DOI 10.5012/bkcs.2011.32.2.719

Population Analysis of the Intermediate Complex States During B-Z Transition of Non-CG-repeat DNA Duplexes Induced by the Zα Domain of Human ADAR1

Eun-Hae Lee, Yeo-Jin Seo, Hee-Eun Kim, Yeon-Mi Lee, Chong-Man Kim,†,* and Joon-Hwa Lee*

*Department of Chemistry and Research Institute of Natural Science, Gyeongsang National University, Gyeongnam 660-701, Korea. *E-mail: joonhwa@gnu.ac.kr*
*†Department of Industrial and Management Engineering, Myongji University, Gyeonggi 449-728, Korea *E-mail: chongman@mju.ac.kr*
Received September 30, 2010, Accepted December 10, 2010

Key Words: NMR, Hydrogen exchange, Z-DNA, Confidence limit, ADAR1

Z-DNA contains nucleic acid bases in alternating *anti*- and *syn*-conformations along the nucleotide chain and has only one groove that is similar to the minor groove of B-DNA.[1-3] Z-DNA is in a higher energy conformation than B-DNA and is stabilized by negative supercoiling generated *in vivo*.[2,3] Human ADAR1 has two left-handed Z-DNA binding domains at its NH_2-terminus, Zα and Zβ, preferentially binds Z-DNA, rather than B-DNA, with high binding affinity.[4-6] The co-crystal structure of the Zα domain of human ADAR1 ($Z\alpha_{ADAR1}$) bound to Z-DNA revealed that one monomeric $Z\alpha_{ADAR1}$ domain binds to one strand of double-stranded DNA and a second $Z\alpha_{ADAR1}$ monomer binds to the opposite strand with two-fold symmetry with respect to the DNA helical axis.[7] A structural study showed that $Z\alpha_{ADAR1}$ binds to the Z-conformation of non-CG-repeat DNA duplexes through a common structural feature rather than by a specific sequence or structural alternations.[8] A previous NMR study on a $d(CGCGCG)_2$-$Z\alpha_{ADAR1}$ complex[9] suggests an *active-mono* B-Z transition mechanism (*see* Fig. 1) in which the $Z\alpha_{ADAR1}$ protein first binds to B-DNA and then converts it to left-handed Z-DNA, a conformation that is then stabilized by the additional binding of a second $Z\alpha_{ADAR1}$ molecule.

Recently, we have reported NMR hydrogen exchange data of complexes between $Z\alpha_{ADAR1}$ and the non-CG-repeat DNA duplexes, $d(CACGTG)_2$ [referred to as CA6] or $d(CGTACG)_2$ [referred to as TA6], with a variety of protein-to-DNA (P/N) molar ratios.[10] The k_{ex} data for the G4b of the CA6-$Z\alpha_{ADAR1}$ complex and for the G2b of the TA6-$Z\alpha_{ADAR1}$ complex showed significant changes as the Z-DNA fraction (f_Z) was increased (meaning that the P/N ratio increased) (*see* Fig. 2). These changes of the k_{ex} data can be explained by the presence of mixtures of two imino protons from B-form DNA (referred to as **B**) and B-DNA-$Z\alpha_{ADAR1}$ complex (referred to as **BP**) in the imino peaks as given by Eq. 1:[10]

$$k_{ex} = \frac{[B]k_{ex}^{B} + [BP]k_{ex}^{BP}}{[B]+[BP]} = k_{ex}^{B} + \frac{[BP]}{1-Z_t}(k_{ex}^{BP} - k_{ex}^{B}) \quad (1)$$

where k_{ex}^{B} and k_{ex}^{BP} are the k_{ex} of the imino protons for the **B** and

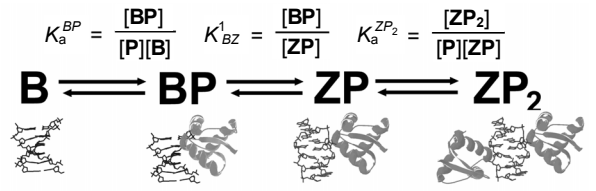

Figure 1. *Active-mono* B-Z transition mechanism of a 6-bp DNA duplex by two Z-DNA binding proteins. **B** and **Z** indicate the B-form and Z-form of the DNA duplex and **P** indicates the Z-DNA binding proteins.

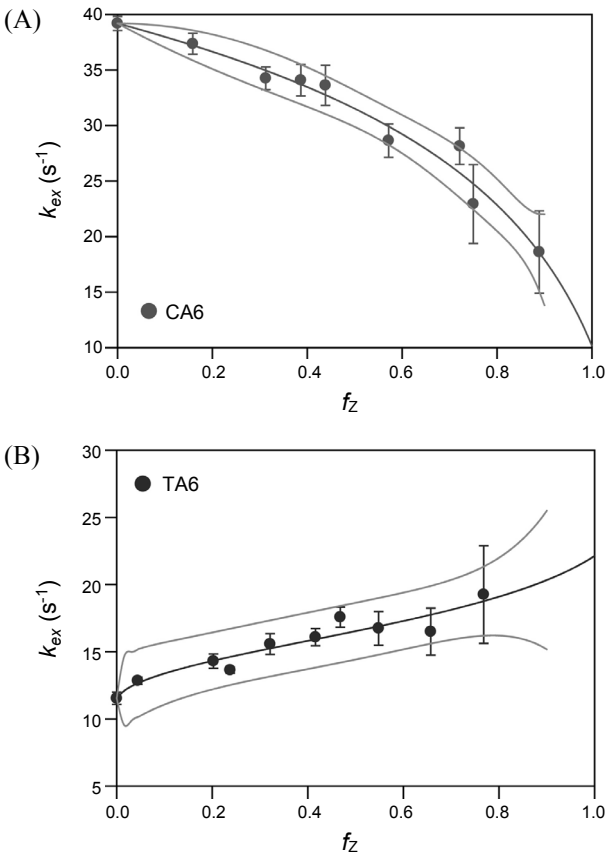

Figure 2. (A) The k_{ex} values of the G4b imino proton for the CA6-$Z\alpha_{ADAR1}$ complex determined at 25 °C and (B) k_{ex} values of the G2b imino proton for the TA6-$Z\alpha_{ADAR1}$ complex determined at 15 °C as a function of the f_Z. Black solid lines are the best fit to Eq. 1, where the k_{ex} data were weighted by the inverse of their variance. The grey lines indicate their upper and lower confidence limits (95% confidence level).

BP states, respectively, and [**B**] and [**BP**] are the concentrations of the **B** and **BP** states, Z_t is the total concentration of Z-conformation. Thus, the correlation between the k_{ex} and f_Z data can be expressed by Eq. 2 as described in previous report:[10]

$$k_{ex} = k_{ex}^{B} + \frac{\left(k_{ex}^{BP} - k_{ex}^{B}\right)}{2(1-\alpha)(1-f_Z)}\left\{1 + (K_{BZ}^{1} - 1)f_Z - \sqrt{(1 + (K_{BZ}^{1} - 1)f_Z)^2 - 4K_{BZ}^{1}(1-\alpha)f_Z(1-f_Z)}\right\} \quad (2)$$

where K_{BZ}^{1} = [**BP**]/[**ZP**], and α (= $K_a^{ZP_2}/K_a^{BP}$) is the ratio of the association constants (K_a) of the **ZP$_2$** and **BP** complex states. In the previous report,[10] the α (CA6: 1.42; TA6 13.9), K_{BZ}^{1} (CA6: 0.4 ± 0.1; TA6: 6.3 ± 3.1), k_{ex}^{B} (CA6: 39.2 ± 0.6 s^{-1}; TA6: 11.5 ± 0.5 s^{-1}), and k_{ex}^{BP} (CA6: 10.2 ± 3.1 s^{-1}; TA6: 22.2 ± 5.3 s^{-1}) values of CA6 and TA6 complexed with Zα$_{ADAR1}$ were determined by curve fitting k_{ex} of the imino protons as a function of f_Z with Eq. 2 (Fig. 2).[10]

In order to estimate the reliability of the proposed model in the previous study, we performed the iterative non-linear curve fitting k_{ex} of the imino protons in the CA6-Zα$_{ADAR1}$ and TA6-Zα$_{ADAR1}$ complexes as a function of f_Z with Eq. 2 using program Origin 7. The upper and lower confidence limits on the k_{ex} data of CA6 and TA6 complexed with Zα$_{ADAR1}$ were evaluated by iterative non-linear curve fitting and the 95% confidence bands of the k_{ex} data are shown in Fig. 2. This result shows that the *active-mono* B-Z transition mechanism, which was proposed in the previous study,[10] is suitable approach to understand the DNA sequence descrimination step of the Zα$_{ADAR1}$ protein during B-Z transition.

The relative population of each complex state (such as **B**, **BP**, **ZP**, and **ZP$_2$**) as a function of the P/N ratio was determined from the f_Z and k_{ex} data, which were reported in previous study,[10] as the following procedure. First, the [**BP**] values are calculated from the exchange data, k_{ex}, k_{ex}^{B}, and k_{ex}^{BP} ,by using Eq. 3:

$$[BP] = \frac{k_{ex} - k_{ex}^{B}}{k_{ex}^{BP} - k_{ex}^{B}}(1 - Z_t) \quad (3)$$

where Z_t are determined from relative peak intensities of the imino proton resonances of the Z-form DNA. Second, the [**B**] values can be calculated by using the equation, [**B**] = 1 − Z_t − [**BP**]. Third, the concentration of the **ZP** state ([**ZP**]) is calculated from the flowing relation, [**ZP**] = [**BP**]/K_{BZ}^{1}. Forth, the concentration of the **ZP$_2$** state ([**ZP$_2$**]) can be calculated by using the equation, [**ZP$_2$**] = Z_t − [**ZP**]. The relative populations (including estimated errors) of the **B**, **BP**, **ZP**, and **ZP$_2$** states in the CA6-Zα$_{ADAR1}$ and TA6-Zα$_{ADAR1}$ complexes as a function of the P/N ratio are shown in Fig. 3 and 4, respectively. Finally, the concentration of free Zα$_{ADAR1}$ ([**P**]) could be calculated by the Eq. 4:

$$[P] = P_t - [BP] - [ZP] - 2[ZP_2] \quad (4)$$

where P_t is the total concentration of Zα$_{ADAR1}$.

From these concentrations, the association constants, K_a^{BP} = [**BP**]/[**B**][**P**] and $K_a^{ZP_2}$ = [**ZP$_2$**]/[**ZP**][**P**], for the CA6-Zα$_{ADAR1}$ and TA6-Zα$_{ADAR1}$ complexes were calculated. The K_a^{BP} and $K_a^{ZP_2}$ values of CA6-Zα$_{ADAR1}$ complex are 3.9 ± 1.3 × 10^3 and 5.5 ± 1.9 × 10^3, respectively.[10] This means that, unlike the d(CGCGCG)$_2$-Zα$_{ADAR1}$ complex,[9] the Zα$_{ADAR1}$ protein can bind to the B and ZP complex states with similar binding affnity. The relative population of each complex state for the CA6-Zα$_{ADAR1}$ complex as a function of the P/N ratio could be calculated from these association constants and equilibirum constants for B-Z transition and the results are shown in Fig. 3 (solid lines). It was observed that [**B**] was gradually decreased, but [**BP**] and [**ZP**] were increased as the P/N ratio increased up to 2 (Fig. 3). In addition, the observation that [**BP**] is always smaller than [**ZP**] could be explained by the fact that K_{BZ}^{1} < 1 (Fig. 3). When the P/N ratio rose to 2, the **ZP$_2$** complex was dominantly produced but [**BP**] and [**ZP**] were decreased as the P/N ratio increased because the added **P** preferentially bound to the **ZP** complex rather than the **B** and **BP** (Fig. 3).

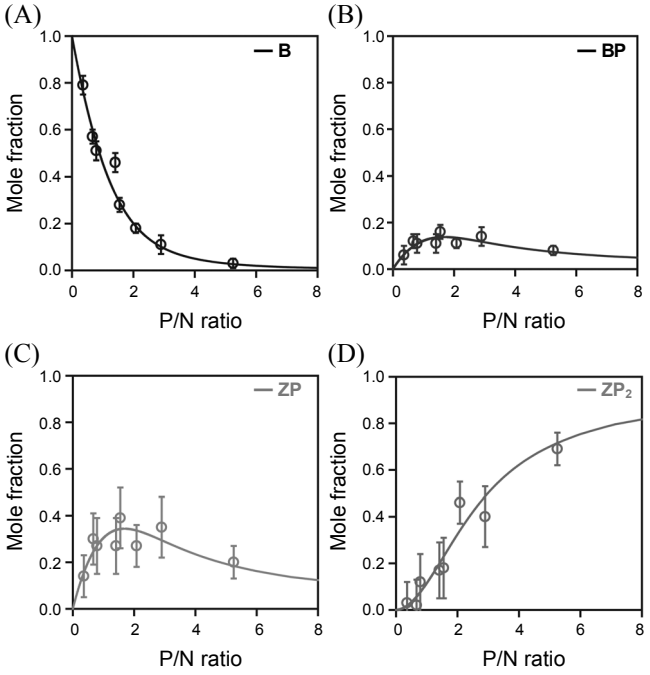

Figure 3. The relative populations of the (A) **B**, (B) **BP**, (C) **ZP**, and (D) **ZP$_2$** states within total DNA populations of the CA6 complexed with hZα$_{ADAR1}$ determined at 25°C. Solid lines are simulated relative population of each complex state determined as described in text.

Similarly, the K_a^{BP} and $K_a^{ZP_2}$ values of the TA6-Zα$_{ADAR1}$ complex are 2.5 ± 0.9 × 10^3 and 3.5 ± 1.3 × 10^4, respectively.[10] The relative population of each complex state for the TA6-Zα$_{ADAR1}$ complex as a function of the P/N ratio are shown in Fig. 4 (solid lines). Similar to the CA6-Zα$_{ADAR1}$ complex In the both complexes, it was observed that [**B**] was gradually decreased, but [**BP**] and [**ZP**] were increased as the P/N ratio increased up to 2 (Fig. 4). However, contrast to the CA6-Zα$_{ADAR1}$ complex, it was observed that [**BP**] is always larger than [**ZP**], indicating that K_{BZ}^{1} > 1, (Fig. 4). When the P/N ratio rose to 2, the **ZP2** complex was dominantly produced but [**BP**] and [**ZP**] were decreased as the P/N ratio increased like CA6 (Fig. 4).

Interestingly, the simulated population (solid line in Fig. 3 and

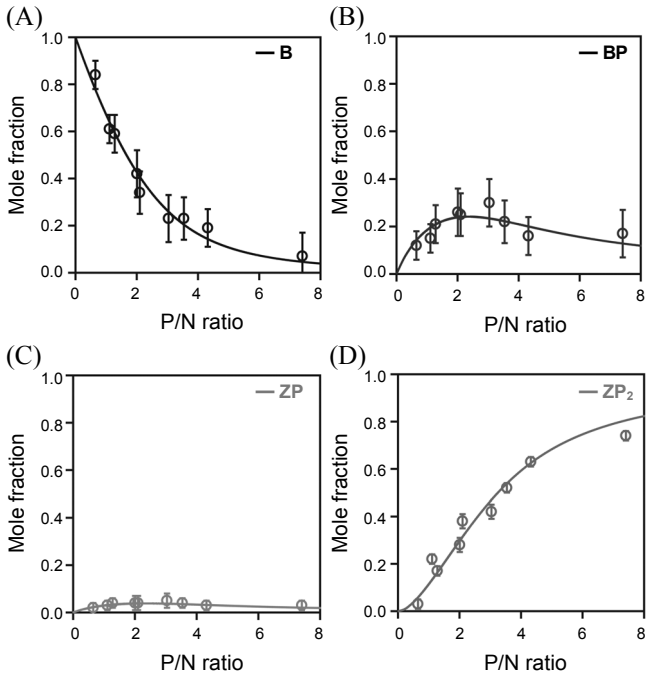

Figure 4. The relative populations of the (A) **B**, (B) **BP**, (C) **ZP**, and (D) **ZP$_2$** states within total DNA populations of the TA6 complexed with hZα$_{ADAR1}$ determined at 15°C. Solid lines are simulated relative population of each complex state determined as described in text.

4) of each complex data determined from the association constants and equilibirum constants for B-Z transition well matched to the experimental value (symbol in Fig. 3 and 4) determined from the f_Z and k_{ex} data. This indicates that our approach is able to calculate successfully the concentrations of the intemediate state during B-Z transition. This correlation between the relative population of each complex state and the P/N ratio as shown Fig. 3 and 4 can explain how the Zα$_{ADAR1}$ protein recognizes the d(CGCGCG) sequence from d(CACGTG) and d(CGTACG) sequences in a long genomic DNA.

In summary, we derived the relative population of each complex state, which is thought to be produced during B-Z transition induced by Zα$_{ADAR1}$, as a function of the P/N ratio. This approach provides the insight into the *active* B-Z transition mechanism and DNA sequence discrimination step of human Z-DNA binding protein, ADAR1.

Acknowledgments. This Work was supported by the National Research Foundation of Korea Grants [2010-0014199; NRF-C1ABA001-2010-0020480 to J.-H.L.] funded by the Korean Government (MEST).

References

1. Rich, A.; Nordheim, A.; Wang, A. H. *Annu. Rev. Biochem.* **1984**, *53*, 791.
2. Herbert, A.; Rich, A. *J. Biol. Chem.* **1996**, *271*, 11595.
3. Herbert, A.; Rich, A. *Genetica* **1999**, *106*, 37.
4. Herbert, A. G.; Rich, A. *Nucleic Acids Res.* **1993**, *21*, 2669.
5. Herbert, A.; Alfken, J.; Kim, Y. G.; Mian, I. S.; Nishikura, K.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8421.
6. Herbert, A.; Schade, M.; Lowenhaupt, K.; Alfken, J.; Schwartz, T.; Shlyakhtenko, L. S.; Lyubchenko, Y. L.; Rich, A. *Nucleic Acids Res.* **1998**, *26*, 3486.
7. Schwartz, T.; Rould, M. A.; Lowenhaupt, K.; Herbert, A.; Rich, A. *Science* **1999**, *284*, 1841.
8. Ha, S. C.; Choi, J.; Hwang, H. Y.; Rich, A.; Kim, Y. G.; Kim, K. K. *Nucleic Acids Res.* **2009**, *37*, 629.
9. Kang, Y.-M.; Bang, J.; Lee, E.-H.; Ahn, H.-C.; Seo, Y.-J.; Kim, K. K.; Kim, Y.-G.; Choi, B.-S.; Lee, J.-H. *J. Am. Chem. Soc.* **2009**, *131*, 11485.
10. Seo, Y.-J.; Ahn, H.-C.; Lee, E.-H.; Bang, J.; Kang, Y.-M.; Kim, H.-E.; Lee, Y.-M.; Kim, K.; Choi, B.-S.; Lee, J.-H. *FEBS Lett.* **2010**, *584*, 4344.