Transworld Research Network
37/661 (2), Fort P.O., Trivandrum-695 023, Kerala, India

DNA Structure, Chromatin and Gene Expression, 2006: 17-26 ISBN: 81-7895-228-9
Editors: Ryoiti Kiyama and Mitsuhiro Shimizu

2 Biochemical and microscopic analyses of DNA conformations

Mikio Kato
Department of Biological Science, Osaka Prefecture University Graduate School of Science, 1-1 Gakuencho, Naka-ku, Sakai 599-8531, Japan

Abstract

In addition to the well-known right-handed double helix (B-DNA), a number of alternative DNA structures can be formed at defined, ordered DNA sequences. For example, inverted repeats may form DNA cruciforms, pyrimidine/purine-biased sequences may form intra- and intermolecular DNA triplexes, alternating purine-pyrimidine sequences may form Z-DNA, and so on. These characteristic sequence tracts occur often within the genomes. And to understand their function, it is necessary to understand their potential to form alternative DNA structures. In this paper, we review several methods for analyzing the

Correspondence/Reprint request: Dr. Mikio Kato, Department of Biological Science, Osaka Prefecture University Graduate School of Science, 1-1 Gakuencho, Naka-ku, Sakai 599-8531, Japan
E-mail: mkato@b.s.osakafu-u.ac.jp

structures of DNA and chromatin, and discuss some of the biological implications of the observed DNA structures.

Introduction

In response to changes in environmental conditions, the structure of DNA may vary in a sequence-dependent manner. Any DNA conformation other than the right-handed double helix (B-DNA) is classified as an "unusual DNA structure," and a variety of DNA sequences with the potential to form unusual structures have been characterized [1-3]. For example, pyrimidine/purine-biased sequences (hereafter, Pyr/Pur sequences) may form inter- and intramolecular triplexes and their derivatives; inverted repeats may form DNA cruciforms; alternating purine-pyrimidine sequences may form Z-DNA (left-handed double helix); the occurrence of phased oligoadenine tracts causes curvature of the DNA axis; direct repeat sequences may cause slipped misalignment; and telomeric guanine-rich repeats may form quadruplexes. Some of these sequences are widespread within genomes, occurring more often than random sequences [4,5], and the unusual (alternative) DNA structures that result are thought (or expected) to affect processes by which gene expression is regulated by modulating global and local DNA topology or by serving as target sites for specific proteins [6-10]. In addition, some unusual DNA structures that appear to have no effect on genome function nevertheless may have a big impact on gene structure and expression. For instance, in the case of triplet repeats and some other simple repeat sequences, the numbers of tandem repeats are unstable during transmission [11]. Although the instability is, itself, not known to affect genome function, expansion of the repeat regions may cause congenital anomalies. Models for expansion and contraction during DNA replication have been proposed based on unusual DNA structures that form at the template or nascent strands, and the mechanisms underlying the diseases caused by expansion of repeat tracts have been discussed [12,13].

In order to clarify how B-DNA transitions into unusual DNA structures, in particular how it occurs *in vivo*, one needs to establish methods for characterizing DNA structures under defined sets of conditions. In this paper, the current methodologies of topological fractionation using gel electrophoresis, chemical and enzymatic probing, and microscopic imaging of DNA and chromatin are reviewed with an emphasis on alternative DNA structures forming at inverted repeats and Pyr/Pur sequences.

Electrophoretic mobility of supercoiled DNA

DNA's electrophoretic mobility is strongly dependent on both the size and shape of the molecule. Larger molecules migrate slower in agarose gel, while

more compact molecules (*e.g.*, supercoiled DNA) migrate faster. Because the formation of intramolecular triplexes results in relaxation of negative supercoils (as does cruciform formation), two-dimensional agarose gel electrophoresis enables one to identify topoisomers at which structural transitions have occurred.

E. coli plasmid DNA molecules are mixtures of topoisomers, and each topoisomer yields an individual spot on two-dimensional gel electrophoresis [14]. In the first dimension, a set of topoisomers is fractionated under conditions in which structural transition may occur in some; in the second dimension they are fractionated under conditions in which no structural transitions occur. Because intercalation of chloroquine reduces the twist angle between neighboring base pairs, electrophoretic buffer containing chloroquine is used in the second dimension to prevent supercoiling-dependent structural transition. We have used this approach to detect formation of unusual DNA structures occurring on supercoiled DNA containing short inverted repeats [15]. The plasmid pBan1 is a pUC19 derivative containing a short inverted repeat made up of 17-bp complementary sequences. The electrophoretic profile of pBan1 DNA exhibiting native bacterial superhelicity is shown in Fig. 1. A structural transition that consumes about 3.5 superhelical turns occurs in

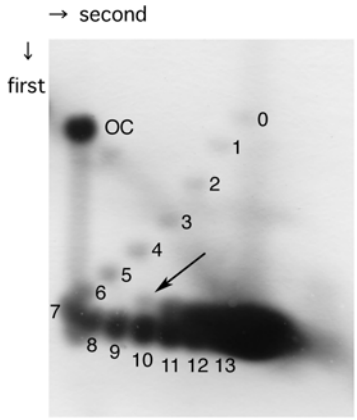

Figure 1. Two-dimensional agarose gel electrophoresis of pBan1 DNA. Plasmid pBan1 exhibiting native bacterial superhelicity was two-dimensionally fractionated on 1.2% agarose gel. In the first dimension, standard TAE buffer was used; 6 μg/ml chloroquine was added in the second dimension. OC represents nicked circular pBan1, and the numbers of negative supercoils present in the samples are given at the respective spots. Structural transition has occurred in the DNA molecules having 10 or more negative supercoils (indicated by arrow), consuming ~3.5 helical turns. Reproduction from Kato *et al*. [15] with permission of the Biophysical Society.

a small population of pBan1 DNA molecules having 10 superhelical turns (indicated by the arrow in the Fig. 1), and in up to 50% of pBan1 DNA molecules having more than 10 superhelical turns. Thus, the negative superhelicity required for structural transition and the numbers of superhelical turns consumed by that transition can be identified using two-dimensional gel electrophoresis.

Pyr/Pur sequences arranged in mirror symmetry have the potential to form DNA triplexes that are induced and stabilized by acidic conditions [1-3]. Moreover, long (more than 100 bp) Pyr/Pur sequences, which appear frequently in eukaryotic genomes [16,17], can form other unusual DNA structures under acidic conditions. Because the center of symmetry within long Pyr/Pur sequences can vary, alternative DNA structures forming at long Pyr/Pur tracts are heterogeneous, enabling triplex stems of various size and multiple triplexes to be seen. Consistent with such heterogeneity, supercoiled DNA molecules show a smeary mobility, rather than discrete bands (or spots), when subjected to electrophoresis under acidic conditions [18-20]. As an example, agarose gel electrophoresis of another pUC19 derivative, pTIR10, which contains a rat genomic DNA fragment with a ~230-bp Pyr/Pur insert is shown in Fig. 2. Note

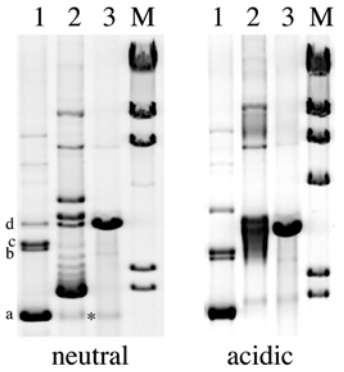

Figure 2. Electrophoretic profiles of plasmid pTIR10 under neutral and acidic conditions. Left panel, electrophoresis on 1% agarose in 40 mM Tris-acetate/ 5 mM sodium acetate/1 mM EDTA (pH7.5); right panel, electrophoresis on 1% agarose in 30 mM sodium acetate/1mM EDTA (pH 4.6). Lane 1, pUC19 DNA; lane 2, pTIR10 DNA; lane 3, pTIR10 DNA linearized by *Hin*dIII digestion; lane M, *Hin*dIII-digested lambda phage DNA size marker. Fully supercoiled molecules, linear molecules, open circles and supercoiled dimeric molecules of pUC19 are labeled a, b, c and d, respectively, in the left panel. The faint DNA band at the bottom of lanes 2 and 3 (marked with an asterisk) might be the supercondensed structure reported previously [25, 26]. Reproduction from Kato *et al*. [20] with permission of FEBS.

the smeary profile produced by electrophoresis in acidic buffer, and that the inserted region is sensitive to S1 nuclease when under superhelical tension [18]. But whereas analysis of electrophoretic mobility is useful for detecting how structural transitions affect the global topology of DNA, it is not designed to show what type of local conformational change has occurred.

Chemical and enzymatic probing of DNA conformation

A number of chemical compounds that react with bases or with the sugar-phosphate backbone are used to define the tertiary structures of DNA and chromatin. Because susceptibility to these chemicals is strongly dependent on the accessibility of their target sites, the sites attacked by chemicals are thought to be exposed to the environment. The nature of these chemical reactions and the results of probes of several DNA structures have been described [2,3].

The Pyr/Pur sequence in the promoter region of the human epidermal growth factor receptor gene (EGFR) can be regarded as possessing either direct repeat or mirror symmetry. The region is sensitive to S1 nuclease [21] and, at acidic pHs, the 5'-half of the polypurine strand is susceptible to attack by diethylpyrocarbonate (Fig. 3), which mainly reacts with unstacked adenine and less with guanine. Acidic pH makes unpaired cytosines more likely to be protonated, thereby promoting formation of C^+:G:C base triads, which is indicative of the potential of the promoter region of EGFR to form intramolecular triplexes [22]. The role of the potential triplex region in regulating EGFR expression is unclear, however. This region was found to be unnecessary for maximum transcriptional activity of the promoter *in vivo* and *in vitro* [23,24], but Johnson *et al*. [21] showed that mutations in this region reduce promoter activity in transient expression assays.

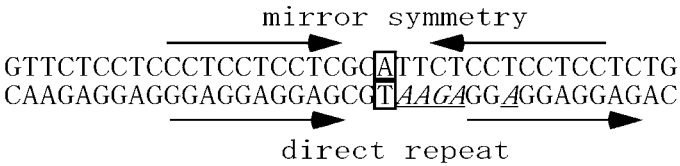

Figure 3. Nucleotide sequence of Pyr/Pur stretch at the human EGFR promoter region. Shown are positions -374 to -337; the numbering is according to Johnson *et al*. [21]. The region can be regarded as possessing either direct repeat or mirror symmetry. The center of symmetry is boxed, and the sites attacked by diethylpyrocarbonate are underlined and italicized. S1 nuclease mainly cleaved the 3' side of the boxed nucleotides (between A and T in the top strand, and between T and G in the bottom strand). The results indicate that the region may produce triplexes occurring at the mirror symmetry rather than DNA slippage occurring at direct repeats. See ref. 22 for details.

It is generally accepted that DNA containing short inverted repeats may assume a cruciform structure [1]. On the other hand, Ohta *et al*. [27] and Kato *et al*. [28] have proposed an alternative DNA structure that is distinct but derived from the DNA cruciform. In that structure, the two stem-loops of the cruciform are arranged in parallel and, possibly, one stem-loop melts to yield a third strand able to form a triple-stranded structure. We tested that idea using potassium permanganate, which efficiently attacks unpaired thymine bases [15]. We found that permanganate reacted with the thymines in the 3'-half of the inverted repeat as well as with those at the middle of the inverted repeat (Fig. 4), which is consistent with the hairpin triplex model proposed by Kato *et al*. [28]. This structure is also formed in *E. coli* cells. Although the permanganate reacted with DNA more efficiently under acidic conditions, the profiles of permanganate oxidation were essentially identical under both neutral (lanes 1, 2, and 3 of Fig. 4) and acidic (lanes 4, 5, and 6 of Fig. 4)

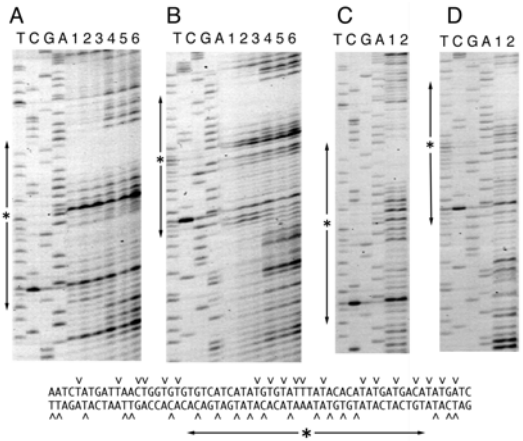

Figure 4. Primer extension assay of permanganate-treated DNA. The products of primer extension were electrophoresed along with control sequencing ladders (lanes T, C, G and A). Permanganate oxidation was carried out *in vitro* (panels A and B) and *in situ* (panels C and D). Panels A and C show results obtained with the M13 forward primer, while panels B and D show those obtained with the M13 reverse primer. Asterisks and arrows at the left of the panels indicate the positions of the inverted repeat symmetry. A summary of sites susceptible to permanganate oxidation (indicated by arrowheads) is shown at the bottom. Panels A and B: lanes 1 and 4, 1 min permanganate reaction; lanes 2 and 5, 2 min permanganate reaction; lanes 3 and 6, 4 min permanganate reaction. The reactions in lanes 1, 2 and 3 were in neutral buffer; those in lanes 4, 5 and 6 were in acidic buffer. Panels C and D: lane 1, permanganate reaction with mid-log growth phase *E. coli* cells; lane 2, permanganate reaction with late-log growth phase *E. coli* cells. Reproduction from Kato *et al*. [15] with permission of the Biophysical Society.

conditions, suggesting that cytosine protonation is not involved in formation of hairpin triplex, but parallel G:G:C and A:A:T triads may be. Such parallel triplexes have been described previously [29,30].

Microscopic imaging of DNA and chromatin

Because biochemical measurements generally describe the average structure of the molecules in a DNA sample, it is difficult to characterize heterogeneous populations of molecules. For example, because long Pyr/Pur sequences may form triplex stems of various size or cause formation of multiple stems in one molecule, electrophoretic profiles are smeary and most of the potential target sites are attacked by chemicals in an uninterpretable manner. In such cases, microscopic imaging is a powerful tool with which to take a snapshot of individual molecules.

We have used atomic force microscopy (AFM) to characterize the heterogeneity of triplex formation in pTIR10 plasmid [20]. AFM imaging of supercoiled pTIR10 revealed the presence of various triplex stems, which was expected based on the biochemical observations (Fig. 5). AFM has the

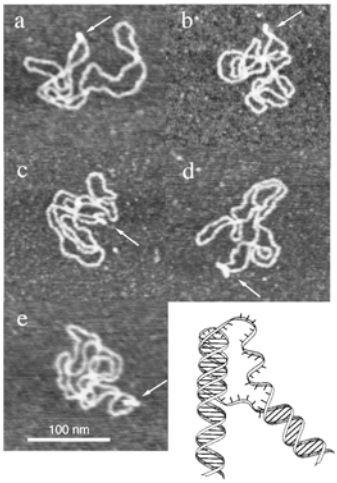

Figure 5. AFM images of the most representative families of irregularities found in pTIR10 DNA. Panel a, short triplex-like stem; panel b, long curved stem; panel c, twin stems; panel d, Π-shaped structure; panel e, Ψ-shaped structure. Scale bar (100 nm) is given at the bottom of panel e and is common to all panels. Potential triplex stem regions are indicated by arrows. Structures shown in panels c, d and e are multiple triplex formations. A model showing a conventional intra-molecular triplex is given at the right bottom. Reproduction with modification from Kato *et al*. [20] with permission of FEBS.

advantage that no staining is required for observation, but there is the disadvantage that, due to a convolution effect [31] caused by the probe tip, the apparent width of the DNA is greater than the actual size. Electron microscopy with heavy metal staining gives sharp (high contrast) images of DNA fiber. Plasmids pBan1 and pTIR10 visualized using electron microscopy with uranyl acetate staining are shown in Fig. 6.

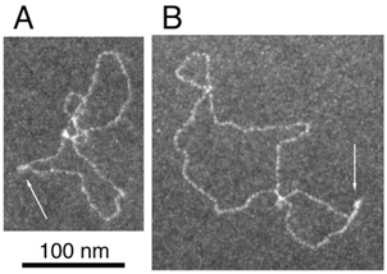

Figure 6. Electron microscopic images of alternative DNA structures. Panel A, supercoiled DNA containing short inverted repeat (pBan1); panel B, supercoiled DNA containing a Pyr/Pur sequence (pTIR10). The stem part is indicated in each panel. Reproduction from Kato *et al*. [15] with permission of the Biophysical Society.

Conventional electron microscopy usually requires artificial fixation of specimens followed by heavy metal staining, so the images obtained should be evaluated cautiously. However, procedures for observing nearly native (unfixed and unstained) samples have been developed with the aim of avoiding chance artifactual deformation of the specimen that can be caused by staining. Bednar *et al*. characterized the structures of nucleosome [32] and transcribing RNA polymerase on the chromatin template [33], and Bouchet-Marquis *et al*. [34] analyzed nuclear architecture by cryo-electron microscopy. In addition, Kaneko *et al*. recently used a novel type of electron microscopy, Hilbert differential transmission electron microscopy, to analyze the subcellular structure of cyanobacteria in a near-living state [35,36]. These techniques will likely prove effective for characterizing alternative DNA structures occurring *in vivo*.

Conclusion

Many sequence tracts within genomes have the potential to form alternative DNA structures; the formation of such structures has been observed *in vitro*, and the kinetics of their structural transition have been studied. For instance, long Pyr/Pur sequences are scattered abundantly within genomes and may assume a variety of alternative DNA structures, including multiple

intramolecular triplexes. Indeed, probing with anti-triplex DNA antibody revealed the presence of numerous triplexes within the interphase nucleus [37]. In addition, intermolecular interaction of two intramolecular triplexes has been reported for the $(GAA/TTC)_n$ triplet repeat in the first intron of the human frataxin gene, which yields sticky DNA [38], and formation of sticky DNA directly correlates with inhibition of transcription [39]. Thus, certain Pyr/Pur loci may offer triplex stems for inter- and intrachromosomal interactions to modulate transcription and replication. That said, little is known about the function of alternative DNA structures within the genome. To define the physiological role of Pyr/Pur sequences and other characteristic DNA sequences, it will be necessary to directly observe alternative DNA structures at work *in vivo* (or *in situ*). To that end, improved microscopic imaging techniques may enable us to detect the *in vivo* dynamics of DNA and chromatin.

References

1. Sinden, R. R., 1994, *DNA Structure and Function*. Academic Press, San Diego. ISBN 0-12-645750-6
2. Wells, R. D. and Harvey, S. C. (editors), 1988, *Unusual DNA Structures*, Springer-Verlag, New York. ISBN 0-387-96631-5
3. Soyfer, V. N. and Potaman, V. N., 1995, *Triple-Helical Nucleic Acids*. Springer-Verlag, New York. ISBN 0-387-94495-8
4. Schroth, G. P., Chou, P.-J. and Ho, S., 1992, *J. Biol. Chem.* **267**, 11846-11855.
5. Schroth, G. P., and Ho, S., 1995, *Nucleic Acids Res*. **23**, 1977-1983.
6. Kohwi Y and Kohwi-Shigematsu, T., 1991, *Genes Dev*. **12**, 2547-2554.
7. Kato, M and Shimizu, N., 1992, *J. Biochem.* (Tokyo) **112**, 492-494.
8. Brukner, I., Belmaaza, A. and Chartrand, P., 1997, *Proc. Natl. Acad. Sci. USA* **94**, 403-406.
9. Chen, S., Supakar, P. C., Vellanoweth, R. L., Song, C. S., Chatterjee, B. and Roy, A. K., 1997, *Mol. Endocrinol*. **11**, 3-15.
10. Zain, R. and Sun, J-S., 2003, *Cell. Mol. Life Sci*. **60**, 862-870.
11. Sinden, R. R., 1999, *Am. J. Hum. Genet*. **64**, 346-353.
12. Pelletier, R., Krasilnikova, M. M., Samadashwily, G. M., Lahue, R. and Mirkin, S. M., 2003, *Mol. Cell. Biol*. **23**, 1349-1357.
13. Krasilnikova, M. M., Mirkin, S. M., 2004, *Mol. Cell. Biol*. **24**, 2286-2295.
14. Hanai, R., and Rocca, J., 1999, *Methods Mol. Biol*. **94**, 19-27.
15. Kato, M., Hokabe, S., Itakura, S., Minoshima, S., Lyubchenko, Y. L., Gurkov, T. D., Okawara, H., Nagayama, K. and Shimizu, N., 2003, *Biophys. J.* **87**, 402-408.
16. Straus, N. A. and Birnboim, H. C., 1976, *Biochim. Biophys. Acta* **454**, 419-428.
17. Birnboim, H. C., 1978, *J. Mol. Biol*. **121**, 541-549.
18. Kato, M., 1993, *Mol. Biol. Reports* **18**, 183-187.
19. Blaszak, R. T., Potaman, V., Sinden, R. R. and Bissler, J. J., 1999, *Nucleic Acids Res*. **27**, 2610-2617.

20. Kato, M., McAllister, C. J., Hokabe, S., Shimizu, N. and Lyubchenko, Y. L., 2002, *Eur. J. Biochem.* **269**, 3632-3636.
21. Johnson, A. C., Jinno, Y. and Merlino, G. T., 1988, *Mol. Cell. Biol*. **8**, 4174-4184.
22. Kato, M., Kudoh, J. and Shimizu, N., 1990, *Biochem. J.* (London) **268**, 175-180.
23. Johnson, A. C., Ishii, S., Jinno, Y., Pastan, I. and Merlino, G. T., 1988, *J. Biol. Chem*. **263**, 5693-5699.
24. Kageyama, R., Merlino, G. T. and Pastan, I., 1988, *J. Biol. Chem*. **263**, 6329-6336.
25. Huang, X. and Chen, X., 1990, *J. Mol. Biol.* **216**, 195-199.
26. Kato, M. and Furuno, A., 1992, *Res. Microbiol*. **143**, 665-669.
27. Ohta, T., Nettikadan, S., Tokumasu, F., Ideno, H., Abe, Y., Kuroda, M., Hayashi, H. and Takeyasu, K., 1996, *Biochem. Biophys. Res. Commun*. **226**, 730-734.
28. Kato, M., Matsunaga, K. and Shimizu, N., 1998, *Biochem. Biophys. Res. Commun.* **246**, 532-534.
29. Shchyolkina, A. K., Borisova, O. F., Minyat, E. E., Timofeev, E. N., Il'icheva, I. A., Khomyakova, E. B. and Florentiev, V. L., 1995, *FEBS Lett*. **367**, 81-84.
30. Shchyolkina, A. K, Timofeev, E. N., Lysov, Y. P., Florentiev, V. L., Jovin, T. M. and Arndt-Jovin, D. J., 2001, *Nucleic Acids Res.* **29**, 986-995.
31. Yang, J., Mou, J., Yuan, J. Y. and Shao, Z., 1996, *J. Microsc*. **182**, 106-113.
32. Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J. and Woodcock, C. L., 1998, *Proc. Natl. Acad. Sci*. USA **95**, 14173-14178.
33. Bednar, J., Studitsky, V. M., Grigoryev, S. A., Felsenfeld, G. and Woodcock, C. L., 1999, *Mol. Cell* **4**, 377-386.
34. Bouchet-Marquis, C., Dubochet, J. and Fakan, S., 2006, *Histochem. Cell Biol*. **125**, 43-51.
35. Kaneko, Y., Danev, R., Nitta, K. and Nagayama K., 2005, *J. Electron Microsc*. (Tokyo) **54**, 79-84.
36. Kaneko, Y., Danev, R., Nagayama, K. and Nakamoto H., 2006, *J. Bacteriol*. **188**, 805-808.
37. Ohno, M., Fukagawa, T., Lee, J. S. and Ikemura, T., 2002, *Chromosoma* **111**, 201-213.
38. Sakamoto, N., Chastain, P. D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J. D. and Wells, R. D., 1999, *Mol. Cell* **3**, 465-475.
39. Sakamoto, N., Ohshima, K., Montermini, L., Pandolfo, M. and Wells, R. D., 2001, *J. Biol. Chem*. **276**, 27171-27177.