ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理 Chinese Physics B

Improved data analysis method of single-molecule experiments based on probability optimization

Weili Zhai(翟伟利), Guohua Yuan(袁国华), Chao Liu(刘超), Hu Chen(陈虎)
Citation: Chin. Phys. B . 2018, 27(1): 018703. **doi:** 10.1088/1674-1056/27/1/018703

Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Direct observation of λ -DNA molecule reversal movement within microfluidic channels under electric field with single molecule imaging technique

Fengyun Yang(杨凤云), Kaige Wang(王凯歌), Dan Sun(孙聃), Wei Zhao(赵伟), Hai-qing Wang(王海青), Xin He(何鑫), Gui-ren Wang(王归仁), Jin-tao Bai(白晋涛)
Chin. Phys. B . 2016, 25(7): 078201. **doi:** 10.1088/1674-1056/25/7/078201

Improvements in continuum modeling for biomolecular systems

Yu Qiao(乔瑜) and Ben-Zhuo Lu(卢本卓)
Chin. Phys. B . 2016, 25(1): 018705. **doi:** 10.1088/1674-1056/25/1/018705

Computational investigations on polymerase actions in gene transcription and replication: Combining physical modeling and atomistic simulations

Jin Yu(喻进)
Chin. Phys. B . 2016, 25(1): 018706. **doi:** 10.1088/1674-1056/25/1/018706

Colloidal monolayer self-assembly and its simulation via cellular automaton model

Wu Yi-Zhi, Chen Chen, Xu Xiao-Liang, Liu Yun-Xi, Shao Wei-Jia, Yin Nai-Qiang, Zhang Wen-Ting, Ke Jia-Xin, Fang Xiao-Tian
Chin. Phys. B . 2014, 23(8): 088703. **doi:** 10.1088/1674-1056/23/8/088703

Mechano-chemical selections of two competitive unfolding pathways of a single DNA i-motif

Xu Yue, Chen Hu, Qu Yu-Jie, Artem K. Efremov, Li Ming, Ouyang Zhong-Can, Liu Dong-Sheng, Yan Jie
Chin. Phys. B . 2014, 23(6): 068702. **doi:** 10.1088/1674-1056/23/6/068702

SPECIAL TOPIC — Soft matter and biological physics

Improved data analysis method of single-molecule experiments based on probability optimization*

Weili Zhai(翟伟利), Guohua Yuan(袁国华), Chao Liu(刘超), and Hu Chen(陈虎)†

Research Institute for Biomimetics and Soft Matter, Fujian Provincial Key Laboratory for Soft Functional Materials Research, Department of Physics, Xiamen University, Xiamen 361005, China

(Received 9 October 2017; revised manuscript received 20 October 2017; published online 13 December 2017)

To extract the dynamic parameters from single molecule manipulation experiments, usually lots of data at different forces need to be recorded. But the measuring time of a single molecule is limited due to breakage of the tether or degradation of the molecule. Here we propose a data analysis method based on probability maximization of the recorded time trace to extract the dynamic parameters from a single measurement. The feasibility of this method was verified by dealing with the simulation data of a two-state system. We also applied this method to estimate the parameters of DNA hairpin folding and unfolding dynamics measured by a magnetic tweezers experiment.

Keywords: probability optimization, Bell's model, DNA hairpin, single molecule manipulation

PACS: 87.80.Nj, 82.37.Rs, 87.15.A–, 02.50.–r **DOI:** 10.1088/1674-1056/27/1/018703

1. Introduction

The single molecule manipulation method has been widely used to study the conformational dynamics of biomolecules, such as DNA unzipping and protein folding/unfolding.[1,2] Important kinetic parameters can be obtained from single molecule experiments. In a single molecule manipulation experiment, the molecule is stretched by force and the extension is measured with nanometer resolution. From the measurement of force and extension, conformational transition of the molecules can be detected.[3]

If a constant force is applied to stretch the molecule, the lifetime of a specific conformation can be obtained.[4] If the molecule is stretched by a constant loading rate (force increasing rate) or constant speed of moving away an optical trap, the transition force can be obtained.[5,6] Distributions of lifetime or transition force are usually used to extract dynamic parameters of conformational transition of the molecule.

For a theoretical model, an analytical expression of the distribution can be derived in some cases. Therefore, fitting of the analytical expression with the measured distribution can be done easily to obtain the dynamic parameters.[7] If there is no analytical expression, a model-dependent simulation is used to compare with the experimental results,[8,9] which is usually done roughly by adjusting the parameter to minimize the difference intuitively. Both methods require the distributions of unfolding force, lifetime, *et al*. Because the conformational transition dynamics of a single molecule is a stochastic process, lots of measurements need to be done to obtain an accurate distribution. With limited experimental data, we cannot obtain an accurate distribution. In this case, the kinetic parameters are difficult to obtain, even rough estimations.

Here we propose a systematic algorithm to obtain the optimized parameters based on the measured time trace directly. We assume that the measured time trace has maximal probability to appear among all possible time traces. In the algorithm, the kinetic parameters are optimized to maximize the probability of the measured time trace. The optimized parameter values are good estimates of their real values. We use the most widely used Bell's model and dynamic simulation of two-state transitions to demonstrate the algorithm. Then we apply this algorithm to the single molecular manipulation experiment of a DNA hairpin.

2. Methods

2.1. Simulation of two-state model

The free energy landscape of the DNA hairpin (Fig. 1(b)) shows two states with local minimal free energy: the folded hairpin state and the unfolded ssDNA state, which can transfer to each other by crossing the transition state with the highest free energy barrier. To test the probability optimization algorithm, we used Monte Carlo simulation to obtain the time trace of the DNA state. Under a constant force, the folding rate and unfolding rate of the DNA hairpin are constant too. In the simulation, if the current conformation is a folded DNA hairpin, it transits to an unfolded ssDNA with probability $k_u\Delta t$ in the time step of Δt. Similarly, if the current conformation is an unfolded ssDNA, it transits to the folded haipin state with probability $k_f\Delta t$ in the time step of Δt. Time step Δt is chosen

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11474237 and 11574310) and the 111 Project, China (Grant No. B16029).

†Corresponding author. E-mail: chenhu@xmu.edu.cn

© 2018 Chinese Physical Society and IOP Publishing Ltd

to make sure $k_u\Delta t \ll 1$ and $k_f\Delta t \ll 1$.

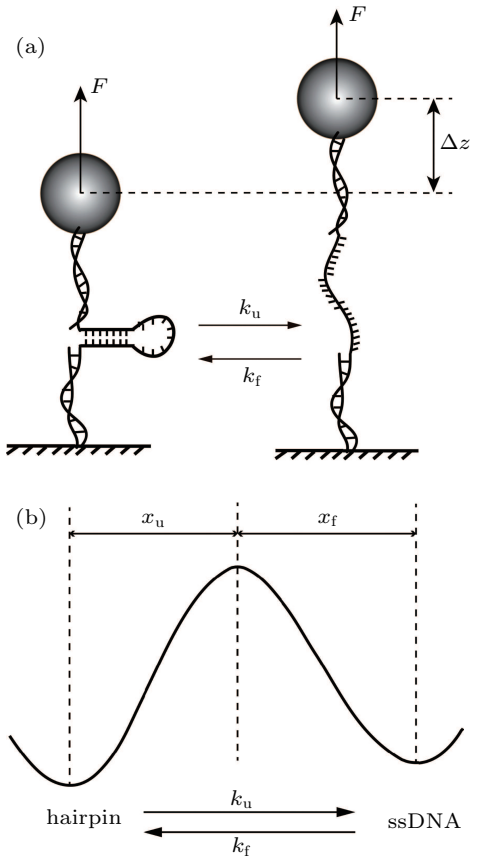


Fig. 1. Sketch of a DNA hairpin structure and its free energy landscape. (a) A DNA hairpin construct is linked between a superparamagnetic bead and a coverslip surface. The DNA hairpin construct will be unzipped to the ssDNA state at a big stretching force, indicated by an extension jump Δz. (b) The free energy landscape of the DNA hairpin at a force close to the critical force shows two energy minima representing the hairpin state and ssDNA state, respectively. The distance between the hairpin state and the transition state defines unfolding distance x_u, while the distance between the ssDNA and the transition state defines folding distance x_f.

If the force changes, the force-dependent transition rate $k(f)$ is usually described by Bell's model[10]

$$k(f) = k_0 \exp(fx/k_B T), \tag{1}$$

where k_0 denotes the transition rate at zero force, f is the stretching force, x is the transition distance, k_B is the Boltzmann constant, and T is the absolute temperature. Parameters k_0 and x define the force-dependent transition rate $k(f)$, and should be determined by a data analysis algorithm.

Based on Bell's model, force-dependent unfolding rate $k_u(f)$ and folding rate $k_f(f)$ of the DNA hairpin can be transformed to

$$k_u(f) = k_{f_c} \exp[(f - f_c)x_u/k_B T], \tag{2}$$

$$k_f(f) = k_{f_c} \exp[(f_c - f)x_f/k_B T], \tag{3}$$

where f_c is the critical force at which $k_u = k_f = k_{f_c}$, and x_u (x_f) denotes the distance from the native state (unfolded state) to the transition state (Fig. 1(b)).

In each cycle of the constant loading rate simulation, the force increased from 6 pN to 13 pN with loading rate of 0.005 pN/s, 0.007 pN/s, 0.01 pN/s, 0.02 pN/s, and 0.05 pN/s, then the force decreased from 13 pN to 6 pN with loading rate of −0.005 pN/s, −0.007 pN/s, −0.01 pN/s, −0.02 pN/s, and −0.05 pN/s, respectively.

2.2. Optimization algorithm

The time trace of states comes from a stochastic process. Therefore, each specific time trace has a certain probability to appear. The state transition process is a Markov process. Therefore, the probability of a specific time trace of states S_1, S_2, ..., S_N is given by

$$P_{S_1,\ldots,S_N} = \prod_{i=1}^{N-1} P_{S_i,S_{i+1}}, \tag{4}$$

where S_i is the state of the i-th data point ($S_i = 0$ denotes the folded hairpin state, and $S_i = 1$ denotes the unfolded ssDNA state), and N is the number of data points. The probability of time trace $S_1, S_2, \ldots, S_N$ is a function of the kinetic parameters in the model, such as k_u and k_f in the constant force process or k_{f_c}, f_c, x_u, and x_f in Bell's model.

The nearest neighboring state S_i, S_{i+1} has four possible combinations with probabilities

$$\begin{aligned} P_{0,1} &= \frac{k_u}{k_u + k_f}(1 - e^{-(k_u+k_f)\Delta t}), \\ P_{0,0} &= 1 - P_{0,1}, \\ P_{1,0} &= \frac{k_f}{k_u + k_f}(1 - e^{-(k_u+k_f)\Delta t}), \\ P_{1,1} &= 1 - P_{1,0}. \end{aligned}$$

When $k_u\Delta t \ll 1$ and $k_f\Delta t \ll 1$, the probabilities converge to

$$\begin{aligned} P_{0,1} &= k_u\Delta t, \\ P_{0,0} &= 1 - k_u\Delta t, \\ P_{1,0} &= k_f\Delta t, \\ P_{1,1} &= 1 - k_f\Delta t. \end{aligned}$$

We assumed that the recorded time trace has the maximal probability among all possible time traces. The probability (Eq. 4) can be maximized by adjusting the parameters that we are interested in. The optimized parameters were supposed to be good estimates of the real values. Note that for a long time trace, probability (Eq. 4) has a very small value. Therefore, the logarithm of $P_{S_1,\ldots,S_N}$ was used as the target function to avoid float underflow failure. The Matlab program was used to maximize the target function with the tolerance 10^{-4} on both the target function and parameters.

2.3. Experimental method

A small DNA hairpin with double-stranded DNA stem shorter than 30 base pairs and single-stranded DNA loop can be modeled as a two-state system: one folded DNA hairpin state stable at low stretching force, and one unfolded ssDNA state stable at high stretching force.[1] At critical force f_c, the probabilities of both states are 50%.

We used the previous reported method to build the DNA hairpin construct.[7,11] A DNA hairpin with sequence of 5'-GAG TCC TGG ATC CTG TTTTTTTT CAG GAT CCA GGA CTC-3' was sandwiched between two GC-rich DNA double helix handles with lengths of 228 bp and 361 bp, respectively. The DNA hairpin construct was labeled with a biotin and thiol group at each end. The thiol group end attached to an SMCC-coated cover slip surface through a covalent bond, while the biotin end was linked to a superparamagnetic bead coated with streptavidin (dynalbead M280, Invitrogen).

Home-made magnetic tweezers were used to apply force to the DNA hairpin construct. The magnetic tweezers controlled the stretching force by changing the distance between the permanent magnets and the sample.[12–14] Overstretching transition of DNA handles at ~ 65 pN was used as an indicator of a single DNA tether.[15,16] Upon unzipping of the DNA hairpin, extension of the DNA construct increased by a well-defined step size, which was determined by the hairpin size and the stretching force. Extension decreased by the same step size upon formation of the DNA hairpin structure.

3. Results

3.1. DNA hairpin unfolds and refolds in constant loading rate experiment

In the magnetic tweezers experiment, after a single DNA hairpin tether was confirmed by the overstretching signal at ~ 65 pN, we applied a constant loading rate of 0.09 pN/s to increase the force from 8.26 pN to 13.13 pN (Fig. 2). At the low force of 8.26 pN, the DNA stayed at the folded hairpin state. With increasing force, extension of the DNA hairpin jumped by a step size of ~ 16 nm at 9.8 pN, which indicated the unfolding of DNA hairpin. Because the loading rate was relatively slow, the DNA hairpin spontaneously refolded to shorten the extension suddenly at 10.25 pN, and unfolded again at 10.6 pN. At the high force of 13.13 pN, the DNA stayed at the unfolded ssDNA state. After that, the force was decreased with a loading rate of -0.09 pN/s. During the force-decreasing process, the DNA hairpin refolded, then unfolded, and refolded again. At 8.26 pN, the DNA returned to the initial folded hairpin state. The same force-increasing and force-decreasing process can be recorded multiple times. Because the nature of conformational transition is a stochastic process, the recorded folding and unfolding processes in different cycles are not identical.

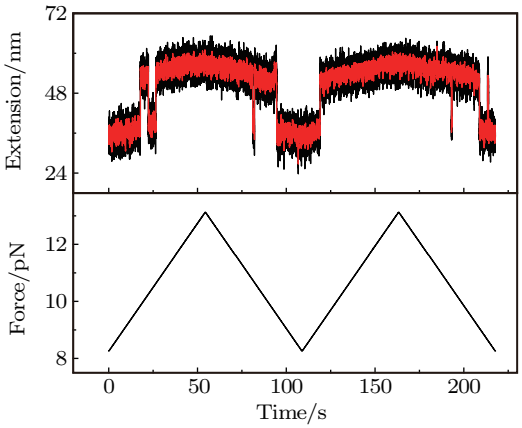


Fig. 2. (color online) Extension time trace of a DNA hairpin from the magnetic tweezers experiment. The force increases from 8.26 pN to 13.13 pN with loading rate of 0.09 pN/s, then decreases to 8.26 pN with loading rate of -0.09 pN/s (low panel). Two cycles of stretching are shown. Extension of the DNA hairpin is recorded with ~ 200 Hz sampling rate (black curve, upper panel). The red curve shows the smoothed curve averaged in 0.1 second time window.

From such a quick measurement, it is difficult to obtain accurate values of the dynamic parameters related to the force-dependent transition rates, such as f_c, x_u, and x_f in Bell's model. Usually, a series of measurements at different constant forces must be carried out to obtain the force-dependent folding rate $k_f(f)$ and unfolding rate $k_u(f)$. Then fitting with Bell's model gives the dynamic parameters that we are interested in. Such a traditional method is usually time-consuming and suffers from tether breakage. A model-based probability maximization algorithm was designed to obtain the dynamic parameters from a quick experiment.

3.2. Probability optimization algorithm applied to simple two-state transition

To test the rationality of the proposed probability optimization algorithm, we first applied it to the case of a simple two-state transition.

At constant force, the measured time trace of the DNA hairpin comes from a simple two-state model with constant transition rates of k_u and k_f. We set $k_u = 0.1\ s^{-1}$, $k_f = 0.2\ s^{-1}$, and obtained the state time trace by Monte Carlo simulation (Fig. 3(a)). In the traditional method, the mean of dwell time in the folded state (or unfolded state) gives the unfolding rate (or folding rate) as the reciprocal of the respective mean dwell time. While in the probability optimization algorithm, the probability equation (Eq. 4) is maximized to obtain the optimized parameters of k_u and k_f.

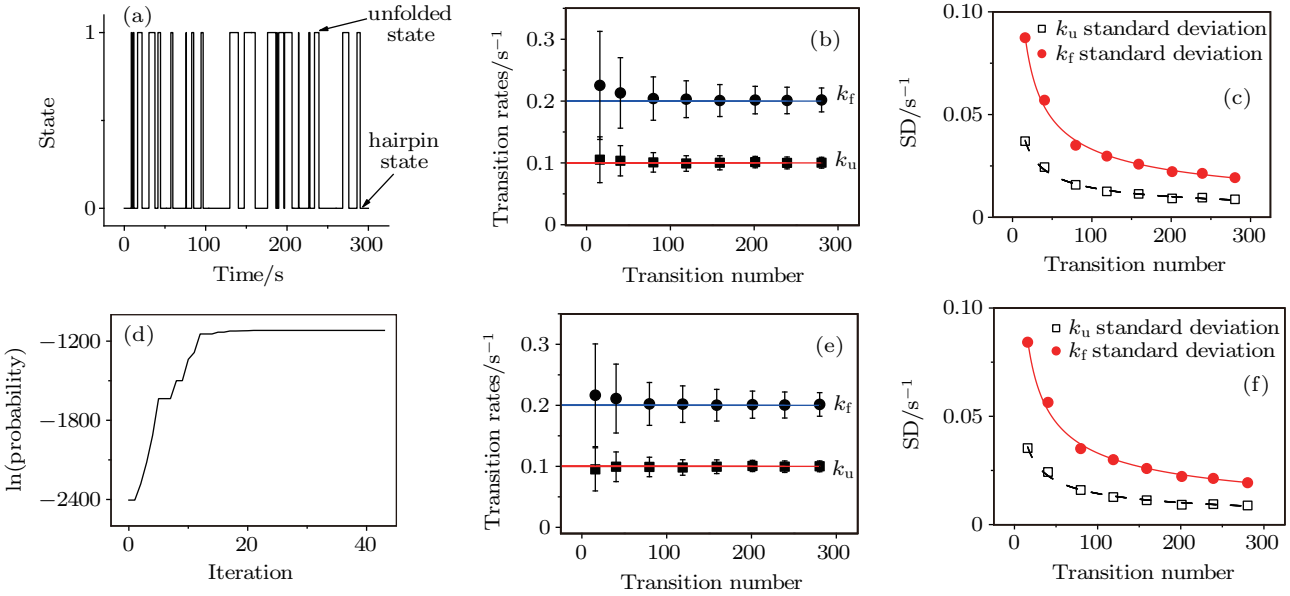


Fig. 3. (color online) The optimization algorithm applies to simulation of simple two-state transitions. (a) State time trace obtained by Monte Carlo simulation with parameters $k_u = 0.1$ s^{-1} and $k_f = 0.2$ s^{-1}. (b) Folding and unfolding rates obtained by the traditional mean dwell time method are shown as functions of the average transition number from simulations of different time lengths. The error bar shows the standard deviation of results from 100 simulations of the same time length. With increasing simulation time, the transition number increases, and the results converge to the correct values (blue and red horizontal lines). (c) Standard deviations of k_u and k_f as functions of the average transition number can be fitted by equation $y = ax^b$ with coefficients: k_u: $b = -0.52$, k_f: $b = -0.53$. (d) Logarithm of probability converges to a maximal value in the optimization process with initial values of $k_u = 1$ s^{-1} and $k_f = 4$ s^{-1}. (e) Folding and unfolding rates obtained from the probability optimization method are similar to the results in panel (b). (f) Standard deviations of results from the optimization method can also be fitted by equation $y = ax^b$ with coefficients: $b = -0.50$ for k_u and $b = -0.52$ for k_f.

Here we estimated the standard deviation of the results by multiple independent simulations. 100 independent time traces were obtained with the same simulation time. From each time trace, a set of parameters k_u and k_f were obtained by the traditional method and the probability optimization algorithm. Both mean and standard deviation were calculated from 100 sets of obtained parameters.

The number of recorded transitions is proportional to the simulation time. The average number of transitions was also calculated. Figure 3(b) shows the results of k_u and k_f by the traditional method as functions of the average number of transitions. With longer time trace, the average number of transitions increases, and the results converge to the input values of k_u and k_f. The standard deviation decreases with increasing the number of transitions.

The probability optimization method gave results similar to the traditional method (Figs. 3(d) and 3(e)), which validated the feasibility of the probability optimization method. The probability optimization method works well for the simple two-state transition.

Figures 3(c) and 3(f) show the standard deviations of k_u and k_f as functions of the transition number, which we tried to fit with function $y = ax^b$. Fitting gave the coefficients $b = -0.52$ (k_u) and $b = -0.53$ (k_f) for the traditional method, while $b = -0.50$ (k_u) and $b = -0.52$ (k_f) for the probability optimization method. The value of b is close to $-1/2$ as expected.

3.3. Probability optimization algorithm applied to constant loading rate simulation

During the measurement, if the force varies, then the transition rates change with time too. Therefore, the traditional mean dwell time method does not work. But the probability optimization algorithm still can be applied if a proper model is used to describe the force-dependent transition rates.

In a small force range, the transition rates usually follow Bell's model (Eqs. (2) and (3), Fig. 4(a)). We set four parameters similar to those of DNA hairpin zipping and unzipping transitions under force: $k_{f_c} = 0.29$ s^{-1}, $f_c = 10.3$ pN, $x_u = 7.1$ nm, and $x_f = 7.7$ nm. A series of time traces were obtained by simulation in the force range from 6 pN to 13 pN with five different loading rates of 0.005 pN/s, 0.007 pN/s, 0.01 pN/s, 0.02 pN/s, and 0.05 pN/s. In one simulation cycle, a force-increasing process was followed by a force-decreasing process. 100 cycles were simulated for each loading rate. The slower the loading rate, the more transitions were recorded in each simulation cycle. In each cycle, at 6 pN, the state always stayed at the folded hairpin state; while at the largest force of 13 pN, the DNA always stayed at the unfolded ssDNA state (Fig. 4(b)).

We optimized the parameters from the initial values of $k_{f_c} = 1$ s^{-1}, $f_c = 8$ pN, $x_u = 7$ nm, and $x_f = 9$ nm. Along with the optimization process (Fig. 4(c)), the probability increased and finally leveled off, then we obtained the optimal parameter values. For each cycle, a set of optimized parameters can

be obtained. Therefore, 100 values for each parameter were obtained from the 100 cycles of simulations, which were used to calculate the mean and standard deviation of each parameter. Figure 5 shows that the probability optimization method gives parameters converging to the simulation parameters with increasing number of transitions, which is reasonable.

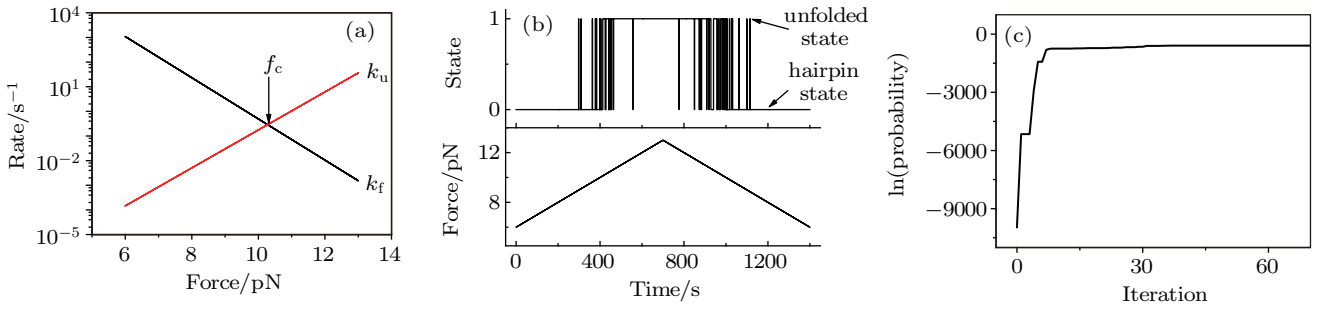


Fig. 4. (color online) Simulation of transition at constant loading rate and optimization method. (a) Force-dependent folding and unfolding rates (Eqs. 2 and 3) of DNA hairpin with parameters $k_{f_c} = 0.29$ s^{-1}, $f_c = 10.2$ pN, $x_u = 7.1$ nm, and $x_f = 7.7$ nm. (b) State time trace at loading rate of 0.01 pN/s. The force increases from 6 pN to 13 pN, then decreases to 6 pN. (c) Logarithm of probability converges to a maximal value in the optimization process with initial values of $k_{f_c} = 1$ s^{-1}, $f_c = 8$ pN, $x_u = 7$ nm, and $x_f = 9$ nm with optimized parameters shown in Fig. 5.

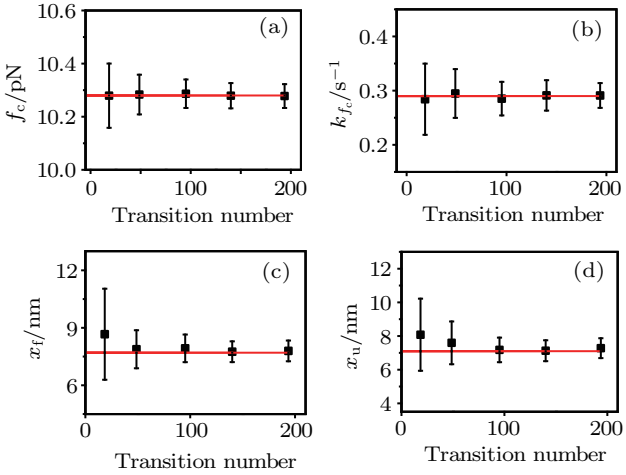


Fig. 5. (color online) Results of the probability optimization algorithm for constant loading rate simulations. Optimized parameters (a) f_c, (b) k_{f_c}, (c) x_f, and (d) x_u are shown as functions of the average transition number from simulations of different time lengths. The error bar shows the standard deviation of results from 100 simulations of the same time length. The red lines are set values of simulations: $f_c = 10.3$ pN, $k_{f_c} = 0.29$ s^{-1}, $x_f = 7.7$ nm, $x_u = 7.1$ nm.

3.4. Optimization algorithm analyses of experimental data under variable force

We applied our algorithm to experimental data of DNA hairpin folding and unfolding dynamics under variable forces which is difficult to analyze reliably. We conducted a set of experiments with eight cycles (Fig. 2). In each cycle of the experiment, the force increased from 8.26 pN to 13.13 pN with loading rates of 0.005 pN/s, 0.01 pN/s, and 0.02 pN/s, then the force decreased from 13.13 pN to 8.26 pN with loading rates of −0.005 pN/s, −0.01 pN/s, and −0.02 pN/s, respectively.

In the extension time trace, the state of the DNA hairpin at each time point was determined from the extension and sharp folding and unfolding transitions. By applying the optimization algorithm on the experimental data, the results of f_c, k_{f_c}, x_f, and x_u are obtained and shown in Fig. 6 as a function of the average transition number, together with the standard deviation as the error bar. We can utilize all experimental data available to optimize the combined probability, which is supposed to give us the most accurate results of $k_{f_c} = 0.12$ s^{-1}, $f_c = 9.5$ pN, $x_u = 7.3$ nm, and $x_f = 6.2$ nm.

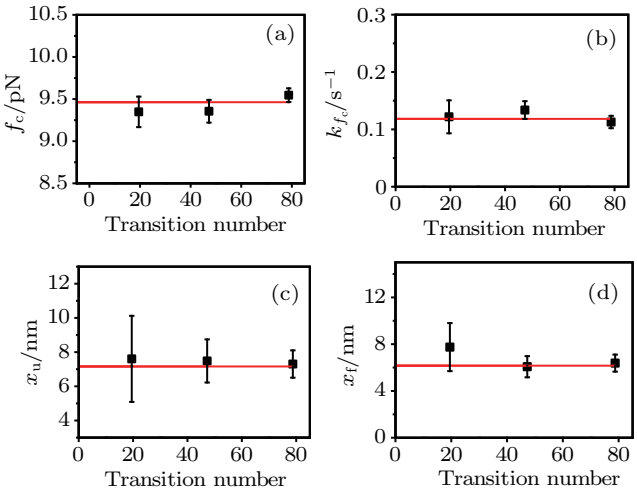


Fig. 6. (color online) Optimization algorithm analyses of experimental data under variable forces. Optimized parameters (a) f_c, (b) k_{f_c}, (c) x_u, and (d) x_f are shown as functions of the average transition number from experiments of different constant loading rates: 0.01 pN/s, 0.02 pN/s, and 0.005 pN/s. Each point is obtained from eight cycles. By optimizing the whole experiment data of the three different loading rates with about 1164 transition numbers, we obtain the optimized parameters values: $k_{f_c} = 0.12$ s^{-1}, $f_c = 9.5$ pN, $x_u = 7.3$ nm, $x_f = 6.2$ nm (red horizontal lines).

4. Discussion and conclusion

We have proposed a probability optimization algorithm, which highly improves the efficiency of experimental data analysis, especially for limited experimental data from which a smooth distribution cannot be obtained. The algorithm was verified by simulation of two-state transitions and applied to experimental measurement of DNA hairpin unzipping. For the simulation data under both constant force condition and force-loading cycle condition, reasonable results were obtain with

very limited data. The results converged to preset parameter values with increasing transition numbers recorded. For the real experiment of the DNA hairpin, the kinetic parameters were obtained with accuracy estimation.

The idea of maximum probability comes from thermodynamic principles. It is strict for a long time trace with an infinite number of transitions. While for a short time trace with a limited number of transitions, it already gives us good estimates of the kinetic parameters. For the same system, probabilities of multiple time traces can be combined together by multiplication. Therefore, all measurements can be utilized to get an estimation of the kinetic parameters. The results will gradually converge to the true value with increasing experimental data.

In magnetic tweezers, the force is controlled while the extension is measured.[13,17] In optical tweezers or atomic force microscopy experiment, the molecule is pulled by changing the position of the optical trap or distance between the sample and the cantilever, at the same time the force is measured.[6,8] For both cases, the maximum probability algorithm can be applied to determine the model parameters if a proper model is used to describe the force-dependent state transition rates.

The simple mean dwell time method can only be applied to experiments with fixed transition rates, such as the constant force experiment. While the probability optimization method can be used to analyze experimental data under any conditions, such as constant force, constant loading rate, and constant pulling speed conditions. Comparing with a Monte Carlo simulation and the unfolding force histogram comparison method, the probability optimization method does not require a large amount of data to obtain a smooth distribution. Both the Monte Carlo simulation method and the probability optimization method rely on a correct model to describe the force-dependent transition rates. In this paper, Bell's model is used to demonstrate the application. The probability optimization method is model-dependent, which might give a meaningless result when using an improper model. In this case, a series of constant force measurements is the most direct method to get the force-dependent transition rate.[4] Dudko's model-independent method also works well if enough experimental data can give a smooth force distribution over the interesting force range.[18]

Acknowledgment

We thank Mingxi Yao and Jie Yan for our insightful discussions.

References

[1] Woodside M T, Behnke-Parks W M, Larizadeh K, Travers K, Herschlag D and Block S M 2006 *Proc. Natl. Acad. Sci. USA* **103** 6190
[2] Chen H, Yuan G, Winardhi R S, Yao M, Popa I, Fernandez J M and Yan J 2015 *J. Am. Chem. Soc.* **137** 3540
[3] Neuman K C and Nagy A 2008 *Nat. Methods* **5** 491
[4] Yuan G, Le S, Yao M, Qian H, Zhou X, Yan J and Chen H 2017 *Angew. Chem. Int. Ed. Engl.* **56** 5490
[5] Chen H, Zhu X, Cong P, Sheetz M P, Nakamura F and Yan J 2011 *Biophys. J.* **101** 1231
[6] Cecconi C, Shank E A, Bustamante C and Marqusee S 2005 *Science* **309** 2057
[7] Liu C, Zhai W, Gong H, Liu Y and Chen H 2017 *Chem. Phys. Lett.* **678** 35
[8] Carrion-Vazquez M, Oberhauser A F, Fowler S B, Marszalek P E, Broedel S E, Clarke J and Fernandez J M 1999 *Proc. Natl. Acad. Sci. USA* **96** 3694
[9] Cao Y, Lam C, Wang M and Li H 2006 *Angew. Chem. Int. Ed. Engl.* **45** 642
[10] Bell G I 1978 *Science* **200** 618
[11] Xu Y, Chen H, Qu Y J, Efremov A K, Li M, Ou-Yang Z C, Liu D S and Yan J 2014 *Chin. Phys. B* **23** 68702
[12] Ran S Y, Sun B and Li M 2007 *Chin. Phys. Lett.* **36** 228
[13] Chen H, Fu H, Zhu X, Cong P, Nakamura F and Yan J 2011 *Biophys. J.* **100** 517
[14] Zheng H Z, Nong D G and Li M 2013 *Chin. Phys. Lett.* **30** 118702
[15] Smith S B, Cui Y and Bustamante C 1996 *Science* **271** 795
[16] Zhang X, Chen H, Fu H, Doyle P S and Yan J 2012 *Proc. Natl. Acad. Sci. USA* **109** 8103
[17] Strick T R, Allemand J F, Bensimon D, Bensimon A and Croquette V 1996 *Science* **271** 1835
[18] Dudko O K, Hummer G and Szabo A 2008 *Proc. Natl. Acad. Sci. USA* **105** 15755