactose fractions changed significantly, more than 5%. After this, we chose among the few remaining combinations the one that presented the smallest number of completed efflux cycles on average when submitted to a $\Delta\tilde{\mu}_{H^+} \gg 0$,

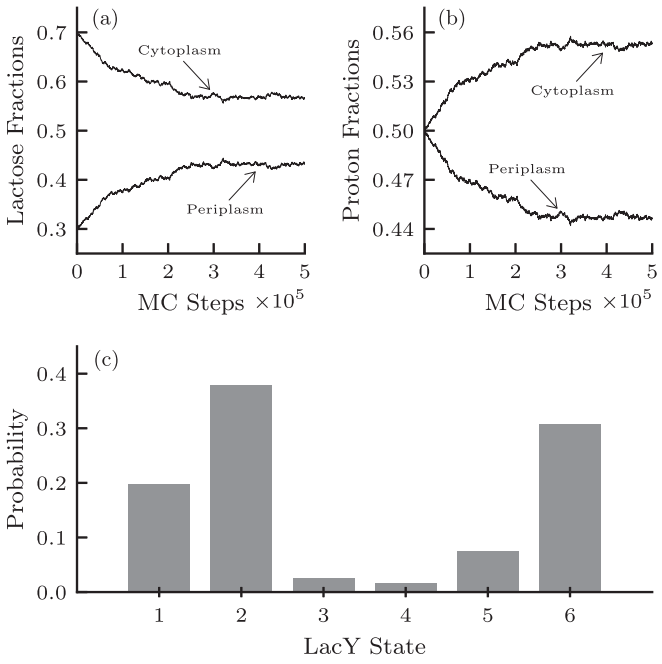

FIG. 3. Dynamics of lactose transport by LacY. In this simulation, the concentrations of substrates in the cytoplasm and periplasm can vary in time. (a) Lactose fractions in the cytoplasm and periplasm. (b) Proton fractions in the cytoplasm and periplasm. (c) Probabilities of the system found in each one of its six states.

since the efflux of lactose should be null under this condition. The resulting free-energy curve is shown in Fig. 1(b).

Our system consists of a cytoplasm and a periplasm with volumes V_c and V_p, respectively, separated by a cell membrane with one embedded LacY. In each simulation, we define an initial number of lactose molecules in the cytoplasm and periplasm, N_c^L and N_p^L, respectively, and an initial number of protons in the cytoplasm and periplasm, $N_c^{H^+}$ and $N_p^{H^+}$, respectively. The total energy available to do work in this system is given by the change in the Gibbs free-energy across the membrane, $\Delta G = \Delta\mu_{H^+} + \Delta\mu_\Psi + \Delta\mu_L$, where the H^+ and lactose gradients are given by [17]

$$\frac{\Delta\mu_{H^+}}{k_B T} = \ln\left(\frac{N_c^{H^+}/V_c}{N_p^{H^+}/V_p}\right) \tag{1}$$

and

$$\frac{\Delta\mu_L}{k_B T} = \ln\left(\frac{N_c^L/V_c}{N_p^L/V_p}\right). \tag{2}$$

In our simulation, we assume that the probability of finding one substrate near the entrance of LacY is proportional to its concentration, e.g., $\Pi_c^L \propto N_c^L/V_c$. At each MC step, LacY has the same probability of moving in the clockwise and counterclockwise direction. However, this movement can be frustrated either by the concentration limitation, or by the Boltzmann factor, described below.

This model that we proposed here is a nonhomogeneous random walk, where the transition probabilities of the Markov chain are determined by the free-energy curve in Fig. 1(b) and the concentrations of substrates in the cytoplasm and periplasm [18]. The elements of the system interact using the Metropolis algorithm [19], a Markov chain Monte Carlo method widely used in reproducing the dynamics of biological systems [20–22]. In this algorithm, each state transition, say $S \to S'$, is accepted with a probability determined by the activation energy between these states, ΔE_a. Transitions with $\Delta E_a \leqslant 0$ are accepted with probability 1, whereas transitions with $\Delta E_a > 0$ are accepted with probability $\exp(-\beta\Delta E_a)$, where $\beta = 1/k_B T$, k_B is the Boltzmann constant, and T is temperature. In our model, we assume that ΔE_a is equal to the energy difference between the states S and S'. This assumption is necessary since there are not enough data in the literature to estimate the ΔE_a of the six transitions in Kaback's model. Our simulation code in the PYTHON language is presented in the Supplemental Material [23].

III. RESULTS AND DISCUSSIONS

We carried out the simulations for three cases. First, we consider an open system in which the substrate concentrations are constant, independent of time. For this case, we assume a cytoplasmic and a periplasmic pH of 7.6 and 5.6, respectively [24]. This periplasmic pH is consistent with the acid environments where *E. coli* can be found, such as the human gastrointestinal tract. Thus, it follows that $\Pi_p^{H^+} = 100\,\Pi_c^{H^+}$. Second, we consider a closed system in which the concentrations of substrates in the cytoplasm and periplasm can vary in time. In this case, to test the behavior of our model in the absence of $\Delta\mu_{H^+}$, the system was initialized with 50% of the protons and 70% of the lactose molecules in the cytoplasm. Third, we validate our model by reproducing two experiments from Ujwal *et al.* [25].

The results for the first simulation are shown in Fig. 2. Here, the electrochemical proton gradient is pointing from the cytoplasm to the periplasm, while the lactose concentration gradient is pointing in the opposite direction. Thus, it is possible to see that the uphill accumulation of lactose [see Fig. 2(a)] is coupled to the downhill movement of H^+ [see Fig. 2(b)]. Note that, despite the fact that these two plots look identical, which is due to the coupling between the lactose and H^+ transport, they are not equal, since there are rare events in which a lactose is transported but the H^+ is not. This happens, for instance, when the cycle goes forward from State 2 to 5 and then moves backward to State 2. Importantly, the system was not expected to evolve toward equilibrium, since we fixed the concentrations of substrates. In Fig. 2(c), we show all the completed influx and efflux cycles and the corresponding MC step in which each one was concluded. Based on that, the turnover number for LacY can be calculated as the absolute difference between the number of completed influx and efflux cycles divided by the total number of MC steps. The result obtained by using 10^3 replicates is $(5.3 \pm 0.4) \times 10^{-5}$ (MC steps)$^{-1}$. Thus, considering the turnover number of 16–20 s^{-1} that was estimated for active lactose transport in Ref. [26], one can conclude that one MC step in our model is equivalent, on average, to 2.9 μs. We will use this estimated physical time scale to describe the experimental data below.

Next, we present the results for the second simulation in Fig. 3. In this case, the system is initialized with a lactose concentration gradient pointing from the periplasm to the

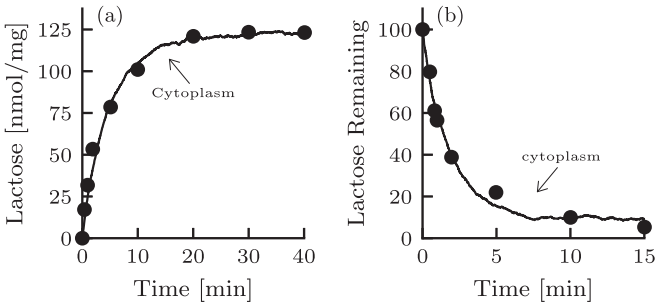

FIG. 4. Time courses of lactose transport by LacY. Circles correspond to experimental data extracted from [25], while solid lines correspond to our simulations. Here, one Monte Carlo step is equivalent to 2.9 μs. (a) Evolution in time of lactose influx. (b) Evolution in time of lactose efflux.

cytoplasm, while $\Delta\mu_{H^+} = 0$. As one can see, the downhill movement of lactose, as shown in Fig. 3(a), initially provides energy to transport protons to the cytoplasm, as shown in Fig. 3(b), which causes the generation of a $\Delta\mu_{H^+} \neq 0$. This proton gradient increases until the system reaches equilibrium (2.5×10^5 MC steps), when the initial downhill translocation of lactose is compensated by its influx.

The probabilities of the system being found in each one of its six states are shown in Fig. 3(c). Based on that, we can discuss the two questions by Guan and Kaback. First, one can see from this figure that LacY spent approximately 60% of its time in the C_{Out} conformation and 40% in the C_{In} conformation. Second, it is notable that the system spent more time in State 2 (38%) than in any other state. We associate this fact with two characteristics of our model: (i) There is a low probability of finding a lactose near the entrance of LacY in the C_{Out} conformation. Hence, the system has to spend a considerable amount of time in this state before the protein captures a lactose from the periplasm ($2 \rightarrow 3$). (ii) There is a large free-energy barrier between the States 2 and 1 [see Fig. 1(b)]. These two characteristics explain the importance of protonation for lactose binding. When LacY captures an H^+ ($1 \rightarrow 2$), the system goes from the highest energy state to the lowest one, which creates the free-energy barrier that keeps the protein in the State 2 while a lactose is not captured. Thus, according to our model, the importance of protonation is in the fact that the H^+ binding causes a large energy drop, reducing the rate of reversing protonation, and allowing relatively slow lactose binding to occur. If there was an appreciable reverse rate from 2 to 1, then LacY would return to the State 1 very easily, and then the next transition would probably be $1 \rightarrow 6$, not $1 \rightarrow 2$, since the transition $1 \rightarrow 6$ is much more likely to occur, as the transition $1 \rightarrow 2$ depends on the availability of H^+ near the entrance of LacY. Therefore, without the free-energy barrier between 1 and 2, LacY would not spend most of its time in the State 2, as required for lactose binding, and hence the influx of lactose would be almost impossible.

In Fig. 4, we use our model to reproduce two experiments from Ujwal *et al.* [25]. The first is the active transport of lactose in *E. coli* T184 expressing wild-type (WT) permease [see Fig. 4(a)], while the second is the equilibrium lactose exchange in *E. coli* T184 right-side-out membrane vesicles containing WT permease [see Fig. 4(b)]. As one can see, the similarity between our model predictions and the experimental data is good, strongly supporting the physical time scale that we calculated earlier. This is an important confirmation that our model has captured the fundamental and essential physics of the lactose-H^+ symport, and that it can provide reliable results within the accuracy of the free-energy curve obtained here. In future studies, it will be interesting to take into account more than one LacY embedded in the cell membrane. It will also be interesting to incorporate into Kaback's model occluded states between the transitions $C_{Out} \leftrightarrow C_{In}$, as well as to consider the activation energy between consecutive states.

In summary, we presented Monte Carlo simulations for the transport cycle of *Escherichia coli* lactose permease, using as a starting point the model proposed by Kaback *et al.* These simulations reproduced well the dynamics of the lactose transport by LacY that we found in the literature. For instance, we calculated the occupancy time of the protein in each one of its six states and obtained that it spends more time in its deprotonated outward-facing state. In addition, we paved the way in this work to solve an open question related to this transport mechanism, which is the importance of protonation for lactose binding. We found that protonation causes a large energy drop, allowing relatively slow lactose binding to occur. Our model can be used to investigate mutations and genetic manipulations, such as the mutant E269D (for details, see Ref. [27]).

ACKNOWLEDGMENTS

This work was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo Grant No. 2017/16910-7. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brasil (CAPES)–Finance Code 001.

[1] J. Abramson, I. Smirnova, V. Kasho, G. Verner, S. Iwata, and H. R. Kaback, The lactose permease of *Escherichia coli*: Overall structure, the sugar-binding site and the alternating access model for transport, FEBS Lett. **555**, 96 (2003).

[2] R. Krämer and C. Ziegler, *Membrane Transport Mechanism: 3D Structure and Beyond*, Springer Series in Biophysics (Springer, Berlin, 2016).

[3] Y. Yin, M. Ø. Jensen, E. Tajkhorshid, and K. Schulten, Sugar binding and protein conformational changes in lactose permease, Biophys. J. **91**, 3972 (2006).

[4] H. R. Kaback, M. Sahin-Tóth, and A. B. Weinglass, The kamikaze approach to membrane transport, Nat. Rev. Mol. Cell Biol. **2**, 610 (2001).

[5] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H. R. Kaback, and S. Iwata, Structure and mechanism of the lactose permease of *Escherichia coli*, Science **301**, 610 (2003).

[6] J. R. Lancaster, Site-exposure model for proton-lactose symport in *Escherichia coli*, J. Theor. Biol. **75**, 35 (1978).

[7] H. R. Kaback, A molecular mechanism for energy coupling in a membrane transport protein, the lactose permease

of *Escherichia coli*, Proc. Natl. Acad. Sci. USA **94**, 5539 (1997).

[8] M. Sahin-Tóth and H. R. Kaback, Arg-302 facilitates deprotonation of glu-325 in the transport mechanism of the lactose permease from *Escherichia coli*, Proc. Natl. Acad. Sci. USA **98**, 6068 (2001).

[9] L. Guan and H. R. Kaback, Binding affinity of lactose permease is not altered by the H^+ electrochemical gradient, Proc. Natl. Acad. Sci. USA **101**, 12148 (2004).

[10] R. J. Naftalin, N. Green, and P. Cunningham, Lactose permease H^+-lactose symporter: Mechanical switch or brownian ratchet, Biophys. J. **92**, 3474 (2007).

[11] M. Jensen, Y. Yin, E. Tajkhorshid, and K. Schulten, Sugar transport across lactose permease probed by steered molecular dynamics, Biophys. J. **93**, 92 (2007).

[12] L. Guan and H. R. Kaback, Properties of a lacy efflux mutant, Biochemistry **48**, 9250 (2009).

[13] I. Smirnova, V. Kasho, and H. R. Kaback, Lactose permease and the alternating access mechanism, Biochemistry **50**, 9684 (2011).

[14] H. R. Kaback, A chemiosmotic mechanism of symport, Proc. Natl. Acad. Sci. USA **112**, 1259 (2015).

[15] L. Guan and H. R. Kaback, Lessons from lactose permease, Annu. Rev. Biophys. Biomol. Struct. **35**, 67 (2006).

[16] G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer, Berlin, 2012).

[17] X. C. Zhang, Y. Zhao, J. Heng, and D. Jiang, Energy coupling mechanisms of mfs transporters, Protein Sci. **24**, 1560 (2015).

[18] M. Menshikov, S. Popov, and A. Wade, *Non-homogeneous Random Walks: Lyapunov Function Methods for Near-Critical Stochastic Systems*, Cambridge Tracts in Mathematics (Cambridge University Press, Cambridge, 2016).

[19] K. Binder and D. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction* (Springer, Berlin, 2010).

[20] A. M. Alencar, J. P. Butler, and S. M. Mijailovich, Thermodynamic origin of cooperativity in actomyosin interactions: The coupling of short-range interactions with actin bending stiffness in an Ising-like model, Phys. Rev. E **79**, 041906 (2009).

[21] N. Rosenblatt, A. M. Alencar, A. Majumdar, B. Suki, and D. Stamenović, Dynamics of Prestressed Semiflexible Polymer Chains as a Model of Cell Rheology, Phys. Rev. Lett. **97**, 168101 (2006).

[22] A. Majumdar, B. Suki, N. Rosenblatt, A. M. Alencar, and D. Stamenović, Power-law creep behavior of a semiflexible chain, Phys. Rev. E **78**, 041922 (2008).

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.99.052411 for the PYTHON code used for our simulations.

[24] J. C. Wilks and J. L. Slonczewski, pH of the cytoplasm and periplasm of *Escherichia coli*: Rapid measurement by green fluorescent protein fluorimetry, J. Bacteriol. **189**, 5601 (2007).

[25] M. L. Ujwal, M. Sahin-Tóth, B. Persson, and H. R. Kaback, Role of glutamate-269 in the lactose permease of *Escherichia coli*, Mol. Membr. Biol. **11**, 9 (1994).

[26] I. Smirnova, V. Kasho, J. Sugihara, and H. R. Kaback, Opening the periplasmic cavity in lactose permease is the limiting step for sugar binding, Proc. Natl. Acad. Sci. USA **108**, 15147 (2011).

[27] A. B. Weinglass, M. Sondej, and H. R. Kaback, Manipulating conformational equilibria in the lactose permease of *Escherichia coli*, J. Mol. Biol. **315**, 561 (2002).