A single trinucleotide, 5'AGC3'/5'GCT3', of the triplet-repeat disease genes confers metal ion-induced non-B DNA structure

Yoshinori Kohwi, Hua Wang and Terumi Kohwi-Shigematsu

La Jolla Cancer Research Foundation, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

Received September 2, 1993; Revised and Accepted October 28, 1993

ABSTRACT

Expansion of (AGC)_n repeats has been associated with genetic disorders called triplet-repeat diseases such as Huntington's disease (HD), myotonic muscular dystrophy (DM) and Kennedy's disease. To gain insight into the abnormal behavior of these repeats, we studied their structural properties in supercoiled DNA. Chemical probing revealed that, under physiological salt and pH conditions, Zn²⁺ or Co²⁺ ions induce (AGC)_n repeats to adopt a novel non-B DNA structure in which all cytosine but none of adenine residues in either strand become unpaired. The minimum size of (AGC)_n repeat that could form this structure independently of neighboring sequences is a single unit of double-stranded trinucleotide, 5'AGC3'/5'GCT3'. Other trinucleotide units of the same nucleotide composition, 5'CAG3'/5'CTG3' or 5'GCA3'/5'TGC3', do not form non-B DNA structures. This unusual DNA structural property adopted by a single 5'AGC3'/5'GCT3' trinucleotide may contribute to expansion of (AGC)_n sequences in triplet-repeat diseases.

INTRODUCTION

The discovery of huge numbers of the triplet-repeat sequence (CGG)_n in the mutant gene responsible for the inherited mental-retardation disorder fragile X syndrome (1–6) triggered triplet-repeat research in other genetic diseases. Patients with Myotonic Dystrophy (DM) contain an increased length of the (AGC)_n repeats of a DM candidate gene (7–12). The AGC triplet-repeat expansion has also been identified in X-linked spinal and bulbar muscular atrophy (SBMA or Kennedy's disease) (13). Most recently, the gene responsible for Huntington's disease (HD) was discovered and revealed that a polymorphic (AGC)_n repeat is expanded and unstable on HD chromosomes (14). How triplet repeats are miscopied in these diseases is unknown. The (AGC)_n repeats may possess intrinsic structural properties, and if so, such properties of the triplet repeats could underlie their abnormal behavior in cells. Such an idea may serve as a first step in deciphering the mystery of the cause of triplet repeat diseases.

Certain DNA sequences, when subjected to negative superhelical strain, adopt non-B DNA structures. These sequences

include an inverted repeat sequence that forms a cruciform structure (reviewed in 15). To test whether an (AGC)_n repeat is also capable of adopting a non-B DNA structure under superhelical strain, we investigated a supercoiled plasmid DNA containing an (AGC) repeat comparable in length to that found in the HD gene (14). We also examined the possible non-B DNA structure of another naturally occurring AGC repeat sequence encoding glutamine stretch found in the nuclear matrix attachment in the gene for HD. We also examined the possible non-B DNA structure of another naturally occurring AGC repeat sequence encoding glutamine stretch found in the nuclear matrix attachment region (MAR) binding protein SATB1 (16). The DNA structure was studied using a chemical probe chloroacetaldehyde that is specific for unpaired adenine and cytosine residues. Chloroacetaldehyde and its analog bromoacetaldehyde have been successfully used to analyze various non-B DNA structures described above (reviewed in 17 and 18).

Metal ions and pH are known to play important roles in the formation of specific non-B DNA structures. For example, under superhelical strain, poly(dG)-poly(dC) sequences adopt a structure containing an intramolecular poly(dG).poly(dG).poly(dC) triplex (H'-DNA, 19) in the presence of Mg²⁺ ions at pH5–8, while the same sequences adopt a poly(dC)⁺.poly(dG).poly(dC) triplex (H-DNA, 20) in the absence of Mg²⁺ ions at pH5 (21). Poly(dG.dA)-poly(dC.dT) sequences, on the other hand, adopt an H-DNA structure that consists of two pyrimidine strands and one purine strand in the presence of Mg²⁺, whereas in the presence of Zn²⁺ ions, the same sequences form an H'-DNA structure that consists of two purine strands and one pyrimidine strand (22). This series of studies led us to examine the (AGC)_n repeats in the absence or presence of Mg²⁺, Zn²⁺ or Co²⁺. The effect of the length of the triplet repeats on the non-B DNA structure formation was also examined.

Our structural analysis of the (AGC)_n repeats revealed that this type of triplet repeats does form a novel non-B DNA structure stabilized by Zn²⁺ and Co²⁺ ions under superhelical strain, which can be detected by the selective reactivity with chloroacetaldehyde of every cytosine residue within the repeats, but none of the adenine residues. Interestingly, one single trinucleotide, AGC, is necessary and sufficient to confer this non-B DNA structure.

MATERIALS AND METHODS

Chemicals

Chloroacetaldehyde was purchased from Fluka and doubly distilled (boiling point, 78–80°C). The purified chloroacetaldehyde was aliquoted in small fractions and stored at –20°C until used. Hydrazine, formic acid and piperidine were purchased from Sigma.

Plasmid DNA

Synthetic oligomer AGC repeat (30 bases) and the complementary GCT repeat (30 bases) was annealed and ligated to make multimers. The ligated double-stranded (AGC/GCT) repeats were inserted at the EcoRV site in Bluescript (Stratagene) after they were blunt-ended by the Klenow fragment of DNA polymerase I. Two plasmid DNAs that contain (AGC)₃₅ or (AGC)₁₁AG(A-GC)₂₅ were prepared in large scale by the method of alkaline-SDS lysis and purified twice by CsCl-density gradient ultracentrifugation. We also employed plasmid DNA which contained naturally occurring AGC repeat sequence of SATB1 cDNA (16).

Mapping of chloroacetaldehyde-modified sites

Supercoiled plasmid DNAs containing AGC repeat sequences were reacted with 2 μ l of doubly distilled chloroacetaldehyde in 100 μ l reaction volume at 37°C under the conditions described in the figure legends. Chloroacetaldehyde-modified DNAs were end-radiolabeled by Klenow fragment of polymerase I as indicated in the Figure legends and further cleaved at a distal site with the second restriction enzyme. The DNA fragment was isolated from a native acrylamide gel and the eluted-DNA from the gel was subjected to hydrazine (HZ) or formic acid (FA) reaction followed by the piperidine reaction as described (22). The adenine or cytosine residues that are unpaired in supercoiled DNAs are modified with chloroacetaldehyde. When unpaired DNA bases are modified with chloroacetaldehyde, the DNA becomes susceptible to cleavage by piperidine at these sites. For chloroacetaldehyde-modified cytosine residues, the cleavages at these sites can be detected by the appearance of new bands on a sequencing gel after the DNA is treated with formic acid and piperidine reactions which normally cleave DNA at adenine and guanine residues. Similarly, the cleavages of DNA due to chloroacetaldehyde-modified adenine and guanine residues can be detected as new bands within a cytosine ladder generated after the hydrazine/piperidine reactions.

RESULTS

Non-B DNA structure formation by synthetic AGC repeats

The supercoiled plasmid DNA containing an (AGC)₂₂AG(A-GC)₁₁ sequence was modified with chloroacetaldehyde either in the absence or presence of Mg²⁺, Co²⁺, or Zn²⁺ ions. The chloroacetaldehyde-modified bases were determined as described in Materials and Methods. In the presence of Zn²⁺ or Co²⁺, all cytosine residues within both strands of the (AGC)₂₂ and (A-GC)₁₁ repeats were reactive with chloroacetaldehyde (Fig. 1A; AGC strand, lanes 9, 10, and 1B; GCT strand, lanes 5, 6; indicated by arrowhead and bar) but all adenine residues remained unreactive (Fig. 1A; lanes 1–5). No such chemical reactivity was detected either with Mg²⁺ or without metal ions (Fig. 1A, lanes 7, 8). This shows that under superhelical strain, Zn²⁺ or Co²⁺ ions stabilize a specific non-B DNA structure formed by the (AGC) repeats.

Figure 1. Structural analysis of a synthetic (AGC)₂₂AG(AGC)₁₁ repeat. The supercoiled plasmid DNA containing the (AGC)₂₂AG(AGC)₁₁ repeat was reacted with 2 μ l of chloroacetaldehyde in 100 μ l of 50mM of Tris-Maleate buffer at pH 5 in the presence of 2mM of MgCl₂, CoCl₂ or ZnCl₂. The fragments as described in (A) and (B) were isolated from a non-denaturing polyacrylamide gel and subjected to either hydrazine (HZ) or formic acid (FA) and subsequently treated with piperidine. The samples were loaded onto a urea/8% polyacrylamide gel. (A) The strand of AGC repeat was analyzed from the XbaI–SalI fragment labeled at the XbaI site by α -³²P dATP and Klenow fragment of polymerase I. Lanes 1,6; control (C; no chloroacetaldehyde modification), Lanes 2,7; in the absence of metal ions, Lanes 3, 8; 2mM MgCl₂, Lanes 4,9; 2mM CoCl₂, Lanes 5,10; 2mM ZnCl₂. (B) The GCT strand of the repeat was analyzed from the XhoI–XbaI fragment labeled at the XhoI site by α -³²P dATP and klenow fragment of polymerase I. Lanes 1,4; control (C; no chloroacetaldehyde modification), Lanes 2,5; 2mM CoCl₂, Lanes 3,6; 2mM ZnCl₂.

The non-B DNA structure of triplet repeats forms under physiological conditions

We tested whether the Zn²⁺-dependent structure also forms in a more physiological condition where the concentration of Mg²⁺ ions is considerably higher than that of Zn²⁺ and at a physiological salt concentration. The Zn²⁺ ion-dependent structure formation was examined in the presence of both Mg²⁺ and Zn²⁺ at varying ratios in a 50mM Tris-maleate buffer, pH7, containing an additional 100mM NaCl. A strong chloroacetaldehyde reactivity with every cytosine residue within a (AGC)₃₅ was detected with 1mM Zn²⁺ and 2mM Mg²⁺, indicating that the Zn²⁺-dependent structure formed regardless of the presence of Mg²⁺ ions (Fig. 2). Even in the presence of 2mM Mg²⁺ ions, as long as 0.2mM Zn²⁺ ions were present, the Zn²⁺-dependent structure still formed. Therefore, the non-B DNA structure adopted by the triplet repeats forms under more physiological conditions and the effect of Zn²⁺ ions prevails over the effect of Mg²⁺ ions.

Figure 2. Structural analysis of an $(AGC)_{35}$ repeat under physiological conditions. The plasmid DNA containing an $(AGC)_{35}$ repeat was modified with chloroacetaldehyde in the presence of $MgCl_2$ and $ZnCl_2$ (as indicated at the top of lanes) in 100mM of NaCl and 50mM of Tris-Maleate buffer at pH 7. The chloroacetaldehyde-modified sites were determined as described in Fig.1 legend and are indicated by arrowheads and bars. Lane 1; 2mM $MgCl_2$ and 1mM $ZnCl_2$, Lane 2; 2mM $MgCl_2$ and 0.5mM $ZnCl_2$, Lane 3; 2mM $MgCl_2$ and 0.2mM $ZnCl_2$, Lane 4; control (C; no chloroacetaldehyde modification).

Zn^{2+} -dependent non-B DNA structure of naturally occurring AGC repeat sequence

We examined whether the metal-induced non-B DNA structure forms only in long, homogenous triplet repeats or forms in much shorter segments of the repeat as well. A naturally occurring $(AGC)_n$ repeat was employed that is interrupted by other triplet sequences. We studied a specific DNA region encoding 15 glutamines in the gene for the nuclear matrix attachment DNA (MAR) associating protein SATB1 (16). This glutamine repeat (CAG repeat) is interrupted three times by other CAA triplets to give short AGC repeats (Figure 2B, The AGC triplets are underlined). In the presence of Zn^{2+} at either pH5 or pH7, chloroacetaldehyde reacted specifically with every cytosine within 5'AGC/GCT3' trinucleotides encoding the glutamine stretch (Fig. 3A, lanes 3, 6, Fig.3B indicated by arrowheads). Cytosines within the interrupting triplets unrelated to the 5'AGC/GCT3' trinucleotide did not confer chloroacetaldehyde reactivity.

What is the minimum length of the $(AGC)_n$ repeat that forms this Zn^{2+} -induced structure? In the polyglutamine encoding region of the SATB1 gene, there are some single units of 5'A-GC/GCT3' trinucleotide that are separated by other triplet sequences. Cytosines in these single 5'AGC/GCT3' trinucleotides also showed cytosine reactivity with chloroacetaldehyde (Fig. 3A, lanes 3, 6 shown by an arrow head, * and number that corresponds to the sequence number shown in Fig. 4A).

Figure 3. Structural analysis of the polyglutamine encoding region of the cDNA for the MAR binding protein SATB1. The chloroacetaldehyde-modified sites in the plasmid DNA (pAT11, ref.16) containing the cDNA sequence of the MAR binding protein SATB1 was determined as described in the legend to Fig. 1 at pH5 or pH7. The chloroacetaldehyde-modified sites are indicated by arrowheads. (A) The AGC strand of the trinucleotide repeats in the cDNA was analyzed from the XbaI–NcoI fragment labeled at the XbaI site by α - ^{32}P dATP and Klenow fragment of polymerase I. Lane 1; control (C; no chloroacetaldehyde modification), Lanes 2,5; 2mM $MgCl_2$, Lanes 3, 6; 2mM $ZnCl_2$, Lanes 4, 7; no metal ions, Lanes 1–4; pH5, Lanes 5–7; pH7. (B) The polyglutamine encoding sequence of the SATB1 cDNA is shown. The chloroacetaldehyde-modified residues within the NcoI site (position 1493) in the SATB1 cDNA and the XbaI site in the Bluecript vector are indicated by arrowheads. Every 5'AGC/GCT3' trinucleotide is underlined. The positions of the first and the last glutamines of the glutamine stretch are indicated by Q1 and Q15. Isolated CAG and GCA trinucleotides are boxed. Isolated AGC trinucleotide units that are next to different neighboring sequences are each denoted with an arrowhead, star, and number that corresponds to the number of each matching sequence shown in Fig. 4B.

Figure 4. Determination of the minimum sequence within the AGC repeats that confers the non-B DNA structure. (A) Plasmid DNAs, pUC19 and Bluescript, were modified with chloroacetaldehyde and the modified sites were analyzed by the formic acid/piperidine reaction as described in Figure 1 legend. Lanes 1 and 2; pUC19, HindIII-end labeled BglII fragment, Lanes 3 and 4; Bluescript, SalI-end labeled PvuII fragment, Lanes 1 and 3; control (C; no chloroacetaldehyde modification), Lanes 2 and 4; 2mM $ZnCl_2$. Single AGC trinucleotide units are each denoted with an arrowhead, star, and number that corresponds to the number of each matching sequence in Fig. 4B. (B) The chloroacetaldehyde-modified cytosines in isolated AGC trinucleotide units and their surrounding sequences described in this report are summarized. The consensus sequence that confers the non-B DNA structure is underlined and the chloroacetaldehyde-modified cytosine residues are indicated by bold type.

Interestingly, we also found a totally isolated 5'AGC/GCT3' triplet in the vector sequences shown in the same polyacrylamide gel, and its cytosine residues also reacted with chloroacetaldehyde (Fig. 1, 3A and 3B; indicated by an arrow head, * and number as described above).

A single trinucleotide AGC is necessary and sufficient for the non-B DNA structure formation

Results described above suggest that even a single 5'AGC/GCT3' trinucleotide is sufficient to adopt the Zn^{2+} -induced structure under superhelical strain, irrespective of its neighboring sequences. This point was further evaluated by placing a 5'A-GC/GCT3' trinucleotide next to a variety of sequences. In

addition to five examples shown in Figures 1 and 3, we tested five more cases from the vector sequences where a single trinucleotide is adjacent to different sequences using pUC19 and Bluescript plasmid DNAs. The results showed that, under native superhelical strain and in the presence of Zn^{2+} , a single trinucleotide unit, 5'AGC/GCT3', forms this non-B DNA structure in all cases examined (Fig. 4A, lane 2 and 4, summarized in Fig. 4B, for sequences 9 and 10, data not shown). Thus, independent of neighboring sequence, a 5'AGC/GCT3' trinucleotide is indeed both necessary and sufficient to confer the metal-induced non-B DNA structure under superhelical strain.

The $(AGC)_n$ repeats can be equally described as $(GCA)_n$ repeats or $(CAG)_n$ repeats. However, these two triplets, 5'GC-A/TGC3' and 5'CAG/CTG3', do not confer any unusual DNA structure as a single triplet unit (see Fig. 3A and 3B, boxed sequences). At least from the viewpoint of DNA structure, the 5'AGC/GCT3' trinucleotide is the building block of the triplet repeats.

DISCUSSION

Recent studies show that triplet repeats may be responsible for the cause of genetic disorders and the mechanisms of the 'triplet repeat diseases' are one of the most exciting topics in molecular genetics today. Nevertheless, there is not even a clue for understanding the possible role of triplet repeats in genetic disorders. What is unusual about triplet repeats that undergo expansion? As a first attempt to provide a clue for answering these questions, we tested a hypothesis that $(AGC)_n$ repeats have unusual structural properties. Unusual structures may affect the biological processes such as replication and transcription. There is evidence strongly suggesting that a poly(dG)-poly(dC) sequence can affect transcription of its downstream gene by changing its structure: when it is double-stranded, a *trans*-acting factor would bind to enhance transcription whereas when it forms a triplex, the factor could no longer bind and the enhancing activity is abolished (23).

Our structural analysis on AGC trinucleotide repeats revealed that these triplet repeats adopt a novel non-B DNA structure, and most surprisingly, only one specific trinucleotide, AGC, is both necessary and sufficient to confer this structure. This was unexpected because one would normally assume that if a non-B DNA structure forms, it would form only as a consequence of the triplet amplification: the longer the sequence is, the easier the formation of the non-B DNA structure is. This was not the case for the AGC triplet repeats. The triplet repeats consist of clusters of non-B DNA structure units formed by each of the AGC trinucleotides. The hallmark of this non-B DNA structure formed by the $(AGC)_n$ sequence under superhelical strain is its cytosine-specific reactivity with chloroacetaldehyde. Although the adenine residues, one at the 5' ends of each double-stranded trinucleotide, must be involved in the formation of the metal ion-induced DNA structure, these adenine residues are totally unreactive with chloroacetaldehyde indicating that their N(6) and N(1) positions are involved in hydrogen bonding. We employed several other chemical probes for the study of the non-B DNA structure of AGC repeat sequences. Neither adenine or guanine residues in the trinucleotide were reactive with diethyl pyrocarbonate (DEPC) which reacts strongly with purines in the *syn* configuration within Z-DNA or in single stranded DNA (reviewed in 17). Their complementary thymine residues were also unreactive with osmium tetroxide, a chemical probe that

reacts with thymine residues in distorted DNA such as B-DNA: Z-DNA junctions (reviewed in 24). We performed a dimethyl sulfoxide (DMS) protection assay (25) in the supercoiled DNA containing the (AGC)_n repeats. If N7 positions of adenine and guanine residues of the triplet repeats are protected from DMS modification, these positions may be occupied as part of the non-B DNA structure. No significant protection of DMS reactivity at the N-7 positions was observed (data not shown). Kohwi previously reported that specific cytosine residues in each direct repeat located at 5' of the chicken adult β -globin gene become reactive to chloroacetaldehyde when subjected to superhelical strain in the presence of Zn²⁺, Co²⁺, or Cu²⁺ (26). There is a 5'AGC/GCT3' trinucleotide in each of the direct repeats and here we delineate that the chloroacetaldehyde reactivity previously observed is attributed to this triplet sequence.

The exact non-B DNA structure formed by the trinucleotide is difficult to elucidate at present. Metal ions, such as Mg²⁺, are known to interact exclusively with the phosphate backbone of DNA, while Zn²⁺ preferentially interact with the nucleotide bases (reviewed in 27). Water molecules may mediate chelation of adenine and guanine residues in the triplet and one of the metal ions to form a specific three dimensional structure. To stabilize such a structure, cytosine residues may have to rotate around the phosphodiester bond to expose their hydrogen bonding sites. Because the non-B DNA structure by the triplet repeats is detected only in supercoiled plasmid DNA, an X-ray crystallographic or NMR analysis for this structure is presently impossible. Because superhelical environment that could stabilize non-B DNA structures appear to exist, at least transiently, in mammalian cells (23), the triplet repeats are likely to change the structure inside cells.

We do not know whether this structural property of AGC repeat is directly relevant for sequence amplification. However, much information could be obtained in the future regarding the mechanism for the AGC amplification based on the DNA structural analysis. For example, the first cDNA cloning for a MAR binding protein was achieved based on the idea that there exists a protein that recognizes the specific structural feature (unpairing property) of MARs (28 and 29). Similarly, there might be a protein that recognizes the structural property specific for the (AGC)_n sequence. DNA structure can play an active role in biological functions. For example, an intramolecular triple helix formation appears to modulate gene expression in mammalian cells (23) and induce the transcription-dependent homologous recombination (30). In the case of the triplet repeats, it is possible that the non-B DNA structure adopted by each 5'AGC3'/5'GC-T3' triplet could trigger multiple copying by the DNA replication apparatus that results in expansion of the (AGC)_n repeats in certain diseases.

ACKNOWLEDGMENTS

We thank Drs M.Kladde and G.Graupner for careful reading of this manuscript. This work was supported by grants from the National Institutes of Health (RO1CA51377) to Y.K.

REFERENCES

1. Dietrich, H., *et al.* (1991) *Nucl. Acids Res.* **19**, 2567–2572.
2. Fu, Y.-H. *et al.* (1991), *Cell* **67**, 1047–1058.
3. Kremer, E.L., *et al.* (1991) *Science* **252**, 1711–1714.
4. Oberle, I., *et al.* (1991), *Science* **252**, 1097–1102.

5. Verkerk, A.J.M.H., *et al.* (1991), *Cell* **65**, 905–914.
6. Yu, S., *et al.* (1991), *Science* **252**, 1179–1181.
7. Harley, H.G., *et al.* (1992), *Nature* **355**, 545–546.
8. Buxton, J., *et al.* (1992), *Nature* **355**, 547–548.
9. Aslanidid, C., *et al.* (1992), *Nature* **355**, 548–551.
10. Brook, J.D., *et al.* (1992) *Cell* **68**, 799–808.
11. Mahadevan, M., *et al.* (1992) *Science* **255**, 1253–1255.
12. Fu, Y.-H., *et al.* (1992) *Science* **255**, 1256–1259.
13. La Spada, A.R., *et al.* (1991) *Nature* **352**, 77–79.
14. Huntington's Disease Collaborative Research Group. (1993) *Cell* **72**, 971–983.
15. Yagil, G. (1991) *Crit. Rev. Biochem. Mol. Biol.* **26**, 475–559.
16. Dickinson, L.A., Joh, T., Kohwi, Y. and Kohwi-Shigematsu, T. (1992) *Cell* **70**, 631–645.
17. Kohwi-Shigematsu, T. and Kohwi, Y., in *Methods in Enzymology* **212**, DNA structure, D.M. Lilley and J.E. Dahlberg Ed. (Academic Press, 1992). p155–180.
18. Leonard, N.J. (1992) *Chemtracts: Biochem. Mol. Biol.* **3**, 273–297.
19. Kohwi, Y., Malkhosyan S. and Kohwi-Shigematsu, T. (1992), *J.Mol.Biol.* **223**, 817–822.
20. Lyamichev, V.I., Mirkin, S.M. and Frank-Kamenetskii, M.D. (1986) *J.Biomol. Struct. Dyn.* **3**, 667–669.
21. Kohwi, Y. and Kohwi-Shigematsu, T. (1988) *Proc.Natl.Acad.Sci. USA* **85**, 3781–3785.
22. Maxam, A.M. and Gilbert, W., in *Methods in Enzymology* **65**, (Academic Press, 1980). p499.
23. Kohwi, Y. and Kohwi-Shigematsu, T. (1991) *Genes & Dev.* **5**, 2547–2554.
24. Palecek, E., in *Methods in Enzymology* **212**, DNA Structure, D.M. Lilley and J.E. Dahlberg Ed. (Academic Press, 1992). p139.
25. Williamson, J.R., Raghuraman, M.K. and Cech, T.R. (1989) *Cell* **59**, 871–880.
26. Kohwi, Y. (1989) *Nucl Acids Res.* **17**, 4493–4502.
27. Barton, J.K. and Lippard, S.J., *Nucleic acid-metal ion interactions* (Spiro, T.G., ed.), pp31–113, John Wiley and Sons, New York.
28. Kohwi-Shigematsu, T. and Kohwi, Y. (1990) *Biochemistry* **29**, 9551–9560.
29. Bode, J., Kohwi, Y., Dickinson, L.A., Joh, T., Klehr, D., Mielke, C. and Kohwi-Shigematsu, T. (1992) *Science* **255**, 195–197.
30. Kohwi, Y. and Panchenko, Y. (1993), *Genes & Dev.* **7**, 1766–1778.