The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Bubble lifetimes in DNA gene promoters and their mutations affecting transcription

Cite as: J. Chem. Phys. 155, 095101 (2021); doi: 10.1063/5.0060335
Submitted: 17 June 2021 • Accepted: 16 August 2021 •
Published Online: 2 September 2021

M. Hillebrand,[1] G. Kalosakas,[2,a)] A. R. Bishop,[3] and Ch. Skokos[1]

AFFILIATIONS

[1] Nonlinear Dynamics and Chaos Group, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa
[2] Department of Materials Science, University of Patras, GR-26504 Rio, Greece
[3] Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

[a)] **Author to whom correspondence should be addressed:** georgek@upatras.gr

ABSTRACT

Relative lifetimes of inherent double stranded DNA openings with lengths up to ten base pairs are presented for different gene promoters and corresponding mutants that either increase or decrease transcriptional activity in the framework of the Peyrard–Bishop–Dauxois model. Extensive microcanonical simulations are used with energies corresponding to physiological temperature. The bubble lifetime profiles along the DNA sequences demonstrate a significant reduction of the average lifetime at the mutation sites when the mutated promoter decreases transcription, while a corresponding enhancement of the bubble lifetime is observed in the case of mutations leading to increased transcription. The relative difference in bubble lifetimes between the mutated and wild type promoters at the position of mutation varies from 20% to more than 30% as the bubble length decreases.

I. INTRODUCTION

As more and more research is carried out on DNA functional activity, there is increased evidence that local base pair fluctuational openings may be one important factor in the process of transcription.[1–3] Hydrogen bond vibrations and local disruptions, leading to transient separations of complementary strands, have been experimentally studied at different time scales.[4–7] Furthermore, various theoretical models have addressed the dynamical and statistical properties of base pair stretchings in the double helix.[8–20]

An apparent correlation between thermally induced large local openings ("bubbles") and functional sites along the DNA sequence, which are relevant for transcription initiation, has been noted in a number of investigations[2,3,21–27] where it has been found that bubbles form with a greater probability at transcriptionally active sites than elsewhere along the promoter sequence. These works have considered the thermal equilibrium or out-of-equilibrium dynamical properties of various gene promoter segments using the Peyrard–Bishop–Dauxois (PBD) coarse-grained model[10] of DNA or extensions of this model.

Examining the dynamics of DNA promoters through Langevin molecular dynamics (MD) simulations, the lifetimes of these bubble openings have been probed as potential indicators for the initiation of transcription.[3,22,28] In particular, the adeno-associated viral (AAV) P5 promoter was shown to exhibit relatively long-lived bubbles with characteristic lengths near the transcription start site (TSS).[28] A broader study of an array of mammalian gene promoters showed the same dynamical signatures; bubbles form with greater frequency and longer lifetimes at active sites in the promoter and particularly at the TSS.[22] The importance of the lifetime as an indicator of transcriptional activity was underlined by the findings that long-lived bubbles formed in functional regions, even when the occurrence probability in those regions was not exceptional. The motif of dynamics-driven transcription was further established by the detailed analysis of the SCP1 so-called "superpromoter". This analysis revealed that not only are long-lived bubbles at the TSS correlated with transcriptional activity, but mutations which significantly reduce the transcription also decrease the occurrence and the lifetimes of bubbles at the TSS but also mutations that significantly reduce the transcription and also decrease the occurrence and

the lifetimes of bubbles at the TSS correlated with transcriptional activity.[3] Consequently, there is a strong interest in further investigating this relationship between bubble lifetimes and transcription using extensive numerical simulations under various conditions of the examined system in order to deepen our understanding of dynamical features with a possible biological role.

In this work, we present a detailed analysis of *inherent* bubble lifetimes (i.e., without any environmental factors or effects from the surroundings, depending only on constant-energy fluctuations in the framework of a microcanonical evolution) across the sequence of different promoters—a viral, a bacterial, and a very strong artificial promoter—and comparisons with a corresponding mutant in each case that is known to either reduce or enhance transcriptional activity. Following a recent study of bubble lifetime distributions,[29] we use the PBD model[10] of DNA with sequence-dependent stacking parameters,[30] to perform constant-energy simulations (complementing earlier works implementing Langevin dynamics[3,22,28]), in order to more closely probe the internal characteristic times of double strand openings.

The methodology of microcanonical MD simulations differs from the approach of Langevin dynamics where the effects of a noisy environment at a particular temperature are simulated through random Gaussian forces. This stochastic element, as well as the accompanying frictional force, requires an additional parameter. This is the dissipation coefficient that governs, along with the temperature, the distribution from which the random force is drawn in accordance with the fluctuation–dissipation theorem. A limitation of this method is that it introduces an essentially arbitrary parameter that may affect the characteristic time scales of the system. In the considered MD approach, we study the inherent lifetimes of the bubbles in the framework of the PBD model.

We examine three promoters, namely, the viral AAV P5 promoter, the bacterial Lac operon promoter, and the artificial SCP1 superpromoter (see Sec. II for their sequences), and one particular mutant of each case exhibiting altered transcriptional activity. The P5 promoter is critical to the genetic activity of AAV DNA by directing relevant expression patterns.[31] We also consider a double mutation of this promoter, resulting in loss of transcription activity.[2] The Lac operon promoter is a thoroughly studied regulatory region in the *E. coli* K-12 bacterium. We investigate the wild type (WT), as well as the mutant Lac UV5, which exhibits increased transcription with no need for activator action.[32] The final promoter is the artificially constructed SCP1 superpromoter, designed to have exceptional transcriptive behavior.[33] The mutant studied here is the m1SCP1 sequence, resulting in reduced transcription.[22]

Sec. II introduces the PBD model, presents the studied promoter sequences, and outlines the methods used in the analysis of the simulation data. Section III lays out the results of our investigation, including a discussion of the relative lifetime changes in the mutated promoters. Finally, Sec. IV summarizes and concludes our work.

II. MODELLING AND NUMERICAL METHODS

In the PBD framework considered here, a coarse-grained model is used for the base pair stretchings and the force fields of the system are approximated through appropriate analytical potentials. The PBD model provides a Hamiltonian for the dynamics, with the on-site intra-base-pair interaction accounted for by a Morse potential V,

$$V(y_n) = D_n\left(e^{-a_n y_n} - 1\right)^2, \tag{1}$$

where y_n represents the relative displacement from the equilibrium of the bases within the nth base pair of a DNA sequence. The site-dependent parameters D_n and a_n distinguish adenine–thymine (A–T) and guanine–cytosine (G–C) base pairs along the sequence.

An anharmonic coupling W models the stacking energy,

$$W(y_n, y_{n-1}) = \frac{K_{n,n-1}}{2}\left(1 + \rho e^{b(y_n + y_{n-1})}\right)(y_n - y_{n-1})^2. \tag{2}$$

The total Hamiltonian of a DNA sequence having N base pairs reads

$$H = \sum_{n=1}^{N}\left[\frac{p_n^2}{2m} + V(y_n) + W(y_n, y_{n-1})\right], \tag{3}$$

where p_n are the conjugate momenta to the canonical displacements y_n. Periodic boundary conditions have been considered here.

Apart from the sequence-dependent spring constants $K_{n,n-1}$ of the stacking energy W, which have been obtained from Ref. 30, all other parameter values we use are from Ref. 34: $m = 300$ amu for the base pair reduced mass; $D_{GC} = 0.075$ eV, $a_{GC} = 6.9$ Å^{-1} and $D_{AT} = 0.05$ eV, $a_{AT} = 4.2$ Å^{-1} for G–C and A–T base pairs, respectively, in the Morse potential; $\rho = 2$ and $b = 0.35$ Å^{-1}. These values have been fitted to successfully reproduce specific melting curves in short oligonucleotides. They have been extensively used in a number of previous works (for example, in Refs. 2, 3, 21–23, 26, 28, and 35–44). The parameters $K_{n,n-1}$ of Eq. (2) take on sequence specific values (see Refs. 29 and 30), and they have been shown to accurately reproduce peculiar denaturation transition temperatures exhibited by homogeneous and periodic DNA oligonucleotides.[30] The PBD model treats the DNA molecule as a purely one-dimensional lattice with the base pair stretching as the only degree of freedom and ignores explicit interactions with other degrees of freedom, such as the twisting angle that describes the unwinding of the double helix. Such effects are phenomenologically taken into account in the effective PBD potential energies for the stretching, but any local unwinding, kinking, or bending of the DNA accompanying a bubble formation cannot be explicitly described in the framework of this model. Attempts to overcome this limitation with more elaborate models that explicitly take into account these coupled degrees of freedom have been considered.[13,14,19,20]

The DNA promoters considered here are presented below. For clarity, only one strand of each sequence is shown, while the complementary strand is implied. The TSS is explicitly indicated as it is preceded by (+1).

- A 69 base pair segment of the viral AAV P5 promoter: 5′-GTGGCC ATTTAGGGTA TATATGGCCG AGTGAGCGAG CAGGATCTCC (+1)ATTTTGACCG CGAAATTTGA ACG-3′.
- A 129 base pair segment of the bacterial Lac operon promoter: 5′-GAAAGCGGG CAGTGAGCGC AACG-CAATTA ATGTGAGTTA GCTCACTCAT TAGGCAC-CCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTTGTGTGG (+1)AATTGTGAGC GGATAACAAT TTCACACAGG-3′.

- A 81 base pair segment of the artificial superpromoter SCP1: 5′-GTACTT ATATAAGGGG GTGGGGGCGC GTTCGTCCTC (+1)AGTCGCGATC GAACACTCGA GCCGAGCAGA CGTGCCTACG GACCG-3′.

In addition to these promoters, one mutant is also examined for each case: The mutated AAV P5 promoter is obtained by changing the base pairs at sites +1 and +2 from A–T and T–A to G–C and C–G, respectively.[2,21] In the Lac UV5 mutant, the sites −9 and −8 of the Lac operon are changed from G–C and T–A base pairs both to A–T.[32] Finally, in the mutant m1SCP1 of the superpromoter SCP1, the base pairs at sites −5 and −4 change from T–A and C–G to C–G and G–C, respectively, and the sites +8 and +15 change both from A–T to G–C.[22]

We have made extensive constant-energy molecular dynamics simulations using the Hamiltonian of Eq. (3) and periodic boundary conditions. We note that in all sequences studied, the regions of interest (i.e., mutations) are sufficiently far from the boundaries of the sequence. As the longest bubbles considered here are of length 10 base pairs, all mutations occur at least 20 base pairs away from the ends of the sequence, and the use of periodic boundary conditions does not affect bubbles in the pertinent regions. The presented bubble lifetime data near the ends of the sequence should not be considered relevant, as this will be affected by the boundary conditions, but our results focus away from this region.

Random initial conditions were implemented, having a fixed energy corresponding to a temperature of 310 K, and the equations of motion were evolved using a symplectic integrator, namely, the symplectic Runge–Kutta–Nyström fourth order integration scheme SRKNb6.[45] The threshold values of $y_{AT}^{thr} = 0.24$ Å and $y_{GC}^{thr} = 0.15$ Å are used in order to consider openings of A–T and G–C base pairs, respectively, which are derived through the characteristic lengths of the corresponding Morse potential and are also consistent with the standard convention that half of the base pairs of a DNA sequence are open at the melting temperature. These thresholds provide a systematic way of studying base pair openings without requiring arbitrary choices of separation values. Additionally, they have values beyond the inflection point of the Morse potential (0.165 and 0.10 Å from the equilibrium for AT and GC base pairs, respectively) by the same percentage of 50% in both cases. The system is evolved for 10 ns to provide thorough thermalization, and then, base pair displacement data are recorded every picosecond for the next nanosecond. We note that the time scales here are for our coarse-grained model.

From these displacement data, the bubble probabilities and lifetime distributions can be calculated (see Ref. 29 for the details of the procedure). Then, the average bubble lifetimes are computed from the corresponding lifetime distributions. In addition to examining bubbles with a fixed length of $l = q$ base pairs (as in Ref. 29), we also take here a more flexible approach of allowing the size of the bubble to fluctuate by considering the bubbles of length $l > q$ for some values of q, starting at a given site. In order to reduce potential issues arising from statistical inadequacy due to the rarity of bubble occurrence, for the calculation of bubble lifetimes, we have used 10 000 simulations with different random initial conditions for each of the considered promoter sequences.

III. AVERAGE BUBBLE LIFETIMES IN THE WILD TYPE AND MUTANT PROMOTERS

We are interested here in the profiles of the average lifetimes $\langle t \rangle$ for bubbles with either a fixed length $l = q$ or a length $l > q$, along the promoter sequences, and in the changes of these profiles with the respective mutations that affect the transcriptional activity. To quantify the effect of the mutations, some of which reduce transcription while others enhance transcription, we consider the sequence dependence of the relative difference in average bubble lifetimes between the mutated and the wild type promoter, which is calculated as

$$\Delta\langle t \rangle_{rel} = \frac{\langle t \rangle_{Mut} - \langle t \rangle_{WT}}{\langle t \rangle_{WT}}, \tag{4}$$

where $\langle t \rangle_{Mut}$ and $\langle t \rangle_{WT}$ correspond to the average lifetimes in the mutated promoter and in the wild type, respectively. Using the relative difference $\Delta\langle t \rangle_{rel}$, we can clearly identify regions along the sequence where the mutations increase or decrease the overall bubble lifetimes, as these areas will exhibit positive or negative relative differences, respectively.

Figure 1 shows the profiles of the average bubble lifetimes in the AAV P5 promoter [Fig. 1(a)] and the mutated P5 promoter [Fig. 1(b)], as well as the relative difference $\Delta\langle t \rangle_{rel}$ between the two variants [Fig. 1(c)], for different bubble lengths $l = q$, with q ranging from 1 up to 10 base pairs. We see that there are regions of more persistent bubbles both upstream and downstream of the TSS, as well as

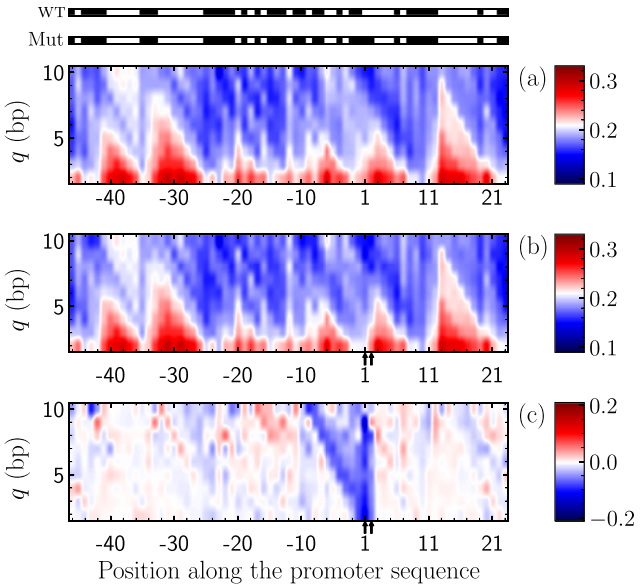

FIG. 1. Average bubble lifetimes for bubbles of length $l = q$ base pairs (bp) starting at a given base pair, as a function of the position of this base pair along the sequence for (a) the AAV P5 promoter and (b) the mutated P5 promoter. Color bars in (a) and (b) indicate bubble lifetimes in ps. The horizontal bars at the top show the distribution of A–T or T–A (white) and G–C or C–G (black) base pairs along the wild type (WT) and the mutated (Mut) sequence, respectively. (c) Relative difference $\Delta\langle t \rangle_{rel}$, Eq. (4), between the lifetimes of bubbles with length $l = q$ in the wild type and the mutated P5 promoter, as shown by the color bar at the right. The arrows below the x-axes in (b) and (c) indicate the mutation sites (see text).

at the TSS [Fig. 1(a)], correlated with the location of A/T rich bands along the sequence (see the horizontal bars above the plots). The effect of the TSS mutation, changing two A–T and T–A base pairs to G–C and C–G, respectively, which leads to loss of transcriptional activity, shows a clear reduction of the bubble lifetime in this region [Fig. 1(b)], which is emphasized in the plot of the relative difference $\Delta\langle t\rangle_{rel}$ in Fig. 1(c). Note that the color scale in Fig. 1(c), and all the relative difference plots shown in the other figures, is set symmetrically so that white regions correspond to zero relative difference, red regions to a positive relative difference (i.e., longer lifetimes in the mutated promoter), and blue regions to a negative relative difference. Here, we see a roughly 10% decrease in bubble lifetimes for bubbles starting or ending near the mutation sites (which coincide with the TSS), significantly reducing the lifetime of bubbles with lengths up to ten base pairs in this region. We note that similar lifetime profiles and relative differences are obtained when using Langevin dynamics with these threshold values, with parameters as used in earlier works.[28]

We have also considered bubbles of length $l > q$ for different values of q up to ten base pairs, which allow for fluctuations of the length of the bubble without effectively destroying and recreating new bubbles constantly in the numerical simulations. These results for the P5 promoter and the considered mutation are shown in Fig. 2, in the same way as the results of fixed length bubbles presented in Fig. 1. It is immediately apparent that the overall lifetimes are longer when the strictness of the fixed bubble length criterion is relaxed, with the longest average bubble lifetimes around 0.5 ps [Figs. 2(a) and 2(b)] for fluctuation-allowed bubbles as compared to 0.3 ps in the fixed-length case [Figs. 1(a) and 1(b)]. The statistics of the bubble lifetime profile is better in this case since all bubbles longer than length q now contribute to the data at each length. The same trends as in Fig. 1 are also apparent in Fig. 2. The effect of the mutation is to significantly shorten the lifetimes of bubbles near and immediately upstream of the TSS, without affecting the lifetimes along the rest of the sequence. However, Fig. 2(c) shows an even stronger reduction in the average lifetimes of bubbles at the TSS due to the mutation, at levels larger than 20% reaching up to 30% as q decreases from 10 to 2 base pairs, suggesting a correlation between significant changes in bubble lifetimes and altered transcriptional activity.

As the base pair opening thresholds $y^{thr}_{AT/GC}$ considered here are relatively small as compared to previous works, the multi-peaked bubble's inherent lifetime profiles shown in Figs. 1 and 2 for the AAV P5 promoter and its mutant are more reminiscent of the equilibrium average bubble probabilities[39] than the out-of-equilibrium probability profiles of much larger amplitude bubbles,[21,28] or even the bubble lifetime distributions obtained through Langevin dynamics.[28] However, all these works demonstrate the significant reduction of the bubble probability or relative lifetime at the position of the mutation at the TSS.

Turning to the bacterial Lac operon promoter, we present the lifetimes for bubbles of length $l > q$, as well as the relative difference between the wild type and mutant Lac UV5 in Fig. 3. The bands of

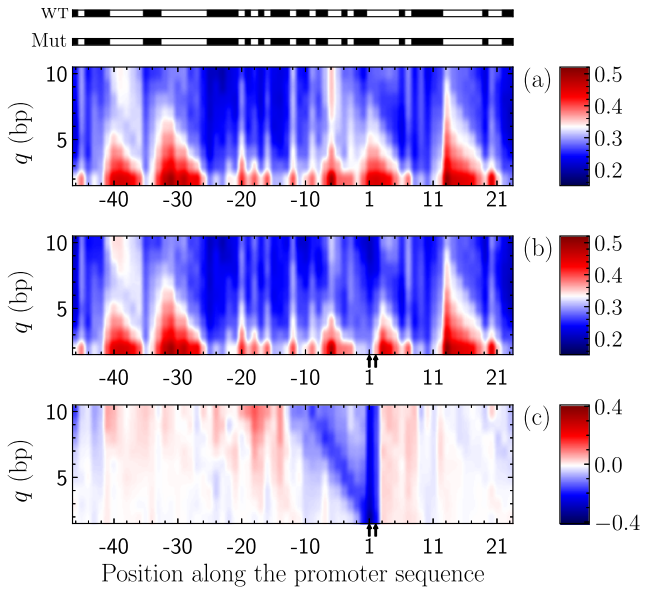

FIG. 2. Average bubble lifetimes for bubbles of length $l > q$ starting at a given base pair, as a function of the position of this base pair along the sequence for (a) the AAV P5 promoter and (b) the mutated P5 promoter. Color bars in (a) and (b) indicate bubble lifetimes in ps. (c) Relative difference $\Delta\langle t\rangle_{rel}$, Eq. (4), between the lifetimes of bubbles with length $l > q$ in the wild type and the mutated P5 promoter, as shown by the color bar at the right. The horizontal bars at the top, as well as the arrows below the axes in (b) and (c), are as in Fig. 1.

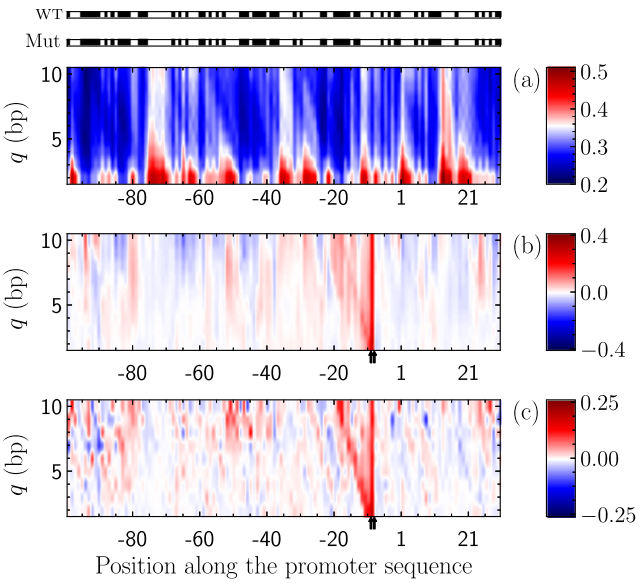

FIG. 3. (a) Average bubble lifetime profile in the wild type Lac operon promoter for bubbles of length $l > q$ starting at a given base pair, as a function of the position of this base pair along the sequence. The color bar indicates bubble lifetimes in ps. The relative difference $\Delta\langle t\rangle_{rel}$, Eq. (4), profile (b) for bubbles of length $l > q$ and (c) for bubbles of length $l = q$ is shown. The horizontal bars at the top of this figure depict the distribution of A–T/T–A (white) and G–C/C–G (black) base pairs in the sequence for the wild type Lac operon (WT) and the Lac UV5 mutant (Mut). The arrows below the axes in (b) and (c) indicate the mutation sites. The color scale in (b) and (c) is set symmetrically so that white regions signify no relative difference.

longer-lived bubbles are once again correlated with the A/T-dense regions of the sequence [Fig. 3(a)], including the region around the TSS. While there is no previous study on bubble lifetimes in the Lac operon, as in the AAV P5 case, the obtained lifetime profile is in accordance with equilibrium base pair opening probabilities in this promoter calculated through Monte Carlo simulations. In particular, the main peaks of the equilibrium bubble opening propensity have been observed[23] in (*i*) a large upstream region extending from around −80 up to −50, (*ii*) near the two binding sites of the polymerase subunit σ factor at around −30 (a larger peak) and in the region from −10 up to the TSS (a smaller one), and finally (*iii*) downstream in the region from +10 up to +20.

In contrast to the AAV P5 mutation examined above, which inhibits transcriptional activity, the Lac UV5 mutation is known to strengthen the promoter.[32,46] Figure 3(b) shows the relative difference $\Delta\langle t\rangle_{rel}$ for bubbles of length $l > q$, while Fig. 3(c) gives the corresponding profile for fixed bubble length $l = q$. We see that at the mutation site, which is located within the −10 element of the promoter (one of the two binding sites of the σ factor), Lac UV5 exhibits bubbles that tend to last longer, from 20% up to more than 30% in the case of $l > q$ depending on the bubble length. As observed previously, the relative difference is smaller, up to 25%, for the fixed length $l = q$ bubbles. Similar to the bubble lifetime enhancement, there is an increase in the equilibrium opening probability in this region in Lac UV5 as compared to the wild type profile.[23] Therefore, we see that the transcription-strengthening mutation results in longer-living bubbles at the transcriptionally functional binding site of the σ factor.

Finally, the profile of the average bubble lifetimes for the SCP1 superpromoter is shown in Fig. 4(a) for bubbles with length $l > q$, while the relative differences of average bubble lifetimes between this promoter and its mutant m1SCP1[3] are depicted in Fig. 4(b) for the case of bubbles with variable length $l > q$ and in Fig. 4(c) for bubbles with fixed length $l = q$. The equilibrium probabilities for large amplitude bubbles show a large peak in the region further upstream from the position −30 and another smaller peak around the TSS.[3] In agreement with these observations are the dominant feature around −30 in our average lifetime profiles and the peak around the TSS [Fig. 4(a)]. However, regarding small amplitude bubbles, as those considered in our case, there are additional peaks further downstream of the TSS, which are not present for bubbles with amplitudes larger than 3.5 Å in equilibrium.[3]

A primary finding obtained by Langevin dynamics simulations of the SCP1 superpromoter was a region immediately downstream from the TSS (located between +1 and +10) where large long-lived bubbles tended to form, while the introduction of the transcription-inhibiting mutations of m1SCP1 led to the destruction of this dominant peak.[3] Our results also show a substantial decrease in the average bubble lifetimes in this region downstream from the TSS, which becomes more clear for larger bubble lengths especially in the $l > q$ case [Fig. 4(b)]. Once more, the relative differences in the average bubble lifetimes around the mutated sites show a decrease, more than 20% for smaller lengths when bubbles of fluctuating ends are considered ($l > q$), for the m1SCP1 mutant that suppresses transcription as compared to the SCP1 superpromoter.

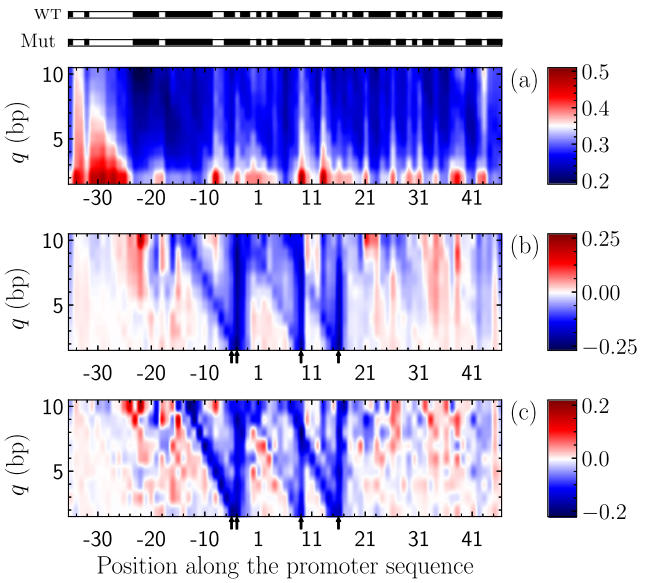

FIG. 4. (a) Average bubble lifetimes in the SCP1 superpromoter for bubbles of length $l > q$ starting at a given base pair, as a function of the position of this base pair along the sequence. The color bar indicates bubble lifetimes in ps. The relative difference $\Delta\langle t\rangle_{rel}$ (4) profile (b) for bubbles of length $l > q$ and (c) for bubbles of length $l = q$ is shown. The horizontal bars at the top of this figure depict the distribution of A–T or T–A (white) and G–C or C–G (black) base pairs in the sequence for the wild type (WT) and mutation (Mut). The arrows below the axes in (b) and (c) indicate the mutation sites. The color scale in (b) and (c) is set symmetrically so that white regions signify no relative difference.

IV. CONCLUSIONS

Using the Peyrard–Bishop–Dauxois coarse-grained model with sequence dependent stacking interactions for the description of base pair openings in DNA and efficient numerical techniques, we have performed extensive microcanonical simulations to investigate average bubble lifetime profiles along the sequence of three different promoters, namely the viral AAV P5 promoter, the bacterial Lac operon, and the artificial SCP1 superpromoter as well as one mutation for each promoter. The time scales of the inherent DNA double strand transient separations have been probed in the framework of this model with no artificial time scales imposed through arbitrary friction coefficient and using a physically motivated base-pair-dependent threshold value for considering base pairs to be open. The inherent bubble lifetimes for relatively small amplitude bubbles of the order of tenths of Å in our constant energy simulations are on the subpicosecond time scale, as opposed to Langevin fluctuational dynamics computations revealing bubble lifetimes of the order of picosecond for larger amplitude bubbles of a few Å.

We found that transcription-inhibiting mutations in the case of the AAV P5 and SCP1 promoters resulted in significant reductions of bubble lifetimes around the transcriptionally relevant mutated sites, while transcription-boosting mutations of the Lac operon promoter at a transcription factor binding site showed significant enhancement of the bubble lifetimes. The corresponding negative or positive relative differences in the average bubble lifetimes between

the mutated and the wild type promoter at the position of the mutation range from 20% for larger bubble lengths up to more than 30% for shorter ones.

ACKNOWLEDGMENTS

M.H. acknowledges support from the National Research Foundation (NRF) of South Africa (Grant No. 129630). G.K. and Ch.S. were supported by the Erasmus+/International Credit Mobility KA107 program. We acknowledge the Center for High Performance Computing of South Africa for providing computational resources for this project.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] H. M. Sobell, Proc. Natl. Acad. Sci. U. S. A. **82**, 5328 (1985).
[2] C. H. Choi, G. Kalosakas, K. Ø. Rasmussen, M. Hiromura, A. Bishop, and A. Usheva, Nucleic Acids Res. **32**, 1584 (2004).
[3] B. S. Alexandrov, V. Gelev, S. W. Yoo, L. B. Alexandrov, Y. Fukuyo, A. R. Bishop, K. Ø. Rasmussen, and A. Usheva, Nucleic Acids Res. **38**, 1790 (2010).
[4] M. Guéron, M. Kochoyan, and J.-L. Leroy, Nature **328**, 89 (1987).
[5] G. Altan-Bonnet, A. Libchaber, and O. Krichevsky, Phys. Rev. Lett. **90**, 138101 (2003).
[6] C. Phelps, W. Lee, D. Jose, P. H. von Hippel, and A. H. Marcus, Proc. Natl. Acad. Sci. U. S. A. **110**, 17320 (2013).
[7] M. González-Jiménez, G. Ramakrishnan, T. Harwood, A. J. Lapthorn, S. M. Kelly, E. M. Ellis, and K. Wynne, Nat. Commun. **7**, 11799 (2016).
[8] L. V. Yakushevich, *Nonlinear Physics of DNA*, 2nd ed. (Wiley-VCH, 2004).
[9] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989).
[10] T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, R44 (1993).
[11] T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, 684 (1993).
[12] T. Dauxois and M. Peyrard, Phys. Rev. E **51**, 4027 (1995).
[13] M. Barbi, S. Cocco, and M. Peyrard, Phys. Lett. A **253**, 358 (1999).
[14] S. Cocco and R. Monasson, J. Chem. Phys. **112**, 10017 (2000).
[15] A. Hanke and R. Metzler, J. Phys. A: Math. Gen. **36**, L473 (2003).
[16] D. Chatterjee, S. Chaudhury, and B. J. Cherayila, J. Chem. Phys. **127**, 155104 (2007).
[17] R. Tapia-Rojo, J. J. Mazo, and F. Falo, Phys. Rev. E **82**, 031916 (2010).
[18] M. Manghi and N. Destainville, Phys. Rep. **631**, 1 (2016).
[19] M. Zoli, J. Chem. Phys. **148**, 214902 (2018).
[20] M. Zoli, J. Chem. Phys. **154**, 194102 (2021).
[21] G. Kalosakas, K. Ø. Rasmussen, A. R. Bishop, C. H. Choi, and A. Usheva, Europhys. Lett. **68**, 127 (2004).
[22] B. S. Alexandrov, V. Gelev, S. W. Yoo, A. R. Bishop, K. Ø. Rasmussen, and A. Usheva, PLoS Comput. Biol. **5**, e1000313 (2009).
[23] A. Apostolaki and G. Kalosakas, Phys. Biol. **8**, 026006 (2011).
[24] R. Tapia-Rojo, D. Prada-Gracia, J. J. Mazo, and F. Falo, Phys. Rev. E **86**, 021908 (2012).
[25] H.-H. Huang and P. Lindblad, J. Biol. Eng. **7**, 10 (2013).
[26] K. Nowak-Lovato, L. B. Alexandrov, A. Banisadr, A. L. Bauer, A. R. Bishop, A. Usheva, F. Mu, E. Hong-Geller, K. Ø. Rasmussen, W. S. Hlavacek, and B. S. Alexandrov, PLoS Comput. Biol. **9**, e1002881 (2013).
[27] R. Tapia-Rojo, J. J. Mazo, J. Á. Hernández, M. L. Peleato, M. F. Fillat, and F. Falo, PLoS Comput. Biol. **10**, e1003835 (2014).
[28] B. S. Alexandrov, L. T. Wille, K. Ø. Rasmussen, A. R. Bishop, and K. B. Blagoev, Phys. Rev. E **74**, 050901 (2006).
[29] M. Hillebrand, G. Kalosakas, Ch. Skokos, and A. R. Bishop, Phys. Rev. E **102**, 062114 (2020).
[30] B. S. Alexandrov, V. Gelev, Y. Monisova, L. B. Alexandrov, A. R. Bishop, K. Ø. Rasmussen, and A. Usheva, Nucleic Acids Res. **37**, 2405 (2009).
[31] J. D. Tratschin, I. L. Miller, and B. J. Carter, J. Virol. **51**, 611 (1984).
[32] W. Gilbert, *RNA Polymerase*, edited by R. Losick and M. Chamberlin (Cold Spring Harbor Laboratory, New York, 1976), pp. 193–205.
[33] T. Juven-Gershon, S. Cheng, and J. T. Kadonaga, Nat. Methods **3**, 917–922 (2006).
[34] A. Campa and A. Giansanti, Phys. Rev. E **58**, 3585 (1998).
[35] N. K. Voulgarakis, G. Kalosakas, K. Ø. Rasmussen, and A. R. Bishop, Nano Lett. **4**, 629 (2004).
[36] S. Ares, N. K. Voulgarakis, K. Ø. Rasmussen, and A. R. Bishop, Phys. Rev. Lett. **94**, 035504 (2005).
[37] G. Kalosakas, K. Ø. Rasmussen, and A. R. Bishop, Chem. Phys. Lett. **432**, 291 (2006).
[38] N. K. Voulgarakis, A. Redondo, A. R. Bishop, and K. Ø. Rasmussen, Phys. Rev. Lett. **96**, 248101 (2006).
[39] Z. Rapti, A. Smerzi, K. Ø. Rasmussen, A. R. Bishop, C. H. Choi, and A. Usheva, Europhys. Lett. **74**, 540 (2006).
[40] S. Ares and G. Kalosakas, Nano Lett. **7**, 307 (2007).
[41] C. H. Choi, Z. Rapti, V. Gelev, M. R. Hacker, B. Alexandrov, E. J. Park, J. S. Park, N. Horikoshi, A. Smerzi, K. Ø. Rasmussen, A. R. Bishop, and A. Usheva, Biophys. J. **95**, 597 (2008).
[42] G. Kalosakas and S. Ares, J. Chem. Phys. **130**, 235104 (2009).
[43] J. J. Traverso, V. S. Manoranjan, A. R. Bishop, K. Ø. Rasmussen, and N. K. Voulgarakis, Sci. Rep. **5**, 9037 (2015).
[44] M. Hillebrand, G. Kalosakas, A. Schwellnus, and Ch. Skokos, Phys. Rev. E **99**, 022213 (2019).
[45] S. Blanes and P. C. Moan, J. Comput. Appl. Math. **142**, 313 (2002).
[46] W. S. Reznikoff, "Formation of the RNA polymerase-Lac promoter open complex," in *RNA Polymerase* (Cold Spring Harbor Laboratory, New York, 1976), pp. 441–454.