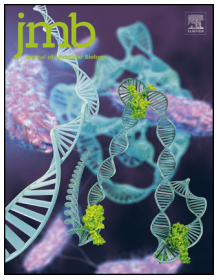

NMR Investigation of the Interaction between the RecQ C-Terminal Domain of Human Bloom Syndrome Protein and G-Quadruplex DNA from the Human c-Myc Promoter

Sungjin Lee[1,†], Ae-Ree Lee[2,†], Kyoung-Seok Ryu[3], Joon-Hwa Lee[2] and Chin-Ju Park[1,]

1 - Department of Chemistry, *Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea*
2 - Department of Chemistry and Research Institute of Natural Science, *Gyeongsang National University, Jinju, Gyeongnam 52828, Republic of Korea*
3 - Protein Structure Research Team, *Korea Basic Science Institute, Ochang, Chungbuk 28119, Republic of Korea*

Correspondence to Joon-Hwa Lee and Chin-Ju Park: *C.-J. Park is to be contacted at: Department of Chemistry, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea. joonhwa@gnu.ac.kr, cjpark@gist.ac.kr*
https://doi.org/10.1016/j.jmb.2019.01.010
Edited by E. Nogales

Abstract

Bloom syndrome protein (BLM) is one of five human RecQ helicases that participate in DNA metabolism. RecQ C-terminal (RQC) domain is the main DNA binding module of BLM and specifically recognizes G-quadruplex (G4) DNA structures. Because G4 processing by BLM is essential for regulating replication and transcription, both G4 and BLM are considered as potential targets for anticancer therapy. Although several studies have revealed the detailed mechanism of G4 unwinding by BLM, the initial recognition of the G4 structure by the RQC domain is unclear. Here, we investigated the interaction between BLM RQC and the G4 DNA from the c-Myc promoter by NMR spectroscopy. While the signals broadened upon reciprocal titrations, the β-wing of RQC had significant chemical shift perturbations and experienced millisecond timescale dynamics upon G4 binding. A point mutation in the β-wing (N1164A) reduced G4 binding affinity. Our hydrogen–deuterium exchange data indicate that imino protons of G4 were exchanged with deuterium much faster in the presence of RQC. We suggest that RQC binds to G4 by using the β-wing as a separating pin to destabilize the G4. By providing information about the RQC–G4 interaction, our study yields insight into potential strategies for preventing G4 processing by BLM.

Introduction

Bloom syndrome protein (BLM) is one of five human RecQ helicases (RECQ1, BLM, WRN, RECQ4, and RECQ5), and it plays essential roles in DNA metabolism [1]. Mutations in the *BLM* gene cause Bloom syndrome, a recessive genetic disorder with the symptoms of premature aging and a predisposition toward various types of cancer [2]. BLM uses the energy of ATP hydrolysis to unwind DNA from the 3′ to 5′ direction [3,4]. BLM protein is known as a binding partner of various DNA structures such as duplexes, Holliday junctions, D-loops, and G-quadruplexes (G4s). It has a binding preference for G4 [5,6].

BLM protein is composed of several domains including a helicase domain, RecQ C-terminal (RQC) domain, and helicase and RNase D C-terminal (HRDC) domain [6–8] (Fig. 1a). The three-dimensional structure of BLM was revealed by several structural studies [7–11] in which the role of the RQC domain was identified as a primary DNA binding module. It contains a winged helix motif and uses the β-wing as a strand-separation pin by recognizing the terminal bases of DNA duplexes [7,9]. Interestingly, BLM RQC domain uses a unique mechanism for binding DNA duplexes compared to other RQC domains of homologous proteins such as Werner syndrome protein (WRN). BLM RQC domain uses the polar contact between the

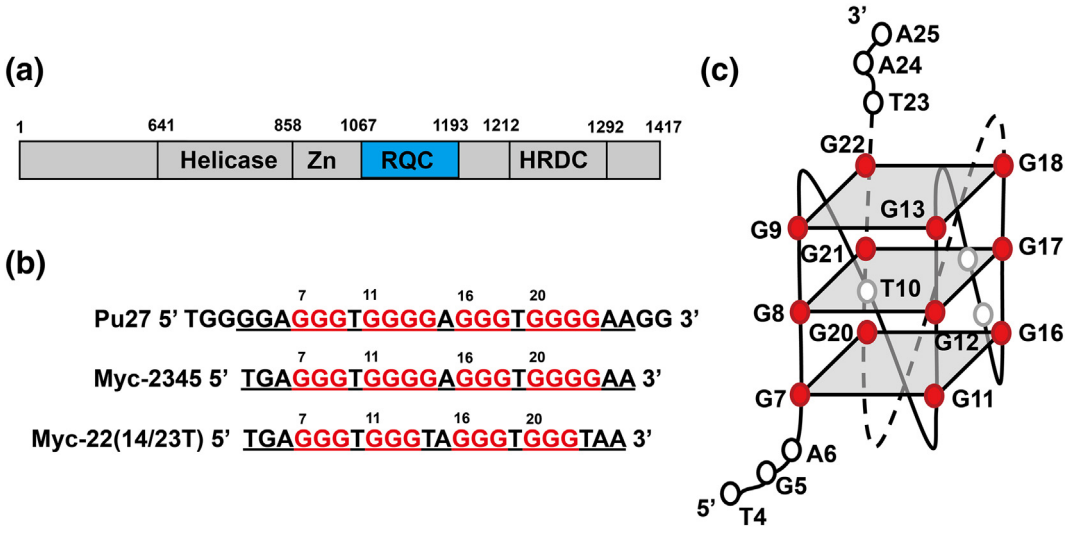

Fig. 1. (a) Domain structure of human BLM protein. (b) The c-Myc gene promoter sequence and its variants. Pu27 is the wild-type 27mer G-rich sequence of c-Myc NHE III. Myc-2345 is the 22mer G-rich sequence that adopts the major c-Myc promoter G4 structure. Myc-22(14/23T) is the major loop isomer of Myc2345. (c) Schematic structure of the conformation of Myc-22(14/23T). It is a propeller-type parallel-stranded G4 structure.

asparagine residue (N1164) at the top of the β-wing and the terminal base of the DNA [7]. However, other RQC domains, such as WRN RQC, use stacking interactions between the aromatic residue placed at the top of the β-wing and the terminal base of the substrate duplexes [12].

G4 is an alternative DNA structure formed by Hoogsteen base pairing between four guanine bases placed on the same plane [13–15]. The G4 structure is stabilized by large monovalent cations at the center of the plane, preferably potassium [16]. In human cells, G4 is mostly found in telomeres and regulatory regions of the genome, such as gene promoters with guanine-rich DNA sequences. One of the essential roles of G4 involves safeguarding DNA replication and transcription. As it is an extremely stable DNA structure, it must be unfolded into single strands to initiate DNA replication or transcription. Furthermore, a previous study revealed that among the G4s, c-Myc sequences found in nuclease-hypersensitive elements (NHE III) play the role of a transcription repressor [17].

Although most studies to date have dealt with *in vitro* G4 structure, these structures are believed to be present *in vivo* as well [18,19]. According to a recent review, there are surprisingly large numbers of potential G4 structures in the human genome, especially near the replication origin [20]. In particular, BLM has been suggested to be one of the active G4 helicases involved in chromosomal telomeric replication and transcription, during which it binds specifically to and unwinds G4. Recently, there have been many studies related to the G4 unfolding function of BLM. Several studies investigated G4 unfolding dynamics in real time by single-molecule Förster resonance energy transfer techniques [21–23]. For effective G4 unfolding, a single-stranded DNA overhang adjacent to the G4 was required [23]. One BLM core mutant that lost ATPase activity (residues 642–1290, K695M) was shown to be able to resolve various intramolecular G4 structures without ATP hydrolysis [24].

Moreover, the Zn–RQC–HRDC domain construct (residues 858–1298) was shown to facilitate ATP-independent G4 unfolding, even in the absence of the helicase domain [23]. The RQC–HRDC construct (residues 1066–1298) also showed significant Förster resonance energy transfer population changes within the folded conformation [23]. This implies that the RQC–HRDC construct can bind to and induce the partial unfolding of the G4, although it cannot mediate complete G4 unfolding. Also, the purified RQC domain with the zinc-binding subdomain (residues 994–1202) showed preferential binding to G4 DNA with high affinity [6]. The affinity to G4 DNA was comparable to that of full-length BLM. This implies that the RQC is a single binding site for G4 DNA and the G4 preference of BLM comes from the RQC [6].

The G4-forming sequence used in this study is Myc-22(14/23T). It is a derivative of Myc-2345 from the promoter of the c-Myc proto-oncogene, one of the most prevalently mutated genes in human cancers. Previous structural studies revealed the properties of the G4s formed in c-Myc promoters, including the 27-nucleotide-long purine-rich isoform Pu27, Myc-2345, and Myc-22(14/23T) [25,26] (Fig. 1b). Each of these has a similar, highly stable, intramolecular, propeller-type, parallel-stranded G4 structure in solutions containing a physiological concentration of K^+ [25,26] (Fig. 1c). Myc-22(14/23T)

has a single conformation. Also, all the chemical shifts of its imino protons, observed by the hydrogen bonding between guanines of the G4 structure, were revealed by a previous study [26].

Although the crystal structure of the BLM helicase core with duplex DNA was presented, how the RQC domain recognizes the G4 DNA structure is still poorly understood. In this study, we investigated the interaction of the BLM RQC domain with G4 DNA using NMR spectroscopy. The binding surface on RQC, mainly composed of the β-wing region, was identified using $^{1}H–^{15}N$ HSQC analysis. Also, we studied the exchange rate, binding affinity, enthalpy, and entropy of RQC–G4 binding using Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion NMR and isothermal titration calorimetry (ITC) experiments. Finally, our hydrogen–deuterium exchange (HDX) data showed that the isolated RQC domain could destabilize the G4 structure. In summary, we propose that the RQC domain binds G4 similarly to its duplex binding and induces partial disruption or destabilization of G4 DNA without any help from the other BLM domains.

Results

Imino proton NMR analysis of G4 upon addition of RQC

In order to investigate changes in G4 DNA upon RQC binding, we titrated RQC into G4 and observed the imino spectra of G4 at 25 °C (Fig. 2a). Imino protons involved in Hoogsteen base pairing appeared between 10 and 12 ppm, and all imino protons were assigned in a previous study [26]. In the presence of RQC, most peaks were shifted, and peak intensities decreased. Among these, G9 showed the most significant chemical shift change ($\Delta\delta_H$) of 0.039 ppm, and others showed similar chemical shift changes of ~0.012 ppm when comparing the spectra without RQC and in the presence of 1.2 equivalents of RQC. G9, G13, G20, and G22 exhibited substantial downfield chemical shift changes, and G11, G12, and G18 showed upfield chemical shift changes. G7, G8, G16, G17, and G21 were not shifted significantly (Fig. 2b). Those non-shifted guanines are located on the 5′ side of the G4 structure (Fig. 1c). This implies that G4 does not use this side as the binding surface for the RQC interaction. Lines were broadened severely in the presence of 2.5 equivalents of RQC and almost completely disappeared with 5 equivalents of RQC. Our one-dimensional (1D) NMR data imply that several imino protons involved in hydrogen bonding within the G4 were affected by RQC binding. Also, reduction of the intensity of imino proton peaks upon RQC binding can be considered as the effect of either the weakening of the hydrogen bonds or the millisecond timescale dynamics of the free and RQC-bound forms of the G4 or the combined effects of both.

Chemical shift perturbation analysis of BLM RQC upon addition of G4

We monitored the changes of the backbone amide signals of BLM RQC in $^{1}H–^{15}N$ HSQC spectra upon addition of Myc-22(14/23T) G4. The amide nitrogens and protons have been assigned previously [10]. Figure 3a shows the perturbed cross-peaks upon addition of G4 at 25 °C. The intensities of most peaks decreased as G4 was added. Some peaks,

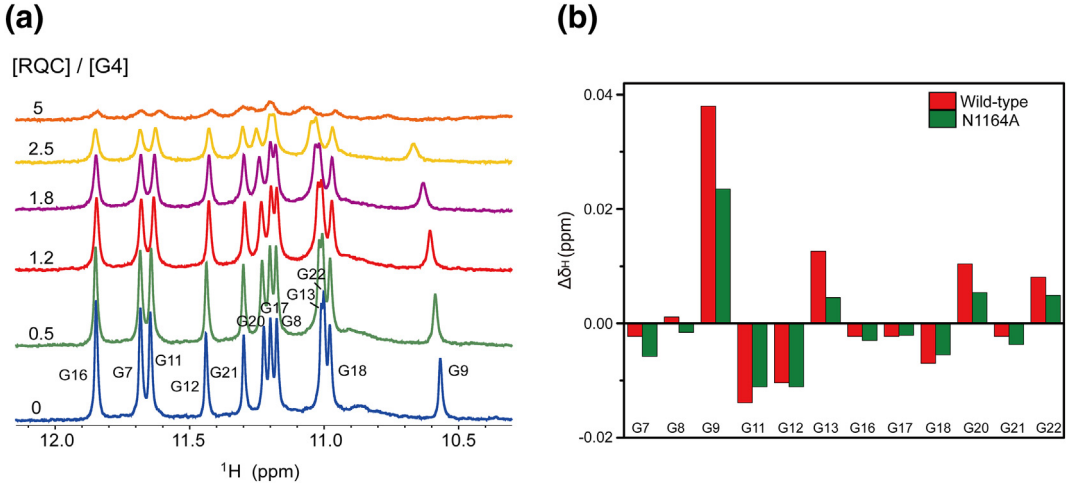

Fig. 2. (a) Stacked 1D imino proton spectra of Myc-22(14/23T) with increasing molar ratios of RQC. (b) $\Delta\delta_H$ values of the imino proton resonance of each guanine in the absence or presence of RQC wild-type (red) and RQC–N1164A mutant (green). $\Delta\delta_H = \delta_{H(bound)} - \delta_{H(free)}$. $\delta_{H(bound)}$ is the imino proton resonance at [G4]: [RQC] = 1: 1.2. $\delta_{H(free)}$ is the imino proton resonance in the free state.

such as S1138 and G1174, had decreased intensity in the presence of 2 molar equivalents of G4 because of line broadening (Fig. 3a). We also performed the same experiment at 35 °C, but the cross-peaks were still broadened with addition of the DNA (data not shown). Several peaks, such as I1161, N1164, I1168, and A1169, showed gradual chemical shift changes as G4 concentration increased. Figure 3b shows the histogram of the average chemical shift perturbation (CSP; $\Delta\delta_{avg}$) of each residue. V1103 in the α1–α2 loop and I1161, N1164, D1165, and I1168 in the β-wing region between β2 and β3 had $\Delta\delta_{avg}$ values greater than one standard deviation above the average. It is

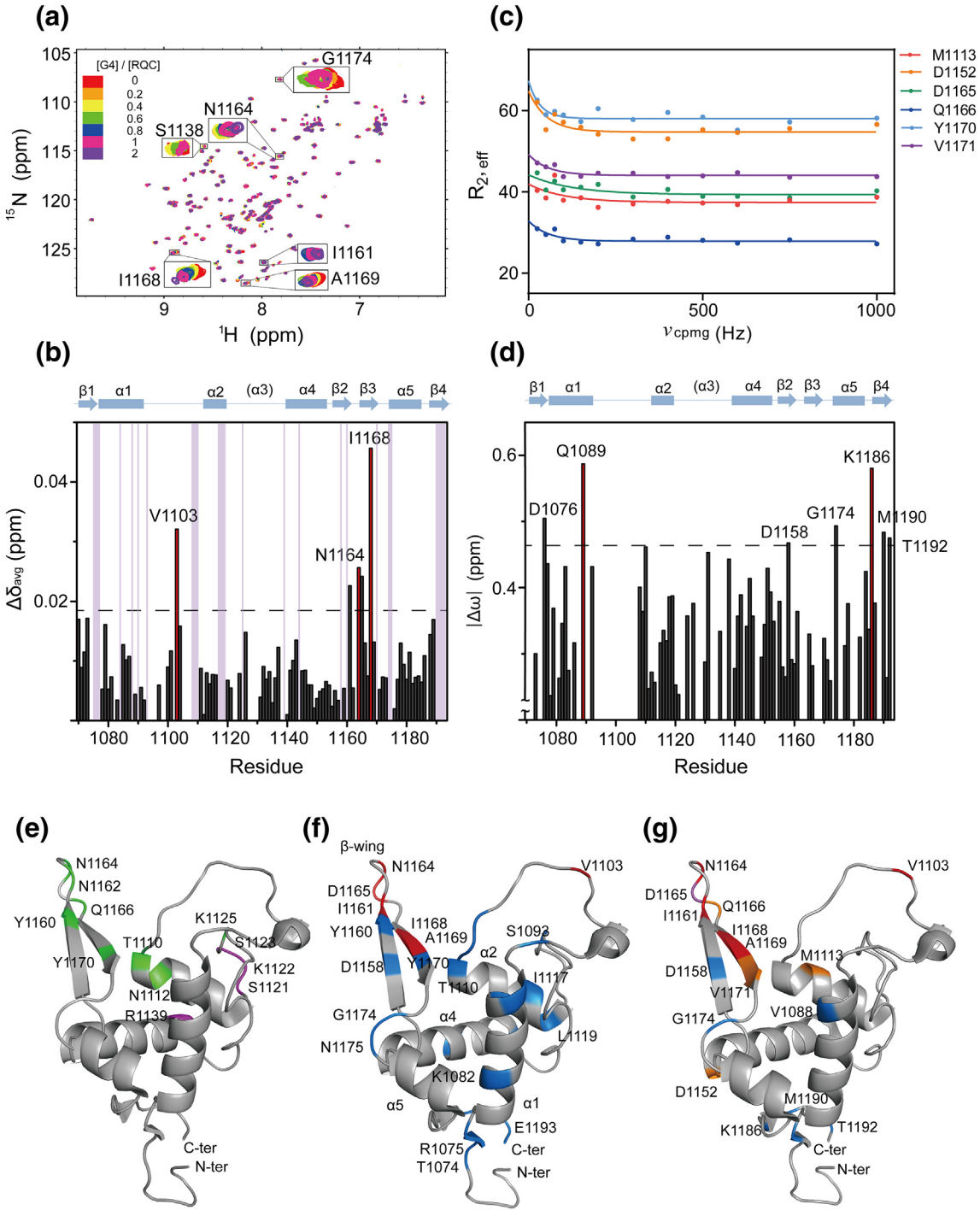

Fig. 3 (*legend on next page*)

noteworthy that N11162, N1164, Q1166, and Y1170 in the β-wing region are involved in the binding of a duplex DNA with a 3′ single-stranded overhang, based on a previous crystal structure (Fig. 3e) [7,9]. Other double-stranded (ds) DNA binding residues such as T1110 and N1112 in α2; S1121, K1122, S1123, and K1125 in the α2–α3 loop; and R1139 in α4 did not show significant changes or could not be analyzed in our experiments.

In our titration experiments, several peaks had broadened, which are shaded in purple in Fig. 3b. Most of those residues are located in the α2 and α4 helices and the N-terminal and C-terminal regions. Figure 3f shows the mapping of significantly perturbed residues and disappearing residues on the solution structure of RQC (PDB ID: 2MH9) [10]. While perturbed residues are clustered in the β-wing region, disappearing residues are dispersed throughout the structure. The line broadening of cross-peaks to the point of disappearance suggests that the RQC–G4 binding equilibrium could be within the millisecond timescale dynamics that are in the intermediate regime of the NMR chemical shift timescale.

^{15}N CPMG relaxation dispersion experiment

We hypothesized that RQC–G4 binding is in the regime of the milliesecond timescale, based on the line broadening of both partners. In order to specify the residues undergoing millisecond timescale dynamics in RQC upon G4 binding, we performed CPMG relaxation dispersion experiments for RQC with 0.05 equivalents of G4 and collected data at one static field. Our data showed that the $R_{2,\mathrm{eff}}$ values of specific residues change according to the ν_{CPMG}, while the $R_{2,\mathrm{eff}}$ values of others exhibit no change. Figure 3c shows the decay curves for residues whose $R_{2,\mathrm{eff}}$ values decreased as the ν_{CPMG} increased. This indicates that those residues experience millisecond timescale dynamics in the presence of G4. Notably, residues located in the β-wing region showed significant changes in the $R_{2,\mathrm{eff}}$ value, including D1165, Q1166, Y1170, and V1171. D1165 in the β-wing showed both a higher CSP than one standard deviation above the average and an $R_{2,\mathrm{eff}}$ curve in the CPMG experiment. M1113, located on the α2 helix, also showed decay in its $R_{2,\mathrm{eff}}$ value. However, this residue could not be observed in the ^{1}H–^{15}N HSQC titration experiment because of line broadening. The rate constant (k_{ex}) was calculated to be 289.3 ± 77.4 s^{-1} by global fitting using Eq. (3), which implies RQC exhibits exchange between two states (free and G4-bound) on the millisecond timescale.

Because we used the Luz–Meiboom exchange model for the fitting, p_{B} and $\Delta\omega$ were inseparable in Eq. (3). In order to extract $|\Delta\omega|$ from the data, we used p_{B} values estimated from the equation $\Delta\delta_{\mathrm{obs}} = p_{\mathrm{B}}\Delta\delta_{\mathrm{B}}$. The chemical shift differences of the state B and A ($\Delta\delta_{\mathrm{B}}$) were obtained from the ^{1}H–^{15}N titration data. The p_{B} value was estimated as 0.057 ± 0.007, which is similar to the theoretically calculated value based on K_{d}.

Figure 3d shows the histogram of $|\Delta\omega|$ of each residue. Several residues (Q1089, D1158, and G1174) that had broadened in the CSP experiment had larger $|\Delta\omega|$ than one standard deviation above the average. Figure 3g shows the residues with significant $\Delta\delta_{\mathrm{avg}}$, $R_{2,\mathrm{eff}}$, and $|\Delta\omega|$ on the three-dimensional structure of the BLM RQC, colored in red, orange, and blue, respectively. Most of them are clustered near the β-wing, which confirms the importance of this wing region for recognizing G4.

NMR analysis of N1164A mutant

Previous research pointed out that the polar contact between N1164 at the tip of the β-wing and the terminal base of the DNA is significant for the RQC-duplex binding [7]. N1164 also showed significant CSP in our ^{1}H–^{15}N HSQC analysis. In order to investigate the importance of this residue for G4 binding, we prepared the N1164A mutant and performed NMR experiments.

We observed the imino spectra of G4 with the N1164A mutant. Peaks were shown to be more

Fig. 3. (a) Overlaid ^{1}H–^{15}N HSQC spectra of ^{15}N-labeled RQC in the absence or presence of increasing molar ratios of Myc-22(14/23T). (b) CSPs ($\Delta\delta_{\mathrm{avg}}$) of ^{15}N-labeled RQC induced by 2.0 equivalents of Myc-22(14/23T). Residues with resonances that broadened upon titration are shaded in purple. The dotted line indicates one standard deviation higher than the average. (c) ^{15}N CPMG relaxation dispersion curves of M1113, D1152, D1165, Q1166, Y1170, and V1171 of RQC in the presence of 0.05 equivalents of Myc-22(14/23T) at 25 °C. (d) Estimated $|\Delta\omega|$ values of ^{15}N-labeled RQC from CPMG analysis. The dotted line indicates one standard deviation higher than the average. (e) Residues interacting with duplex DNA are mapped on the solution structure of RQC [7,10] (PDB ID 2MH9). Residues bound to the 3′ single-stranded overhang are colored in green, and those interacting with the major groove of DNA are colored in magenta. (f) Significantly perturbed residues are colored in red and disappearing residues are colored in blue on the solution structure of RQC. (g) Mapping of residues affected by G4 binding on the solution structure of RQC. Significantly perturbed residues upon addition of G4 are marked with red, residues with $R_{2,\mathrm{eff}}$ decay in the CPMG experiment are marked with orange, and residues that showed high $|\Delta\omega|$ values are marked with blue. D1165 showed both a large CSP ($\Delta\delta_{\mathrm{avg}}$) in the ^{1}H–^{15}N HSQC spectra and a decay curve of $R_{2,\mathrm{eff}}$ in the CPMG experiment is shown in purple.

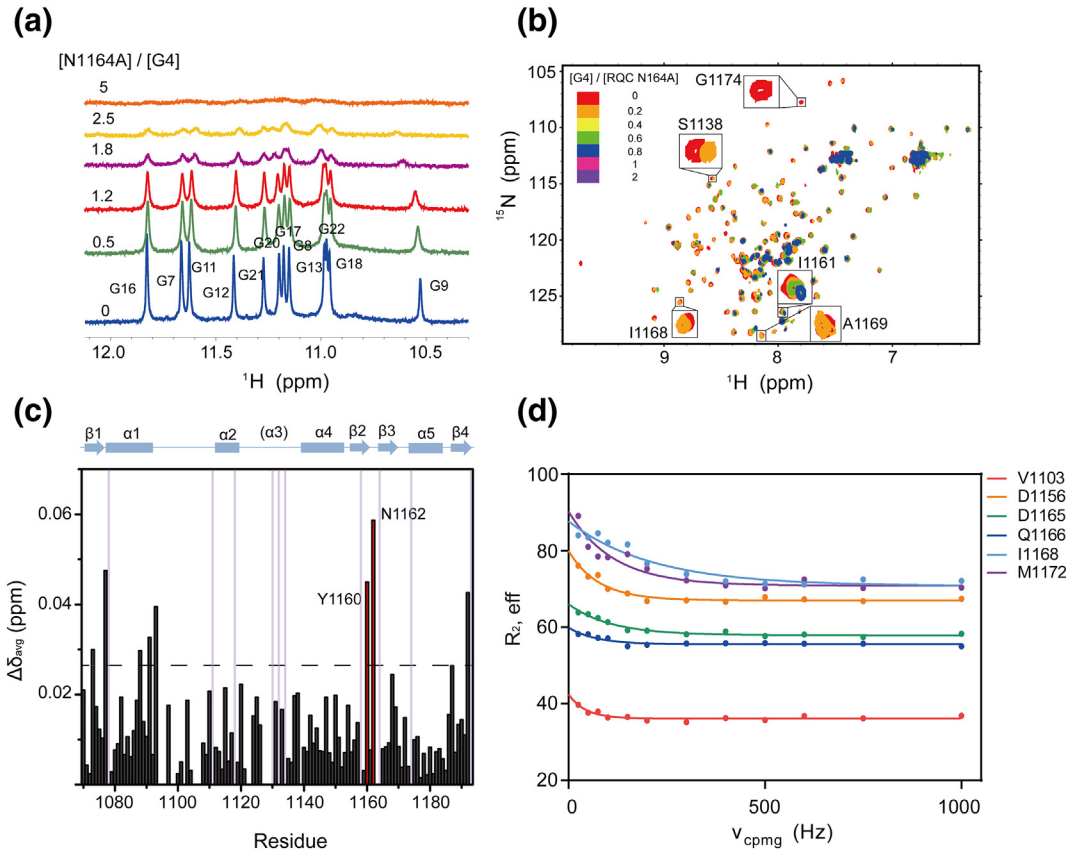

Fig. 4. (a) Stacked 1D imino proton spectra of Myc-22(14/23T) with increasing molar ratios of N1164A mutant. (b) Overlaid ^{1}H–^{15}N HSQC spectra of ^{15}N-labeled RQC–N1164A in the absence or presence of increasing molar ratios of Myc-22(14/23T). (c) CSPs ($\Delta\delta_{avg}$) of ^{15}N-labeled RQC–N1164A mutant induced by 0.2 equivalents of Myc-22(14/23T) binding. Resonances that broadened upon binding are shaded in purple. The dotted line indicates one standard deviation higher than the average. (d) ^{15}N CPMG relaxation dispersion curves of V1103, D1156, D1165, Q1166, I1168, and M1172 of the N1164A mutant in the presence of 0.05 equivalents of Myc-22(14/23T) at 25 °C.

broadened at the same [RQC]/[G4] value. The shifted imino peaks were identical to the peaks that shifted in the wild-type RQC spectra (Figs. 4a and 2a). We performed the ^{1}H–^{15}N HSQC titration experiments with G4 (Fig. 4b). Interestingly, the same peaks had broadened at much lower concentrations of G4 compared to wild-type RQC. With 0.4 equivalents of G4, most of the amide peaks of the N1164A mutant had broadened. $\Delta\delta_{avg}$ values were thus calculated using CSP data of RQC with 0.2 equivalents of G4 (Fig. 4c). The magnitude of the CSP of N1164A with 0.2 equivalents of G4 was comparable to the wild-type RQC with 2.0 equivalents of G4 (Fig. 4c and Fig. 3b). The β-wing region (Y1160 and N1162) of the mutant showed significant CSP upon G4 interaction (Fig. 4c). We observed that V1103 and I1168 showed large CSPs in the wild-type but not in the mutant.

Finally, the CPMG experiment was performed with the N1164A mutant. Figure 4d shows the $R_{2,eff}$ values of residues in the RQC–N1164A mutant, which decreased as the ν_{CPMG} increased. Similar to the results for the wild-type RQC, the residues located on the α1–α2 loop (V1103) and β-wing (D1165, Q1166, and I1168) showed a decay curve. D1165 and Q1166 on the tip of the β-wing showed a decay curve in both the wild-type and the N1164A mutant. V1103, D1156, I1168, and M1172 showed changes in $R_{2,eff}$ value in the mutant, but not in the wild-type. V1103 and I1168 were significantly perturbed residues in the titration experiment of wild-type RQC. D1156 and M1172 are located near the β-wing region that participates in G4 binding in wild-type RQC. The rate constant (k_{ex}) was estimated to be 885.5 ± 101.0 s^{-1}, a value 3-fold larger than for the wild-type protein. We applied the same methods used with the wild-type to estimate p_B from the titration data. The obtained p_B was 0.050 ± 0.011. Our data imply that the N1164A mutant could interact with G4, and the exchange rate between G4-bound and G4-unbound states is higher than that of the wild-type RQC.

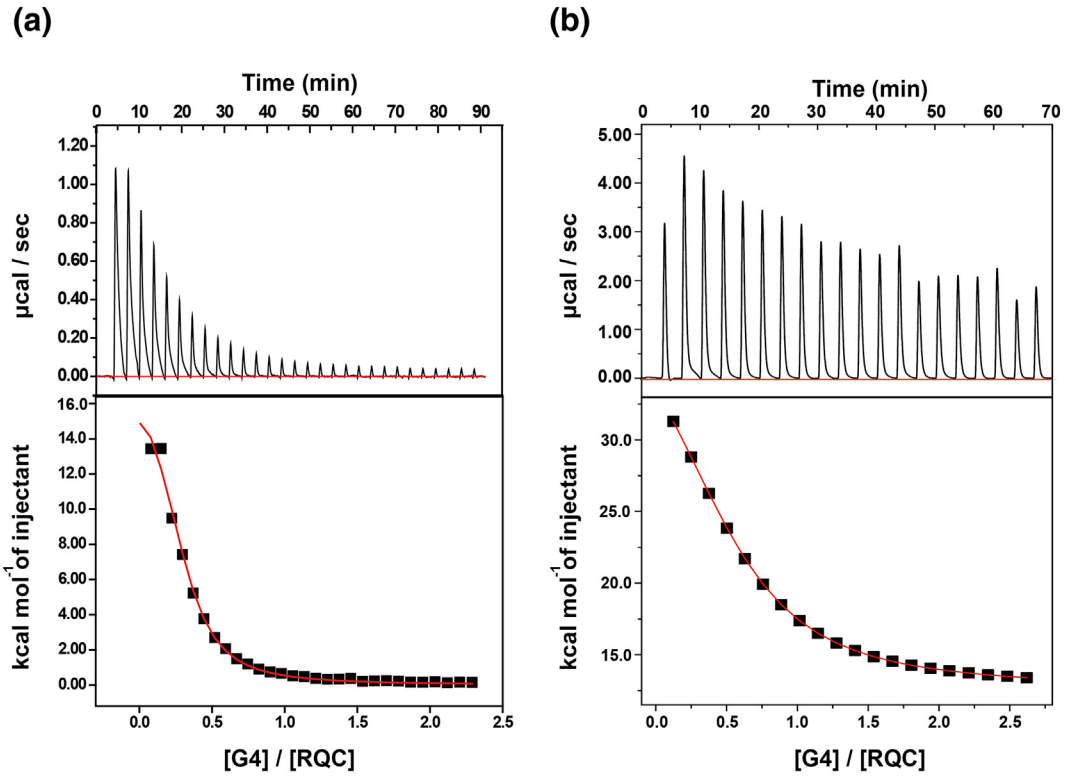

Fig. 5. The ITC binding isotherm of Myc-22(14/23T) titrated into (a) RQC wild-type and (b) RQC–N1164A mutant at 25 °C. Raw heat data (top) and the integrated heat data with the nonlinear regression fits (bottom) are shown.

Thermodynamic parameters of RQC binding to G4

To further characterize the RQC–G4 interaction, we employed ITC. Figure 5a shows the ITC binding isotherm of wild-type RQC with G4. The enthalpy change for the binding was highly positive (ΔH = +77.0 ± 3.5 kJ·mol^{-1}). The change of entropy also showed a positive value (ΔS = +364 J·K^{-1}·mol^{-1}). These data showed that RQC–G4 binding is entropically driven. The dissociation constant, K_d, was calculated to be 2.9 ± 0.4 µM by a one-site model fitting (Table 1). Figure 5b shows the results of the ITC experiment using the BLM RQC–N1164A construct. The interaction was still endothermic. The K_d value was calculated to be 13.9 ± 3.47 µM, which is about 5-fold higher than wild-type RQC (Table 1).

Table 1. Dissociation constants (K_d) and association/dissociation rate constants (k_{on} and k_{off}) for the G4 DNA binding of the wild-type RQC and N1164A mutant

Protein	K_d (µM)	k_{on} (×10^7 M^{-1}s^{-1})	k_{off} (s^{-1})
Wild-type	2.9 ± 0.4	9.45 ± 2.52	274.0 ± 73.0
N1164A	13.9 ± 3.47	6.05 ± 0.69	841.2 ± 96.0

These results showed that the mutation of N1164 to alanine, which had significant CSPs, weakened the binding affinity. Considering k_{ex} and K_d together, we can calculate k_{off} ($k_{off} = k_{ex} \times p_{free}$) of the wild-type as 274.0 ± 73.0 s^{-1}, and k_{on} ($k_{on} = k_{off}/K_d$) as (9.45 ± 2.52) × 10^7 M^{-1}s^{-1}. The values of the N1164A mutant are as follows: k_{off} 841.2 ± 96.0 s^{-1} and k_{on} (6.05 ± 0.69) × 10^7 M^{-1}s^{-1} (Table 1). The mutant thus has a 1.5-fold slower k_{on} and a 3-fold faster k_{off} than the wild-type.

Hydrogen–deuterium exchange

In order to investigate the stability of the Hoogsteen base pairings of G4 in the presence of RQC, we performed HDX experiments. Figure 6a shows the 1D NMR spectra of G4 6, 500, and 1000 min after D_2O was added. Only the peaks of G8, G12, G17, and G21, which are located in the middle plane of the G4 structure, the region most protected from exposure to the buffer, were observed. In the absence of RQC, these imino peaks remained visible 1000 min after deuterium was added, whereas they broadened in the presence of RQC. Figure 6b shows the normalized intensity of the imino peaks *versus* time. The decay time constant (τ) of G4 without RQC was 4994 to 7850 min, and that of G4 with 0.3 equivalents of RQC

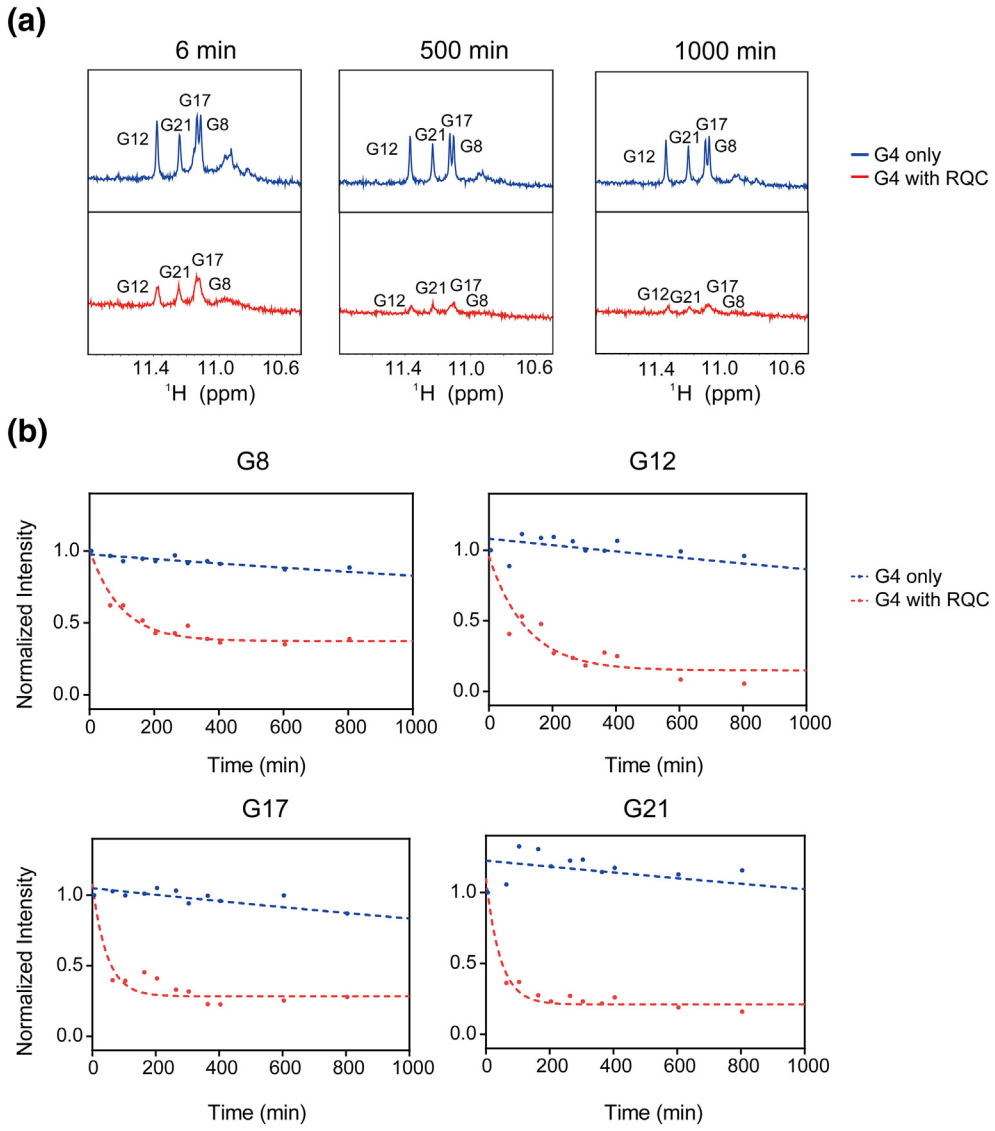

Fig. 6. (a) 1D NMR spectra at 6, 500, and 1000 min after deuterium was added to samples of G4 only (blue) or G4 mixed with 0.3 equivalents of RQC (red). (b) The normalized imino peak intensities of G8, G12, G17, and G21 after buffer exchange from hydrogen to deuterium as a function of time in the absence (blue) or presence (red) of 0.3 equivalents of RQC.

was 40.18 to 119.40 min (Table 2). The decay time constant became about 85-fold shorter in the presence of RQC, on average. Our data show that the HDX rates of G4 imino protons are accelerated in the presence of RQC. This strongly indicates that the BLM RQC–G4 interaction contributes to the destabilization of G4 base pairing.

Table 2. Decay time constants (τ) of guanine imino protons with and without RQC

Base	DNA only (min)	With RQC (min)
G8	4994 ± 183	96.36 ± 34.33
G12	7850 ± 437	119.40 ± 36.73
G17	5221 ± 310	40.18 ± 16.43
G21	7728 ± 521	45.42 ± 12.20

Discussion

Previous reports have shown that BLM RQC is responsible for specific recognizing G4 structures [27] and that the combined RQC–HRDC domain (residues 858–1298) can unfold G4 DNA in an ATP-independent manner [23]. These data imply that the detailed information of RQC–G4 interaction is necessary for the understanding of the G4 processing of BLM. However, how RQC recognizes G4 is still unclear. Furthermore, it has not been known

whether the RQC alone, without the other domains (helicase and HRDC), can disrupt or unfold the G4 structure.

In this study, we investigated the RQC–G4 interaction using NMR spectroscopy. Upon RQC–G4 binding, several imino protons of the G4 and amide signals of the RQC showed notable perturbations. The imino proton signals that did not shift were in the guanines positioned on the 5′ side of the plane (G7, G8, and G16). This implies that the 5′ side of the G4 plane does not make contact with RQC, consistent with the fact that BLM unfolds its substrates in the 3′ to 5′ direction [3,4].

According to the crystal structure of the BLM helicase core and dsDNA complex, the β-wing region is the binding surface for the minor groove face of duplex DNA with a 3′ single-stranded overhang [7,9]. Previous studies revealed that the polar interface between N1164 on the tip of the β-wing and the terminal base of the DNA is the most crucial contact for the duplex–RQC interaction [7,9]. In addition to that, residues in the β-wing region (Y1160, N1162, Q1166, Y1170) and α2 (T1110, N1112) also bind with the phosphate ions on the DNA backbone [7] (Fig. 3e). This is consistent with our CSP analysis, in which I1161, N1164, D1165, and I1168 showed significant chemical shift changes upon G4 binding (Fig. 3f). This means that the β-wing region makes contacts with G4 DNA, possibly from the 3′ side. Our ITC data also support that the β-wing region is important for the G4 binding. The N1164A mutant had a 5-fold higher K_d than the wild-type protein, suggesting that the contact between RQC and G4 is mediated by N1164. In the previous crystal structure, residues in the α2–α3 loop (S1121, K1122, S1123, and K1125) were also pointed toward the duplex DNA major groove binding region [7]. Those residues were not perturbed significantly in the current study, which implies that that region does not participate in the G4 interaction.

Interestingly, severe line broadening was observed in the NMR spectra of reciprocal titrations. Both imino peaks of DNA and amide signals of protein broadened upon addition of the binding partner. In particular, we found that G4 DNA–N1164A binding induces more line broadening than wild-type RQC binding. The line broadening phenomenon implies that exchange between the free and the bound forms occurs on the millisecond timescale, which is in the intermediate regime of the NMR chemical shift timescale. At the same time, the G4 destabilization by RQC could also contribute to the line broadening. We hypothesize that the N1164A mutant destabilizes the G4 more than the wild-type based on the greater broadening of the imino signals at the same protein/DNA ratio. Consequently, this could cause more heterogeneous dynamic states of the mutant protein to be observed as severe line broadening. In this way, the $^{1}H–^{15}N$ HSQC titration data could reflect the intrinsic heterogeneous conformational states caused by the G4 destabilization. These could be the reason why the nitrogen chemical shift differences between the free and the G4-saturated states of the protein ($\Delta\delta_{N,HSQC}$) were not perfectly matched with $\Delta\omega$ from the CPMG analysis in this study.

Our CPMG analysis showed that residues with changes in $R_{2,eff}$ were clustered in the β-wing region (D1152, D1165, Q1166, Y1170, and V1171) and the α2 helix (M1113) (Fig. 3g). This is consistent with our CSP analysis and a previous crystal structure [7]. This result strongly indicated that the G4–RQC interaction is mainly achieved by the β-wing region of RQC.

Taken together, our data showed that the β-wing region and the α2 helix are essential to G4 binding. Compared to the duplex binding of RQC [7], G4 binding is similar except that the G4 does not contact the α2–α3 loop region, while the major groove of duplex DNA does. The possible reason could be the different geometry of the DNA structures. The previous crystal structure [7] revealed a detailed binding site for the duplex DNA binding, and the distance between atoms could be measured. The distance between the β-wing binding site and the α2–α3 loop binding site was ~23 Å (about 7 base pairs). As the G4 used in this study is more compact than B-form DNA, the distance from the 3′ end to the bottom of the guanine plane is ~15 Å. Furthermore, our 1D NMR experiment showed that G4 does not utilize the 5′ side for the binding with RQC, making the total length of the interacting part of G4 even shorter than 15 Å. Based on this, we speculate that the G4 is not long enough to use both binding sites (β wing and α2–α3 loop), while the duplex is (Fig. 7).

In addition to the exchange between the free and the bound form, disappearance of the imino peaks of the 1D spectrum could also be caused by the weakening of hydrogen bonds. It has been pointed out that loss of hydrogen bonding, which results in conformational heterogeneity, induces the reduction of imino signals from the G4 structure [28,29]. Therefore, this suggests that RQC might be able to destabilize the G4 without other parts of BLM by breaking some hydrogen bonds and inducing conformational changes. The HDX experiment confirms this weakening of hydrogen bonds caused by RQC. The HDX NMR technique is a powerful tool for investigating the folding stability of the molecule by observing individual hydrogen bonding [30]. As the calculated decay time constant became much shorter in the presence of RQC, it shows that RQC alone can destabilize the G4. This is consistent with our ITC data, which showed that the RQC–G4 interaction is endothermic and entropically driven. The positive entropy change could result from the destabilization of G4. We hypothesize that RQC

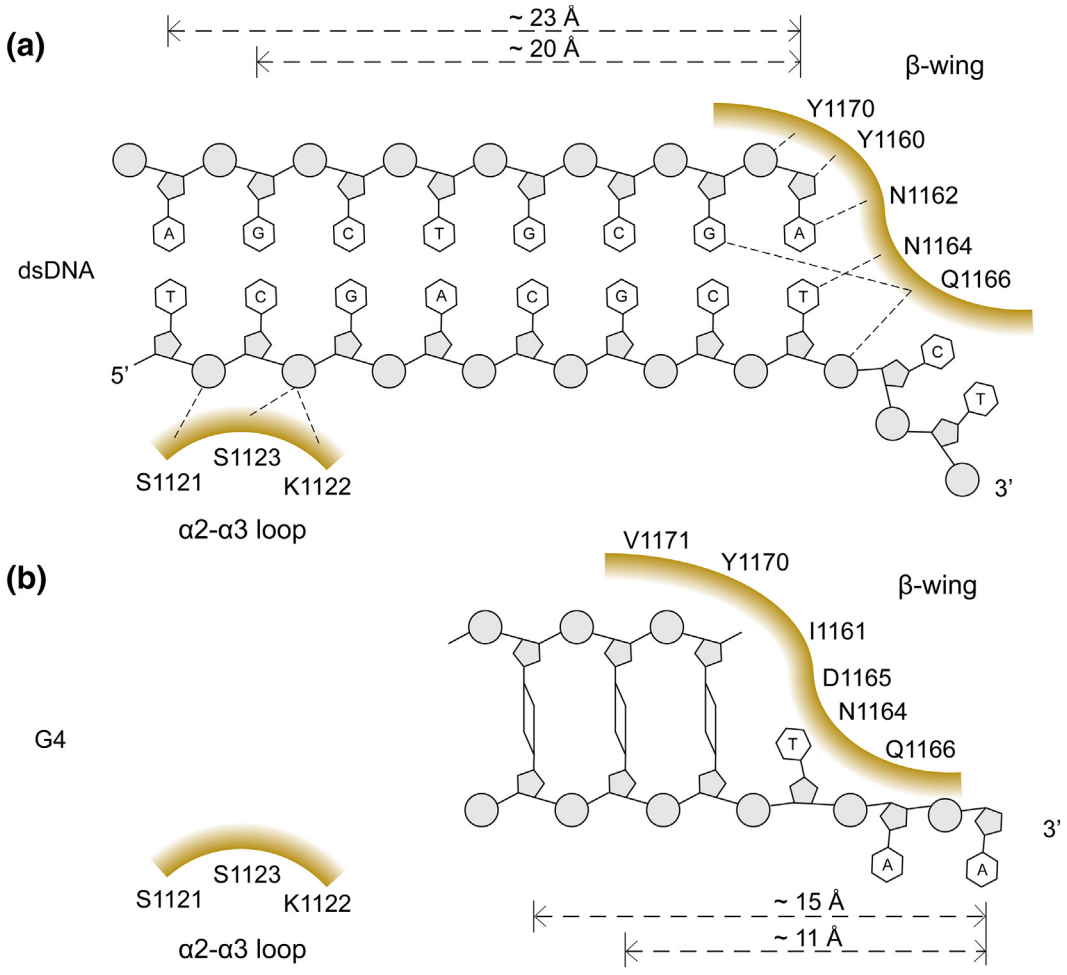

Fig. 7. Schematic drawing of RQC contacts with (A) dsDNA [7] and (B) G4.

binding destabilizes G4 and breaks some hydrogen bonds. Then, the potassium cation placed in the middle of each plane might release, leading to an increase in entropy. A similar entropically driven binding process, caused by ion release induced by protein binding, was reported previously [31]. However, the magnitude of the van't Hoff enthalpy of folding of a similar G4 structure was much larger than the measured enthalpy value in this study [32]. This suggests that RQC binding alone does not provide enough energy to unfold the G4 structure completely. Based on this, we suggest that RQC binding can only partially disrupt or destabilize the G4 structure.

In summary, we provide information on the interaction of BLM RQC with the G4 from the Myc promoter. Our CSP analysis and CPMG data indicate that the β-wing region of the RQC is important for G4 DNA binding, most likely on the 3′ side of G4. HDX results and ITC data show that the RQC destabilizes hydrogen bonds within G4 DNA. This understanding of initial G4 recognition by RQC gives new insights into strategies for the regulation of G4 metabolism.

Materials and Methods

Sample preparation

The BLM RQC domain (residues 1067–1210) was expressed and purified as previously described except for using a HiLoad 16/600 Superdex 75 pg gel filtration column (GE Healthcare) with 20 mM Tris and 100 mM KCl (pH 7.0) buffer for chromatography [10]. The ^{15}N-labeled BLM RQC domains were prepared by using $^{15}NH_4Cl$ (Cambridge Isotope Laboratories) as the sole source of nitrogen in M9 minimal media. Myc-22 (14/23T) was purchased from IDT Inc. We dissolved Myc-22(14/23T) in 20 mM Tris and 100 mM KCl (pH 7.0) buffer to prepare 1 mM of the G4 solution stock. The prepared DNA solutions were heated to 95 °C for 10 min and cooled to room temperature for 1 h. After cooling, samples were stored at 4 °C.

^{1}H and ^{1}H–^{15}N HSQC NMR experiments

All NMR experiments were performed on a Bruker Avance II 900 MHz spectrometer equipped with a cryogenic probe (Korea Basic Science Institute, Ochang) except the HDX experiment. All NMR samples were prepared in 20 mM Tris and 100 mM KCl buffer (pH 7.0). Topspin (Bruker) was used for data processing and SPARKY software was used for data analysis [33].

1D proton spectra of Myc-22(14/23T) in the absence or presence of RQC were obtained at 25 °C. ^{1}H–^{15}N HSQC spectra of ^{15}N-labeled RQC in the absence or presence of Myc-22(14/23T) were also obtained at 25 °C. The concentration of the ^{15}N-labeled RQC was fixed at 0.3 mM and G4 was added for each ratio. CSP values ($\Delta\delta_{avg}$) were calculated using the following equation.

$$\Delta\delta_{avg} = \sqrt{(\Delta\delta_H)^2 + (\Delta\delta_N/5.88)^2} \quad (1)$$

CPMG experiments

The CPMG experiment was performed using 0.7 mM ^{15}N-labeled RQC and RQC mixed with 0.05 molar equivalent of G4 at 25 °C. Experiments employed a constant relaxation delay T_{relax} of 20 ms and 12 values of $\nu_{CPMG} = 1/(2\tau_{CP})$ ranging from 25 to 1000 Hz, where τ_{CP} is the delay between consecutive pulses. Transverse relaxation rates $R_{2,eff}$ were calculated for each cross-peak signal at each value using the following equation.

$$R_{2,eff}(\nu_{CPMG}) = -\frac{1}{T_{relax}} \ln\left\{\frac{I(\nu_{CPMG})}{I_0}\right\} \quad (2)$$

$I(\nu_{CPMG})$ is the intensity of each peak at values of each ν_{CPMG}, and I_0 is the intensity of the peak with 0 Hz for ν_{CPMG}. Calculated $R_{2,eff}$ values are globally fitted to the following equation to deduce k_{ex}.

$$R_{2,eff}(\nu_{CPMG}) = R_2^0 + \frac{\Phi_{ex}}{k_{ex}}\left\{1 - \frac{4\nu_{CPMG}}{k_{ex}} \tanh\left(\frac{k_{ex}}{4\nu_{CPMG}}\right)\right\} \quad (3)$$

R_2^0 is the intrinsic transverse relaxation rate, and Φ_{ex} is calculated by the following equation: $\Phi_{ex} = p_A p_B(\Delta\omega)^2$, where $\Delta\omega$ is the chemical shift difference between the two states, and p_A and p_B are the relative populations of state A and state B. k_{ex} is the exchange rate between the two states.

p_B was estimated by using $\Delta\delta_{obs} = p_B\Delta\delta_B$. $\Delta\delta_{obs}$ is the chemical shift difference between the sample used in the CPMG experiments and the free protein, and $\Delta\delta_B$ is obtained from the ^{1}H–^{15}N HSQC titration data.

HDX experiments

The HDX experiment was performed using a Varian Inova 700 MHz spectrometer equipped with a cryogenic probe (GNU, Jinju). Samples of 0.7 mM G4 and 0.7 mM G4 mixed with 0.3 molar equivalents of RQC were lyophilized, and deuterium oxide (99%) was added right before the NMR experiment acquisition. 1D proton NMR spectra were observed continuously every 2 min after all the hydrogen of the buffer was exchanged to deuterium at 25 °C. The first NMR spectra were observed 6 min after the deuterium was added because of the locking/shimming process. The intensity of each guanine peak was normalized using the methyl peak of thymine, which is not affected by HDX and appears at 1.54 ppm. The decay time constant was calculated by fitting to a one-phase decay model in GraphPad Prism version 7.01 for Windows (GraphPad Software, La Jolla, CA, USA; www.graphpad.com).

Isothermal titration calorimetry

A Microcal VP-ITC was used for the ITC experiment (KAIST, Daejeon). Concentrated Myc-22(14/23T) DNA solution (450 μM) and a 45 μM solution of BLM RQC domain were used. Protein and DNA samples were extensively dialyzed against 4 L of 20 mM Tris and 100 mM KCl (pH 7.0) buffer. For the first injection, 1 μL of DNA stock was added to the protein sample. From the second injection on, 10 μL of DNA solution was added into the protein stock each time, and 30 injections were performed. Other experimental conditions were as follows: interval, 180 s; stirring speed, 350 rpm; and cell temperature, 25 °C.

CRediT authorship contribution statement

Sungjin Lee: Conceptualization, Data curation, Formal analysis, Writing - original draft. **Ae-Ree Lee:** Data curation, Writing - original draft. **Kyoung-Seok Ryu:** Data curation, Formal analysis.
Joon-Hwa Lee: Conceptualization, Formal analysis, Funding acquisition, Writing - original draft, Writing - review & editing. **Chin-Ju Park:** Conceptualization, Formal analysis, Funding acquisition, Writing - original draft, Writing - review & editing.

Acknowledgments

This work was supported by the National Research Foundation of Korea (Grants 2018R1A2B6004388 to C.-J.P. and 2017R1A2B2001832 to J.-H.L.), which is

funded by the Korean government (MSIT); by the Korean Basic Science Institute under R&D program (Project No. D39700 to C.-J.P.) supervised by MSIT, Korea; by GIST Research Institute funded by the GIST in 2018 (to C.-J.P.); and by the Samsung Science and Technology Foundation (SSTF-BA1701-10 to J.-H.L.). We thank Prof. B.-S. Choi (KAIST, Daejon) for the use of their ITC instrument. We thank the GNU Central Instrument Facility and the high-field NMR facility at the Korean Basic Science Institute (KBSI, Ochang) for performing NMR experiments, and Dr. M. Stauffer of Scientific Editing Solutions for editing the manuscript.

Declarations of Interest: None.

Received 12 October 2018;
Received in revised form 1 January 2019;
Accepted 5 January 2019
Available online 10 January 2019

Keywords:
G-quadruplex;
Bloom syndrome protein;
RQC domain;
NMR;
hydrogen–deuterium exchange

†S.L. and A.-R.L. contributed equally to this work.

Abbreviations used:
BLM, Bloom syndrome protein; RQC, RecQ C-terminal domain; G4, G-quadruplex; ITC, isothermal titration calorimetry; HDX, hydrogen–deuterium exchange; CPMG, Carr–Purcell–Meiboom–Gill; CSP, chemical shift perturbation.

References

[1] R.J. Monnat, Human RECQ helicases: roles in DNA metabolism, mutagenesis and cancer biology, Semin. Cancer Biol. 20 (2010) 329–339.
[2] N.A. Ellis, J. Groden, T.-Z. Ye, J. Straughen, D.J. Lennon, S. Ciocci, M. Proytcheva, J. German, The Bloom's syndrome gene product is homologous to RecQ helicases, Cell 83 (1995) 655–666.
[3] V.A. Bohr, Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance, Trends Biochem. Sci. 33 (2008) 609–620.
[4] G.H. Nguyen, W. Tang, A.I. Robles, R.P. Beyer, L.T. Gray, J.A. Welsh, A.J. Schetter, K. Kumamoto, X.W. Wang, I.D. Hickson, N. Maizels, R.J. Monnat, C.C. Harris, Regulation of gene expression by the BLM helicase correlates with the presence of G-quadruplex DNA motifs, Proc. Natl. Acad. Sci. 111 (2014) 9905–9910.
[5] A. Kamath-loeb, L.A. Loeb, M. Fry, The Werner syndrome protein is distinguished from the Bloom syndrome protein by its capacity to tightly bind diverse DNA structures, PLoS One 7 (2012), e30189.
[6] M.D. Huber, M.L. Duquette, J.C. Shiels, N. Maizels, A conserved G4 DNA binding domain in RecQ family helicases, J. Mol. Biol. 358 (2006) 1071–1080.
[7] J.A. Newman, P. Savitsky, C.K. Allerston, A.H. Bizard, Ö. Özer, K. Sarlós, Y. Liu, E. Pardon, J. Steyaert, I.D. Hickson, O. Gileadi, Crystal structure of the Bloom's syndrome helicase indicates a role for the HRDC domain in conformational changes, Nucleic Acids Res. 43 (2015) 5221–5235.
[8] Y.M. Kim, B.-S. Choi, Structure and function of the regulatory HRDC domain from human Bloom syndrome protein, Nucleic Acids Res. 38 (2010) 7764–7777.
[9] M.K. Swan, V. Legris, A. Tanner, P.M. Reaper, S. Vial, R. Bordas, J.R. Pollard, P.A. Charlton, J.M.C. Golec, J.A. Bertrand, Structure of human Bloom's syndrome helicase in complex with ADP and duplex DNA, Acta Crystallogr. Sect. D Biol. Crystallogr. 70 (2014) 1465–1475.
[10] C.-J. Park, J. Ko, K.-S. Ryu, B.-S. Choi, Solution structure of the RecQ C-terminal domain of human Bloom syndrome protein, J. Biomol. NMR 58 (2014) 141–147.
[11] A. Sato, M. Mishima, A. Nagai, S.-Y. Kim, Y. Ito, T. Hakoshima, J.-G. Jee, K. Kitano, Solution structure of the HRDC domain of human Bloom syndrome protein BLM, J. Biochem. 148 (2010) 517–525.
[12] K. Kitano, S.-Y. Kim, T. Hakoshima, Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN, Structure 18 (2010) 177–187.
[13] B. Heddi, A.T. Phan, Structure of human telomeric DNA in crowded solution, J. Am. Chem. Soc. 133 (2011) 9824–9833.
[14] A. Artese, G. Costa, S. Distinto, F. Moraca, F. Ortuso, L. Parrotta, S. Alcaro, Toward the design of new DNA G-quadruplex ligands through rational analysis of polymorphism and binding data, Eur. J. Med. Chem. 68 (2013) 139–149.
[15] A.I. Karsisiotis, M.W. da Silva, Structural probes in quadruplex nucleic acid structure determination by NMR, Molecules 17 (2012) 13073–13086.
[16] S. Haider, G.N. Parkinson, S. Neidle, Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex, J. Mol. Biol. 320 (2002) 189–200.
[17] T. Lemarteleur, D. Gomez, R. Paterski, E. Mandine, P. Mailliet, J.F. Riou, Stabilization of the c-myc gene promoter quadruplex by specific ligands' inhibitors of telomerase, Biochem. Biophys. Res. Commun. 323 (2004) 802–808.
[18] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H.J. Lipps, Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo, Nat. Struct. Mol. Biol. 12 (2005) 847–854.
[19] N. Maizels, Dynamic roles for G4 DNA in the biology of eukaryotic cells, Nat. Struct. Mol. Biol. 13 (2006) 1055–1059.
[20] O. Mendoza, A. Bourdoncle, J.-B. Boulé, R.M. Brosh, J.-L. Mergny, G-quadruplexes and helicases, Nucleic Acids Res. 44 (2016) 1989–2006.
[21] J.B. Budhathoki, E.J. Stafford, J.G. Yodh, H. Balci, ATP-dependent G-quadruplex unfolding by Bloom helicase exhibits low processivity, Nucleic Acids Res. 43 (2015) 5961–5970.
[22] W.-Q. Wu, X.-M. Hou, M. Li, S.-X. Dou, X.-G. Xi, BLM unfolds G-quadruplexes in different structural environments through different mechanisms, Nucleic Acids Res. 43 (2015) 4614–4626.
[23] S. Chatterjee, J. Zagelbaum, P. Savitsky, A. Sturzenegger, D. Huttner, P. Janscak, I.D. Hickson, O. Gileadi, E. Rothenberg, Mechanistic insight into the interaction of BLM

helicase with intra-strand G-quadruplex structures, Nat. Commun. 5 (2014) 5556.
[24] J.B. Budhathoki, S. Ray, V. Urban, P. Janscak, J.G. Yodh, H. Balci, RecQ-core of BLM unfolds telomeric G-quadruplex in the absence of ATP, Nucleic Acids Res. 42 (2014) 11528–11545.
[25] A.T. Phan, Y.S. Modi, D.J. Patel, Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter, J. Am. Chem. Soc. 126 (2004) 8710–8716.
[26] A. Ambrus, D. Chen, J. Dai, R.A. Jones, D. Yang, Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization, Biochemistry 44 (2005) 2048–2058.
[27] P. Mohaghegh, J.K. Karow, R.M. Brosh, V.A. Bohr, I.D. Hickson, The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases, Nucleic Acids Res. 29 (2001) 2843–2849.
[28] Z.A.E. Waller, S.A. Sewitz, S.-T.D. Hsu, S. Balasubramanian, A small molecule that disrupts G-quadruplex DNA structure and enhances gene expression, J. Am. Chem. Soc. 131 (2009) 12628–12633.
[29] M.J. Morris, K.L. Wingate, J. Silwal, T.C. Leeper, S. Basu, The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells, Nucleic Acids Res. 40 (2012) 4137–4145.
[30] M.-H. Li, Z.-F. Wang, M.H.-J. Kuo, S.-T.D. Hsu, T.-C. Chang, Unfolding kinetics of human telomeric G-quadruplexes studied by NMR spectroscopy, J. Phys. Chem. B 118 (2014) 931–936.
[31] A. Loregian, E. Sinigalia, B. Mercorelli, G. Palù, D.M. Coen, Binding parameters and thermodynamics of the interaction of the human cytomegalovirus DNA polymerase accessory protein, UL44, with DNA: implications for the processivity mechanism, Nucleic Acids Res. 35 (2007) 4779–4791.
[32] E. Hatzakis, K. Okamoto, D. Yang, Thermodynamic stability and folding kinetics of the major G-quadruplex and its loop isomers formed in the nuclease hypersensitive element in the human c-Myc promoter: effect of loops and flanking segments on the stability of parallel-stranded intramolecul, Biochemistry 49 (2010) 9152–9160.
[33] T.D. Goddard, D.G. Kneller, SPARKY 3, University of California, San Francisco, CA, 2003.