Evidence for an Intermediate State in the B-to-Z Transition of DNA

Vassili Ivanov,* Kazimierz Grzeskowiak, and Giovanni Zocchi

Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095-1574

Received: June 5, 2003; In Final Form: September 16, 2003

We report here the study of the transition of a B-form of d(CGCGCGCGCGCG) to the Z-form as a function of the ions strength of PBS buffer. We used circular dichroism (CD) spectroscopy. Decomposition CD spectra by linear superposition provides evidence for one or more intermediate states that are distinct from the B- and Z-forms of DNA. Linear decomposition is performed according to the locally linearized model (LLM) modified by the explicit introduction of the weight function essentially present in any quadratic integral of the spectra.

Introduction

Changing from familiar right-handed B-DNA to elongated left-handed Z-DNA occurs most readily in segments with specialized sequences, favored largely by alternating deoxycytidines and deoxyguanosines.[1] The crystal structure of a DNA hairpin, formed from partially self-complementary d(CGCGCGTTTTCGCGCG) hexadecamer indicated the duplex stem region, is in the left-handed Z-conformation.[2] Circular dichroism (CD) spectra of this 16-mer hairpin demonstrated that the transition from B- to Z-form of the duplex stem is a two-state process, and the transition depends on the sequence of the loop.[3] Such transition usually requires an extreme salt concentration (up to 4 M of Na or other monovalent cations), or an increase in temperature, the presence of ethanol or cations such as cobalt hexamine and polyamines.[4] The helix elongation from 3.4 Å per base step (in B-DNA) to 3.7 Å (in Z-DNA) combines with an inversion of guanine from anti to syn configuration. Although even in the most favorable sequence, ...CGCGCGCGCGCG..., under normal conditions the Z form is much less stable than the B form, a biological role for Z-DNA is quite possible because negative supercoiling favors its formation. Indeed there are proteins that fix DNA sites with Z-forming sequences.[5] To further understand the mechanism of potential transitions occurring between both types of DNA secondary structures, we investigate the CD spectra of d(CGCGCGCGCGCG) upon increasing the ionic strength of phosphate buffered saline (PBS). We corrected the locally linearized model (LLM) by the explicit introduction of the weight function essentially present in any quadratic integral of the spectra. Using corrected LLM, we find the evidence for intermediate state(s) during the transition of a B- to Z-form.

Materials and Methods

Reagents were purchased from Sigma and d(CGCGCGCGCGCG) was from Operon Inc. The salt concentration of DNA solution was adjusted by adding the appropriate amount of solid PBS (Sigma P-4417) and was measured in total concentration of sodium and potassium. CD spectra of d(CGCGCGCGCGCG), 4.5 uM/mL equal to 0.7 O.D were measured at 0.1, 1.0, 1.5, and 2.0 M and in increments of 0.2 M thereafter up to 3.4 M, and 4.0 M PBS in a 1 cm path length quartz cell using a spectropolarimeter (Jasco J600) with temperature stabilized at 20 °C. Ten scans at 50 nm/minute were made over the wavelength range from 210 to 325 nm. The spectra in a range from 215 to 325 nm were used for analysis.

Linearity of the CD Spectra and Linear Functional Space. The locally linearized model (LLM)[6] is usually used for the determination of the protein secondary structure (percentage of the α-helixes, β-sheets, and β-turns within the same protein). According to the LLM, the CD spectrum of a mixture of two diluted solutions is the linear combination of the spectra of the components with weights equal to their concentrations. If α is a concentration of the first component having secondary structure of the first type with normalized CD spectrum $f(\lambda)$ and β is the concentration of the second component with normalized CD spectrum $g(\lambda)$, then the CD spectrum of the mixture of two components is $y(\lambda) = \alpha f(\lambda) + \beta g(\lambda)$, where concentrations α and β are indeed the *coordinates* of the function $y(\lambda)$ in the basis of $f(\lambda)$ and $g(\lambda)$. The scalar product on the linear functional space (LFS) of CD spectra is the quadratic integral of the spectra over wavelength λ with positive weight function $w(\lambda)$:

$$(f,g) \equiv \int_{\lambda_1}^{\lambda_2} w(\lambda) f(\lambda) g(\lambda)\, d\lambda \quad f^2 \equiv (f,f) \tag{1}$$

where the spectra $f(\lambda)$ and $g(\lambda)$ are measured from λ_2 to λ_1, and $f^2 > 0$ for $\forall f \neq 0$.

If $f(\lambda)$ and $g(\lambda)$ are linearly independent functions that are used as a basis of the LFS, the CD spectrum y of the mixture of two components can be measured and the functional equation

$$h^2(\alpha,\beta) \equiv (y - \alpha f - \beta g)^2 \tag{2}$$

can be minimized to find concentrations α and β. We call the function

$$h(\lambda) \equiv y(\lambda) - \alpha f(\lambda) - \beta g(\lambda) \tag{3}$$

the *perpendicular* to the basis $f(\lambda)$, $g(\lambda)$ part of the spectrum $y(\lambda)$, since $(h,f) = (h,g) = 0$, and the function

$$y(\lambda) - h(\lambda) = \alpha f(\lambda) + \beta g(\lambda) \tag{4}$$

the *parallel* to the basis $f(\lambda)$, $g(\lambda)$ part of the spectrum $y(\lambda)$. If the value of $y(\lambda)$ is only a mixture of $f(\lambda)$, $g(\lambda)$, then $y(\lambda) =$

* Corresponding author. Phone (310) 206-0715. E-mail vivanov@ucla.edu

10.1021/jp035593p CCC: $25.00
Published on Web 10/23/2003

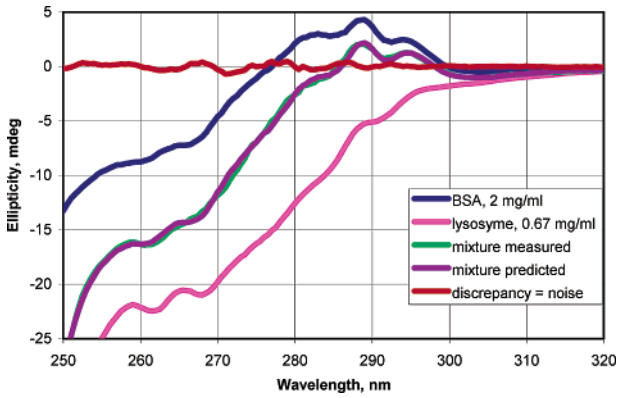

Figure 1. CD spectrum of BSA (dark blue curve) and lysozyme (pink curve) mixture (green curve) decomposed in the basis of the components. Concentrations of both proteins in the mixtures are derived.

$\alpha f(\lambda) + \beta g(\lambda)$, $h^2(\alpha,\beta) = 0$, and correspondingly $h(\lambda) = 0$. By this way, the CD spectrum of the mixture can be decomposed into a linear combination of the CD spectra of components.

The weight $w(\lambda)$ is an arbitrarily positive function. We assume that the weight function is essential for the LLM. In fact, the spectra can be either integrated over the wavelength $d\lambda$, $w(\lambda) = 1$, or over energy $dE \propto d\lambda/\lambda^2$, $w(\lambda) = 1/\lambda^2$, or weight function can be chosen to suit some specific properties of the spectra. The decomposition of the CD spectrum of the mixture is unique and does not depend on the weight function choice, if the basis is complete, and there is no noise in measurements. If the basis is not complete, i.e., some unknown components are present in the solution, the solution spectrum cannot be decomposed completely into the basis spectra, the function $h(\lambda)$ does not equal to zero, and the perpendicular part of the spectrum $h(\lambda)$ and the coordinates in the basis depend on the weight function choice. The perpendicular part of the spectrum in this case is not a spectrum of the unknown components (means not the basis components), but it is the part of the spectrum of unknown components perpendicular to the basis within the current choice of the weight function. The concentrations of the basis components cannot be calculated using solely LLM in this case.

To test the method, the bovine serum albumin (BSA) 2.0 mg/mL and lysozyme 1.0 mg/mL were prepared and mixed in the ratio 1:2. The spectrum was integrated from 250 to 340 nm with weight $w(\lambda) = 1$. Figure 1 shows the CD spectrum of the mixture (green curve) decomposed in the basis of the components (BSA, 2.0 mg/mL, dark blue curve; lysozyme, 0.67 mg/mL, pink curve). The mixture spectrum was decomposed into parts parallel (brown curve) and perpendicular (red curve) to the basis. The parallel part coincides with measured spectrum of the mixture (brow and green curves coincide), while the perpendicular part of the spectrum (red curve) is equal to zero up to the noise. The concentrations of the proteins calculated from the CD spectrum were in agreement with the prepared concentrations of stock solution of the mixture of two proteins.

Procedure. The evidence of the perpendicular state can be obtained from decomposition of the spectra into the parallel and perpendicular parts. If, in addition to B- and Z-forms, the other component is present, then the basis would be incomplete and $h(\lambda) = y(\lambda) - \alpha f(\lambda) - \beta g(\lambda) \neq 0$. The perpendicular and parallel parts of the spectrum $y(\lambda)$ depend on the choice of weight function $w(\lambda)$, which can be used to suppress the high noise part of the spectrum.

The characteristic CD spectra of B- to Z-DNA transition lie in the 185 to 320 nm range. The spectra have large peaks near 200 nm. Due to dramatic increase of optical density below 210 nm, this part of the spectrum also has a higher noise and should be measured separately in more diluted conditions. The peaks near 200 nm have 10 times larger amplitude than the part of the spectrum from 215 to 320 nm that is analyzed in this paper. Therefore, these peaks will contribute 100 times more than the rest of the spectrum in the quadratic scalar products, and the part of the spectra from 215 to 320 nm might be suppressed into 1% noise. We can avoid this problem by an appropriate choice of the weight function, or simply by restricting limits of integration. Thus, we were able to detect the presence of an intermediate state.

The spectrum was measured from 210 to 320 nm; weight function $w(\lambda)$ was equal to 1 from 215 to 320 nm, and equal to zero below 215 nm and above 320 nm. By this choice we avoid the high noise part of the spectra below 215 nm. The large peak near 200 nm is also not included into our calculation, and within an order of magnitude all measured spectra have the same noise and intensity.

The decomposition of the intermediate ionic strength solution spectra is accomplished in the basis of B-DNA (0.1 M PBS) and Z-DNA (4 M PBS) CD spectra. The basis function $f(\lambda)$ is B-DNA spectrum, α is the coordinate of $f(\lambda)$ or, in other words, the normalized concentration of B-DNA. Correspondingly $g(\lambda)$ and β are the basis function and the coordinate, respectively, for Z-DNA. The coordinates for B- and Z-DNA are normalized to $\alpha = 1$, $\beta = 0$ at 0.1 M PBS, and $\alpha = 0$, $\beta = 1$ at 4.0 M PBS correspondingly. The measured spectrum at the intermediate ionic strength is $y(\lambda)$. Plots of the perpendicular $h(\lambda) = y(\lambda) - \alpha f(\lambda) - \beta g(\lambda)$ and parallel $\alpha f(\lambda) + \beta g(\lambda)$ spectra versus salt concentration were constructed minimizing the functional $h^2(\alpha,\beta)$, see eq 2.

For the two-state transition from B- to Z-DNA through the mixture of B- and Z-forms, the sum of the concentrations of B- and Z-DNA in the solution must be the constant value of 1 according to the chosen normalization

$$\alpha + \beta = 1 \tag{5}$$

In this case the spectrum $y(\lambda)$ measured at the intermediate ionic strength must be completely decomposed into the B- and Z-DNA spectra $y(\lambda) = \alpha f(\lambda) + \beta g(\lambda)$, and the perpendicular part of the spectrum $h(\lambda) = y(\lambda) - \alpha f(\lambda) - \beta g(\lambda)$ is equal to zero.

It is to be mentioned that the LLM technique works best for measurements of the concentrations in the solution of non-interacting (nonreacting and nonaggregating) molecules with linear independent CD spectra.[7] In our case, the primary structure of DNA is well defined during the measurements; however, it is also possible that two DNA helices may associate into semidimeric duplexes during the B- to Z-DNA transition.

Results and Discussion

The CD spectra d(CGCGCGCGCGCG), (Figure 2) shows the change of long-wavelength from a positive−negative pattern at low salt concentration 0.1 M PBS (B-form, dark blue curve) to nearly the mirror image negative−positive at high salt 4.0 M PBS (Z-form brown curve). Between 0.1 M PBS and 2 M PBS the spectrum does not change dramatically, so we focus on the 2.2 to 3.4 M range. The CD spectra at intermediate ionic strengths from 2.2 M PBS (pink curve) to 3.4 M PBS (blue curve) measured with 0.2 M PBS ionic strength increments. The ellipticity probes in the range 270 to 300 nm and 250 to 270 nm are related, respectively, to the direction of the transition vector of guanine base stacked between two cytosine vectors (8), and to the unwinding and expansion of double stranded

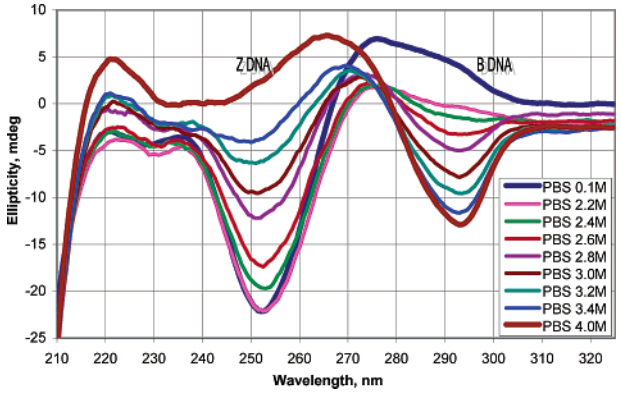

Figure 2. CD spectra of d(CGCGCGCGCGCG) measured at salt concentration varying from 0.1 M (dark blue curve) to 4.0 M PBS (brown curve). In the range 270 to 277 nm all the CD graphs from 2.2 (pink curve) to 3.4 M (blue curve) PBS are below both B and Z spectra (dark blue and brown curves), which is completely impossible for a mixture of only B and Z forms.

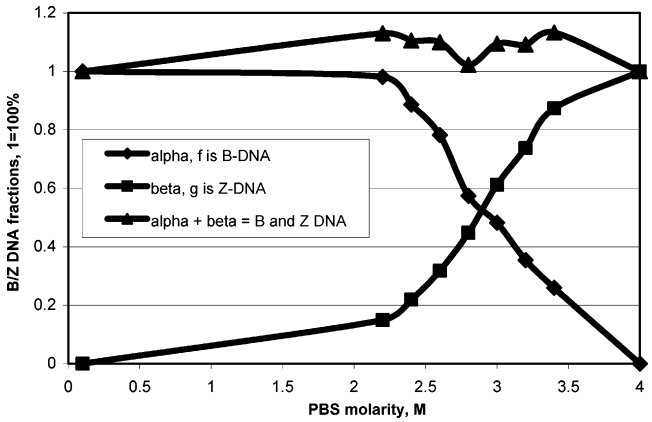

Figure 3. Graph of a B- and Z-form of d(CGCGCGCGCGCG) concentrations (dark blue and brown curves) measured at different buffer concentrations; (I) 0.1 M PBS only B-DNA is present and $\alpha = 1$, $\beta = 0$. (II) 4.0 M PBS only Z-DNA is present and $\alpha = 0$, $\beta = 1$. (III). For mixture of only B- and Z-DNA present the sum of $\alpha + \beta$ (red curve) is equal to 1. In the range from 2.2 to 3.4 M PBS except 2.8 M PBS a sum of $\alpha + \beta = 1.13$ indicates on existence of two intermediate states.

helix (9). As the ionic strength of the buffer increases from 0.1 to 4.0 M PBS, positive ellipticity of the band at 275 nm decreases until finally it is inverted completely to a negative value, shifted to 293 nm in 4.0 M PBS−characteristic features of Z-DNA. Similarly, the long-wavelength negative CD at 253 nm changes accordingly as the ionic strength of PBS buffer increases, and in 4.0 M PBS it reaches a positive value expanding from 250 to 275 nm. This indicates that at 0.1 M PBS the B-state content α is 1 and monotonically decreases to reach a value of zero at 4.0 M, whereas the Z-form content β continues to rise from zero to 1 with increasing buffer concentration. In the range 270 to 277 nm, all the CD graphs from 2.2 to 3.4 M PBS are below both B and Z spectra, which is completely impossible for a mixture of B- and Z-DNA because the spectrum of the mixture is a linear combination of two forms, B and Z, with positive coefficients α and β, which sum equals to 1 (see eq 5). It suggests that the transition from B-DNA to the Z form may proceed through an intermediate state.

Figure 3 shows calculated coordinates of B-form α (dark blue curve) and Z-form β (brown curve), and the sum of them (red curve) calculated for different ionic strengths by minimizing the functional $h^2(\alpha,\beta)$ (equation 2). The sum $\alpha + \beta$ calculated

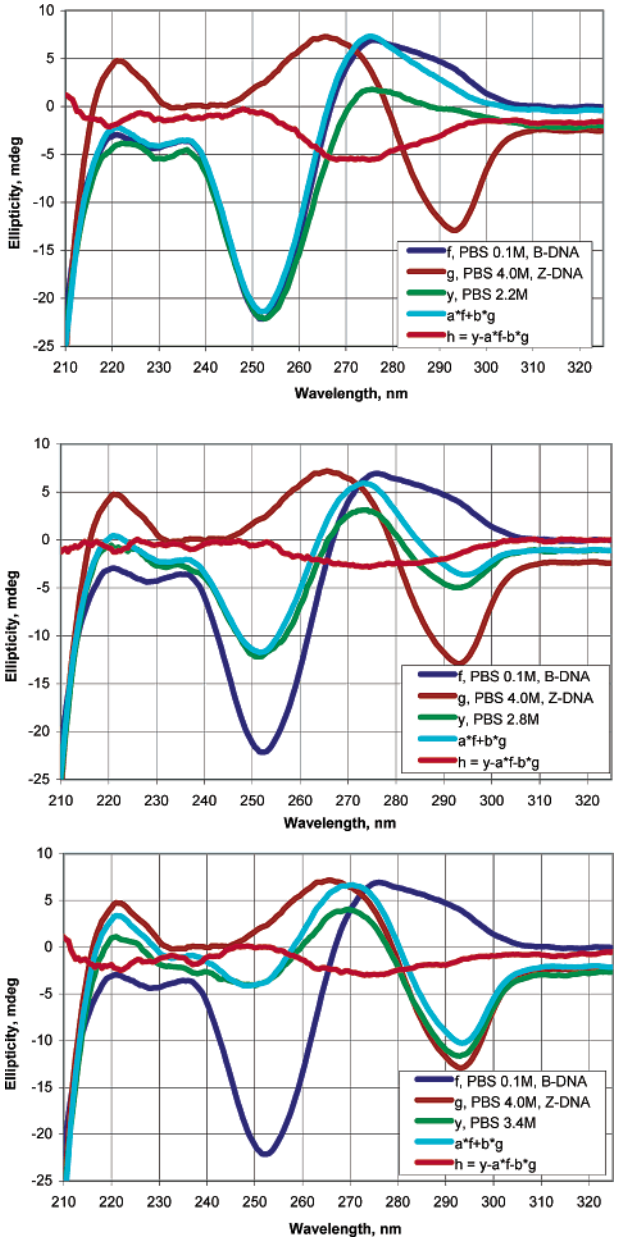

Figure 4. (a, b, c). Graphs of CD spectra of d(CGCGCGCGCGCG) (green curve) decomposed in the basis of B- and Z-DNA (dark blue and brown curve) into the parallel (light blue curve) and perpendicular (red curve) to the basis functions calculated at concentrations of (a) 2.2 M, (b) 2.8 M, and (c) 3.4 M PBS. Presence of nonzero perpendicular part (red curve), which cannot be decomposed as a linear combination of B-and Z-DNA spectra, is the evidence of intermediate state.

for all concentrations from 2.2 M to 3.4 M PBS reaches up to 1.13 or 13% above the value of 1, with the exception 2.8 M PBS. This is another evidence that the B−Z transition is not a two state process through B and Z forms only.

Figure 4a,b,c shows the decomposition of CD spectra $y(\lambda)$ (green curve) measured for 2.2, 2.8, and 3.4 M PBS. The best-fit linear combination of B- and Z-DNA spectra $\alpha f(\lambda) + \beta g(\lambda)$ (light blue curve) is significantly different from the decomposed CD spectra (green curve). Really, the perpendicular part of the measured spectra $h(\lambda) = y(\lambda) - \alpha f(\lambda) - \beta g(\lambda)$ (red curve) is the discrepancy between measured spectrum $y(\lambda)$ and the best-fit spectrum $\alpha f(\lambda) + \beta g(\lambda)$. The graphs obtained in the range 210 to 325 nm show that the discrepancy $h(\lambda)$ is far from zero. This is another evidence of the presence of an intermediate state during a B-to-Z transition. Figures 4a and 4c, CD spectra for

2.2 M PBS and 3.4 M PBS, show that the perpendicular part of the spectra is present over the entire measured range, with exception of the area near 250 nm, while for 2.8 M PBS, Figure 4b, the perpendicular part of the spectrum is localized from 255 to 300 nm only.

Conclusion

Using a linear decomposition of the CD spectra into B-DNA and Z-DNA components, we show the existence of a third component orthogonal to both the B- and Z-DNA spectra. This calculated orthogonal component of the spectra is nonzero in the ionic strength range from 2.2 to 3.4 M. This suggests the presence of an intermediate state between the B- and Z-DNA forms. In comparison to previous studies[3] the critical element, which allows us to detect the presence of an intermediate state, is that the analysis was performed on a restricted range in wavelengths, from 215 to 320 nm. This excludes a part of the spectrum which otherwise contributes disproportionately highly to the quadratic integrals of LLM, suppressing other contributions. In the chosen range of wavelength, all the spectra have similar values of CD signal for all ionic strengths. The optical density does not change significantly, and all factors contribute to a good signal-to-noise ratio.

The presence of the intermediate state raises the question of the mechanism of the structural rearrangement of the DNA bases and the sugar conformation during the transition. Structures from X-ray crystal analysis indicate that in B-DNA the phosphate conformation is related to base stacking and overlap, and this has sequence dependence. For example, the BII phosphates are observed only at steps of the type pyrimidine-purine, and C−G steps are most conductive to roll bending.[9] In a Z-form of d(CGCGCG), the configuration (ZI or ZII) of phosphate groups is controlled by the occupancy of cations around these phosphates.[10] Considering the results of our measurements and the X-ray crystal data of DNA oligomers, a picture emerges of a quasi-continuous base deformation in either direction beginning at the 5′-phosphates surrounded by cations followed by guanine bases. This would lead to the formation of an intermediate state, which gradually converts into the B- or Z-form in response to salt concentration changes.

Acknowledgment. This research was partly supported by the U.S.-Israel Binational Science Foundation under grant no. 2000298 and National Science Foundation grant DMR-0105903.

References and Notes

(1) Haniford, D. B.; Polleybank, D. E. *Nature* **1983**, *302*, 632−34.
(2) Chattopadhyaya, R.; Grzeskowiak, K.; Dickerson, R. E. *J. Mol. Biol.* **1990**, *211*, 189−210.
(3) Amaraturnga, M.; Pancoska, P.; Paner, T. M.; Benight, A. S. *Nucleic Acids Res.* **1990**, *18*, 577−582.
(4) Gosule, L. C.; Schellman, J. A. *J. Mol. Biol.* **1978**, *121*, 311−326.
(5) Schwartz, T.; Rould, M. A.; Lowenhaupt, K.; Herbert, A.; Rich, A. *Science* **1999**, *284*, 1841−1845.
(6) van Stokkum, I. H. M.; Spoelder, H. J. W.; Bloemendal, M.; Grondelle, R.; van Groen, F. C. A. *Anal. Biochem.* **1990**, *191*, 110−118.
(7) Venyaminow, S.; Yu.; Yang, J. T. *Determination of protein Secondary Structure, 69−107, from Circular Dichroism and the Conformational analysis of Biomolecules*, Fasman, G. D., Ed.; Plenum Press: New York, 1996.
(8) Harada, N.; Nakanishi, K. *Circular Dichroic Spectroscopy: Exciton Coupling in Organic Stereochemistry*; University Sciences Books, Mill Valley, CA, 1983.
(9) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67−111.
(10) Ohishi, H.; Nakanishi, I.; Tomita, K. *Biochem. Biophys. Res. Commun.* **1997**, *236*, 146−150.