Bloom's syndrome protein unfolding G-quadruplexes in two pathways*

Zhen-Ye Zhao(赵振业)[1,2], Chun-Hua Xu(徐春华)[1,2], Jing Shi(史婧)[3], Jing-Hua Li(李菁华)[4], Jian-Bing Ma(马建兵)[1,2], Qi Jia(贾棋)[1,2], Dong-Fei Ma(马东飞)[1,2], Ming Li(李明)[1,2], and Ying Lu(陆颖)[1,2,†]

[1] *Key Laboratory of Soft Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing* 100190, *China*
[2] *Beijing National Laboratory for Condensed Matter Physics, Institute of Physics and School of Physical Sciences, University of Chinese Academy of Sciences, Beijing* 100190, *China*
[3] *College of Life Science, Northwest A & F University, Xianyang* 712100, *China*
[4] *Material and Energy School, Guangdong University of Technology, Guangzhou* 510006, *China*

(Received 24 March 2017; revised manuscript received 12 April 2017; published online 20 June 2017)

The Bloom helicase (BLM) gene product encodes a DNA helicase that functions in homologous recombination repair to prevent genomic instability. BLM is highly active in binding and unfolding G-quadruplexes (G4), which are non-canonical DNA structures formed by Hoogsteen base-pairing in guanine-rich sequences. Here we use single-molecule fluorescence resonance energy transfer (smFRET) to study the molecular mechanism of BLM-catalysed G4 unfolding and show that BLM unfolds G4 in two pathways. Our data enable us to propose a model in which the HRDC domain functions as a regulator of BLM, depending on the position of the HRDC domain of BLM in action: when HRDC binds to the G4 sequence, BLM may hold G4 in the unfolded state; otherwise, it may remain on the unfolded G4 transiently so that G4 can refold immediately.

Keywords: G-quadruplexes, BLM, helicase, smFRET

PACS: 87.14.G–, 87.15.H–, 87.15.kj **DOI:** 10.1088/1674-1056/26/8/088701

1. Introduction

The Bloom helicase (BLM) belongs to the RecQ helicase family.[1,2] It contains two RecA-like domains which are involved in DNA binding and hydrolysis of adenosine triphosphate(ATP) and provides the ability to translocate the BLM on single-stranded DNA (ssDNA) in the 3′ to 5′ directions.[2,3] Deficiencies in the BLM could cause a serious genetic disease named Bloom syndrome (BS), the characteristics of which include genome instability, dwarfism, immunodeficiency, reduced fertility, and elevated levels of many types of cancers.[4,5] Cells from patients with BS show chromosomal instability characterized by higher rates of chromatid gaps and breaks, and sister chromatid exchanges.[6,7] BLM can resolve many different DNA substrates, such as Watson–Crick's duplex DNA, G-quadruplexes (G4)[8–10] and Holliday junctions.[11–13] It is generally accepted that BLM is ATP-dependent as it catalyzes G4 unfolding because it is an ATP hydrolysis-driven helicase.[10] However, a study showed that co-incubation of G4 with very a high concentration BLM (100 nM–2 M) may lead to 16%–45% unfolding of G4 in the absence of ATP or in the presence of the non-hydrolysable ATP analog ATP-γ-S.[9]

BLM contains an HRDC domain which is specific to the RecQ family. Previous studies showed the HRDC domain of the BLM works mainly as an assistant in DNA binding and translocation along ssDNA.[14,15] However, BLM mutants lacking HRDC show deficiency in strand annealing[16] and double Holliday junction dissolution[17] while they act similarly as the wild type.[16–18] The BLM binds G4 structures with high specificity.[19] BLM without the HRDC domain can also unfold G4.[18,20] It therefore seems that HRDC is not necessary for BLM to unfold G4. Yet there are reports that HRDC binds to ssDNA,[21,22] to assist the BLM to dissolve G4.[20] It was also proposed that HRDC could regulate the process of ATP hydrolysis of the BLM.[23] Here we use single-molecule fluorescence resonance energy transfer (smFRET) to study the interaction between BLM and G4. Surprisingly, we find that BLM with and without the HRDC domain unfolds G4 differently. Further studies indicate that the difference depends on the concentration of ATPs. We propose a model for explaining the results and make a suggestion that the HRDC domain is needed by BLM to hold the G4 in the unfolded state, possibly waiting for other proteins to act on the G4 sequence.

2. Experiments

2.1. DNA preparations

G4 is a non-canonical structure formed by Hoogsteen base-pairing in guanine-rich DNA sequences.[24,25] These regions of genome are significant for maintaining the human

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11674382, 11574381, and 11574382) and the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (Grant No. QYZDJ-SSW-SYS014).

†Corresponding author. E-mail: yinglu@iphy.ac.cn

© 2017 Chinese Physical Society and IOP Publishing Ltd

genome. In our experiments, G4 was composed of a DNA sequence of $(GGGTTA)_3GGG$. It was linked to a 29-bp duplex DNA at its 5′ end and a 15-nt ssDNA at its 3′ end (Fig. 1). All the oligonucleotides required to make the DNA substrates were purchased from Sangon Biotech (Shanghai, China). Their sequences are listed in Table 1.

Table 1. Sequence of DNA substrates used in the experiments.

	Substrates for single-molecule FRET(5′-3′)
G4	GCGTGGCACCGGTAATAGGAAATAGGAGAGGGTTAGGGTTAGGGTTA
	GGGT(icy3)TTTTTTTTTTTTT
stem	TCTCCT(icy5)ATTTCCTATTACCGGTGCCACGC-biotin
Y	CGACTGCTAGCATGTCGTACTAGCACTGTAGCTGTAGTACT(icy3)TTTTTTTTTTTTT
G4+Y	CGACTGCTAGCATGTCGTACTAGCACTGTAGCTGTAGT(icy3)ACGGGT
	TAGGGTTAGGGTTAGGGTTTTTTTTTTTTT
stem	TTTTTTTTTTTTTTTTTTTTTTTTT(icy5)GTACTACAGCTACAGTGCTA
	GTACGACATGCTAGCAGTCG-biotin

2.2. Protein expression and purification

BLM mutations (Figs. 1(a) and 4(a)) were used in our experiments. The $BLM^{642-1290}$ (called BLM from the plasmid encoding human gene in the present work) mutation contains an HRDC domain and the $BLM^{610-1298}$ (called $BLM^{\Delta HRDC}$ from the plasmid encoding gallus gene in the present work) does not contain an HRDC. Truncated $BLM^{642-1290}$ was expressed and purified as previously described.[12]

2.3. Data acquisition and analyses

All smFRET experiments were carried out with a home-built objective-type total-internal-reflection microscope at room temperature. Coverslips (Fisher Scientific) and slides were cleaned thoroughly by a mixture of sulfuric acid and hydrogen peroxide, then the surfaces of the coverslips were coated with a mixture of 99% monomethoxy-polyethylene glycol (m-PEG-5000, Laysan Bio, Inc.) and 1% of biotin-PEG (biotin-PEG-5000, Laysan Bio, Inc.). Streptavidin (1 mg/mL) in buffer containing 50-mM NaCl, 20-mM Tris-HCl, pH7.5 was added to the microfluidic chamber made of the PEG-coated coverslip and was incubated for 10 min. After washing, 20 pM–50 pM DNAs were added to the chamber and allowed to be immobilized for 10 min. Then free DNA was removed by washing with the reaction buffer. After that, the chamber was filled with the reaction buffer with an oxygen scavenging system (0.8% D-glucose, 1-mg/mL glucose oxidase, 0.4-mg/mL catalase, and 1-mM Trolox). The exposure time is 100 ms in all of our experiments. The raw fluorescence intensity trajectories were three-point averaged. Then, the FRET efficiency was calculated by using $I_A/(I_D+I_A)$, where I_D and I_A represent the intensity of the donor and the acceptor, respectively.

3. Results

We first measure the activity of 2-nM BLM at 200-μM ATP which is nearly the saturation concentration. About 15% of the substrates could be unfolded by BLM. Upon the addition of BLM, we observe repetitive unfolding signals (Figs. 1(c) and 1(d)), with the FRET values switching between 0.8 and 0.2. The repetitive signals are consistent with a recent result of BLM unfolding G4 (31).[26] We, however, observed a surprising phenomenon in which some G4's stay in the low FRET state (0.2) for a long period of time. A histogram of the dwell times of the low FRET state is shown in Fig. 1(g), and we find that the histogram can be fitted with a two-exponential function with a short characteristic time 2.7 ± 0.1 s and a very long characteristic time 93 ± 2 s. It is obvious that there are two different pathways of G4 unwound by BLM. One way is by short-time unwinding and the other is by long-time unwinding.

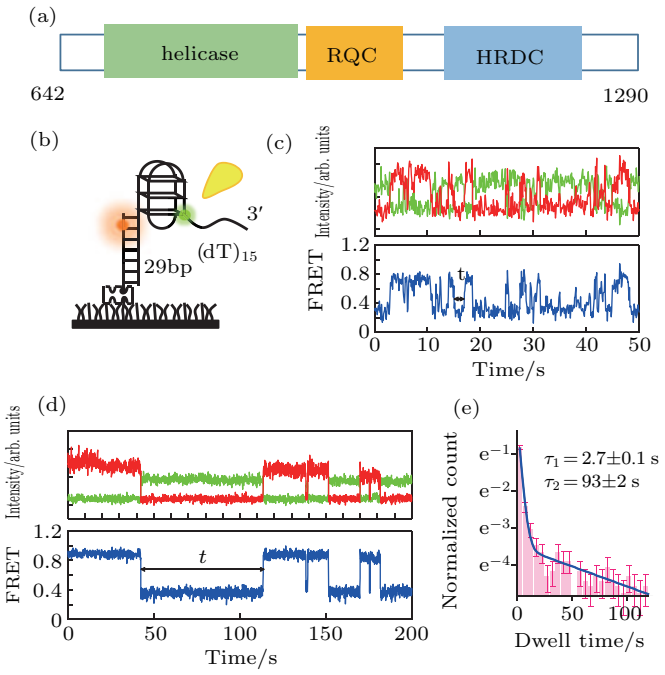


Fig. 1. (color online) BLM unfolds G4 at 200-μM ATP. (a) BLM protein functional domains used in this study. (b) Schematic diagram of Cy3 (green)- and Cy5 (red)-labeled DNA construct (G4) with G4 having three G-quartet planes. G4 strand and complementary stem strand are annealed to form a duplex stem. Biotin is used to immobilize the DNA to a streptavidin-coated coverslip surface. (c) and (d) Traces of fluorescence intensities of Cy3 and Cy5 (upper panel) and FRET trace (lower panel) with 2-nM BLM and 200-μM ATP. The dwelling time of the completely unfolded state is denoted as t. (e) Histograms of the dwelling times for the completely unfolded G4. Error bar = s.e.m.

The results above show that the BLM helicase either unfolds a G4 repetitively or holds it in the unfolded state for a long time. Now, we come to see whether such a feature of the BLM would affect its unwinding of a DNA duplex (Figs. 2(a) and 2(b)). We design two DNA forks, each with an FRET pair (Cy3 and Cy5) near the entrance of the duplex. One of them is a simple fork with two poly(T) tails (hereafter called Y DNA), and the other has a G4 on the 3′ tail right in front of the entrance (hereafter called G4+Y DNA). Both the DNA structures will yield similar FRET signals when the duplex is unwound by BLM. The experimental conditions are similar to those in Fig. 1, namely 2-nM BLM with 200-μM ATP. It turns out that BLM almost does not unwind the duplex with a G4 in its translocation way. We believe that this indicates that BLM has a strong affinity with the G4 sequence so that it hinders BLM.[20]

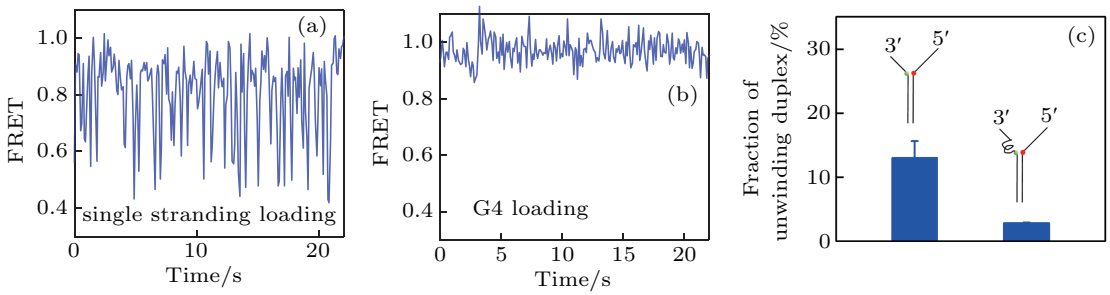


Fig. 2. (color online) G4 on the loading overhang hindering BLM unwinding DNA duplex. (a) and 2(b) A typical FRET trace of BLM unwinding a DNA duplex without (a) and with (b) a G4 on the 3′ overhang. The schematic diagrams of the two substrates are shown in the inset of panel (c). (c) Comparison between the fractions of duplexes with and without a G4 unwound by BLM. Error bar = s.e.m.

We lower the ATP concentrations to 20 μM and 1 μM respectively, and repeat the unfolding experiments. The phenomenon of long-time dwelling almost disappears, while BLM still repetitively unfold G4 (Fig. 3). Less than 1% of the unfolding evens are observed to exhibit long-time dwelling at 1-μM ATP. This is in strong contrast with the value of $> 30\%$ when the ATP concentration is 200 μM. It suggests that ATP hydrolysis is necessary for BLM to keep G4 in the unfolded state for a long time. To confirm this, we perform similar experiments with ATPγS which is often used as a nonhydrolyzable substrate. No long-time dwelling events are observed even with using 1-mM ATPγS.

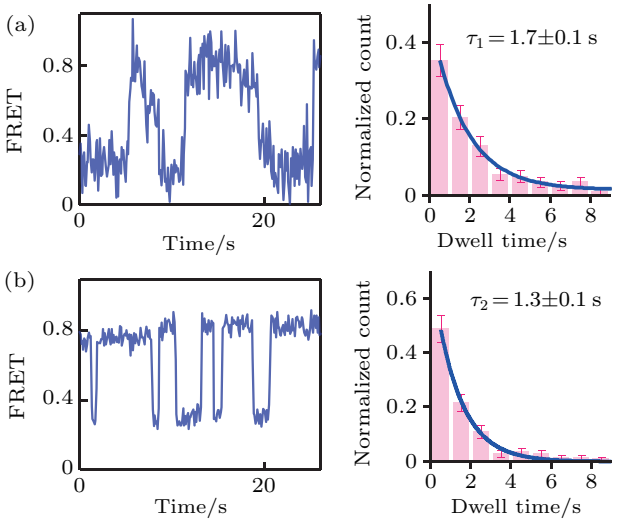


Fig. 3. (color online) Long-time dwelling that is less frequently observed at low ATP concentrations. (a) and (b) FRET traces and dwelling time distributions for G4 unfolding by 2-nM BLM at (a) 1-μM and (b) 20-μM ATP.

We use a BLM mutation with the HRDC domain being truncated ($BLM^{\Delta HRDC}$ (Fig. 4(a)) to see whether HRDC plays a role in the long-time dwelling. It turns out that $BLM^{\Delta HRDC}$ can unfold G4 just like a normal BLM with the HRDC domain. However, even at high ATP concentrations, only the unfolding events with short dwelling time are observed (Figs. 4(b) and 4(c)). The histogram of the short dwelling times for $BLM^{\Delta HRDC}$ is almost the same as that for the normal BLM. The results suggest that the HRDC domain of BLM is responsible for the observed long-time dwelling.

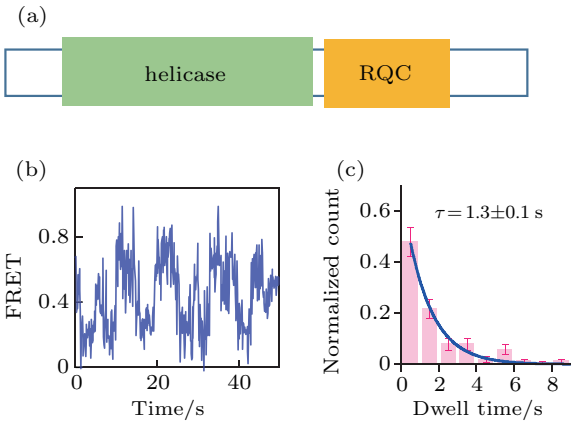


Fig. 4. (color online) Unfolding of G4 by BLM without the HRDC domain. (a) BLM mutation without the HRDC domain used in this study. (b) A FRET trace of unfolding G4 with 200-μM ATP. (c) Histograms of dwelling times for the completely unfolded G4.

4. Discussion

The HRDC domain has long been a mystery for the RecQ family helicases. BLM has been found to be able to unfold G4 even in the absence of HRDC. HRDC is also reported to be not necessary for BLM to unwind DNA duplexes.[18] However, our present study indicates that HRDC is not just a redundant domain of BLM. We observe that BLM with the HRDC domain could hold G4 in an unfolded state for a very long time. Without the HRDC domain, the binding of BLM to the G4 sequence is so transient that G4 can refold immediately after it

has been unfolded by BLM. A recent structural study shows that the HRDC domain is connected to the core of BLM by a long random coil so that HRDC can move freely around the core.[23] Based on our experimental results and the structure of BLM, we propose a model for BLM unfolding G4 (Fig. 5). Upon binding to G4, the HRDC domain is detached from the BLM core, either moving freely in solution or binding to the ssDNA. If HRDC is in the free state, BLM functions like a truncated BLM$^{\Delta HRDC}$, repetitively unfolding G4 (Fig. 5(a)). But if the HRDC domain binds to the G4 sequence, the BLM will cover the G4 consequence and prevents G4 from refolding (Fig. 5(b)).

Our results also show that ATP hydrolysis is needed for BLM to hold the unfolded G4 firmly. We propose that the ATP hydrolysis should cause the BLM core to be translocated on the DNA overhang. But the motion is blocked by the HRDC ahead of it so that the DNA overhang is pushed way, hence preventing G4 from refolding (Fig. 5(b)). The result is in agreement with the fact that HRDC is only loosely connected to the core of BLM.[23] At low ATP concentrations, the BLM core is free of ATP in time (before a new ATP is coupled to the core), the BLM affinity for DNA becomes low so that it unbinds the G4 sequence readily, allowing the unfolded G4 to refold. A large number of proteins have been found to coordinate with BLM to resolve G4.[27,28] It is possible that the long-time dwelling of G4 in the unfolded state enables BLM to prepare a good condition for the recruited protein to work effectively.

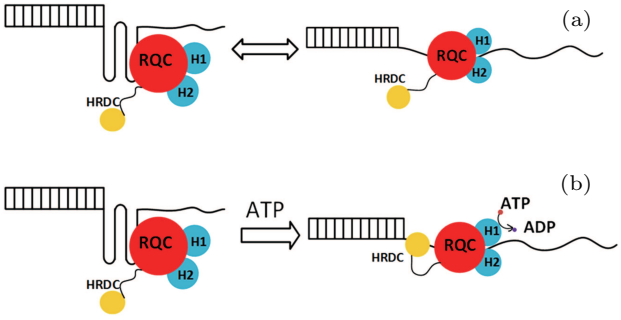


Fig. 5. (color online) Proposed model for the two unfolding pathways of G4 by BLM. (a) HRDC domain of BLM, which is free in G4 unfolding. (b) HRDC domain, which binds to G4 when BLM unfolds G4.

Acknowledgment

We thank Drs. J Diao and X Hou for their help in the FRET experiments.

References

[1] German J 1993 *Medicine* **72** 393
[2] Karow J K, Chakraverty R K and Hickson I D 1997 *J. Biol. Chem.* **272** 30611
[3] Yodh J G, Stevens B C, Kanagaraj R, Janscak P and Ha T 2009 *EMBO J.* **28** 405
[4] Cheok C F, Bachrati C Z, Chan K L, Ralf C, Wu L and Hickson I D 2005 *Biochem. Soc. Trans.* **33** 1456
[5] Xi X G 2010 *Methods in molecular biology* (Clifton, N.J.) **587** 173
[6] German J, Sanz M A, Ciocci S, Ye T Z and Ellis N A 2007 *Hum. Mutat.* **28** 743
[7] Ellis N A, Groden J, Ye T Z, Straughen J, Lennon D J, Ciocci S, Proytcheva M and German J 1995 *Cell* **83** 655
[8] Budhathoki J B, Ray S, Urban V, Janscak P, Yodh J G and Balci H 2014 *Nucleic Acids Res.* **42** 11528
[9] Budhathoki J B, Stafford E J, Yodh J G and Balci H 2015 *Nucleic Acids Res.* **43** 5961
[10] Wu W Q, Hou X M, Li M, Dou S X and Xi X G 2015 *Nucleic Acids Res.* **43** 4614
[11] Liu J Q, Chen C Y, Xue Y, Hao Y H and Tan Z 2010 *J. Am. Chem. Soc.* **132** 10521
[12] Sun H, Karow J K, Hickson I D and Maizels N 1998 *J. Biol. Chem.* **273** 27587
[13] Rezazadeh S 2013 *Mol. Biol. Rep.* **40** 3049
[14] Sato A, Mishima M, Nagai A, Kim S Y, Ito Y, Hakoshima T, Jee J G and Kitano K 2010 *J. Biochem.* **148** 517
[15] Kocsis Z S, Sarlos K, Harami G M, Martina M and Kovacs M 2014 *J. Biol. Chem.* **289** 5938
[16] Cheok C F, Wu L, Garcia P L, Janscak P and Hickson I D 2005 *Nucleic Acids Res.* **33** 3932
[17] Wu L, Chan K L, Ralf C, Bernstein D A, Garcia P L, Bohr V A, Vindigni A, Janscak P, Keck J L and Hickson I D 2005 *EMBO J.* **24** 2679
[18] Gyimesi M, Harami G M, Sarlos K, Hazai E, Bikadi Z and Kovacs M 2012 *Nucleic Acids Res.* **40** 3952
[19] Huber M D, Duquette M L, Shiels J C and Maizels N 2006 *J. Mol. Biol.* **358** 1071
[20] Chatterjee S, Zagelbaum J, Savitsky P, Sturzenegger A, Huttner D, Janscak P, Hickson I D, Gileadi O and Rothenberg E 2014 *Nat. Commun.* **5** 5556
[21] Bernstein D A and Keck J L 2005 *Structure* **13** 1173
[22] Killoran M P and Keck J L 2008 *Nucleic Acids Res.* **36** 3139
[23] Newman J A, Savitsky P, Allerston C K, Bizard A H, Ozer O, Sarlos K, Liu Y, Pardon E, Steyaert J, Hickson I D and Gileadi O 2015 *Nucleic Acids Res.* **43** 5221
[24] Gellert M, Lipsett M N and Davies D R 1962 *Proc. Natl. Acad. Sci. USA* **48** 2013
[25] Sen D and Gilbert W 1988 *Nature* **334** 364
[26] Tippana R, Hwang H, Opresko P L, Bohr V A and Myong S 2016 *Proc. Natl. Acad. Sci. USA* **113** 8448
[27] Lillard-Wetherell K, Machwe A, Langland G T, Combs K A, Behbehani G K, Schonberg S A, German J, Turchi J J, Orren D K and Groden J 2004 *Hum. Mol. Genet.* **13** 1919
[28] Selak N, Bachrati C Z, Shevelev I, Dietschy T, van Loon B, Jacob A, Hubscher U, Hoheisel J D, Hickson I D and Stagljar I 2008 *Nucleic Acids Res.* **36** 5166