Series

What history tells us
IX.
Z-DNA: when nature is not opportunistic

Michel Morange
Centre Cavaillès, Ecole normale supérieure, 29 rue d'Ulm, 75230 Paris Cedex 05, France
(Fax, 33-144-323941; Email, morange@biologie.ens.fr)

1. Introduction

In December 1979, Alexander Rich and his collaborators described in *Nature* the structure of a left-handed double helical DNA fragment (Wang *et al* 1979). This structure differed significantly from the B-DNA structure described 26 years before by Watson and Crick (figure 1). It was named Z-DNA because the deoxyribose-phosphate backbone follows a zig-zag course instead of the regular one found in the B-structure. But the last letter of the alphabet was also chosen to indicate that the new form was highly different from the previously described A and B structures. This discovery was not totally unexpected: since Watson and Crick, many other models had been proposed for DNA, including left-handed models, just like Z-DNA. Nevertheless, it was a big shock, and it attracted a lot of attention.

From the beginning, Z-DNA was a structure in search of a function. From 1981 to 1983, preliminary results suggested that Z-DNA might exist in physiological conditions, and play a role in the regulation of transcription. No obvious major experimental support was further obtained. In fact, negative results were reported in the following years and interest in Z-DNA progressively declined.

The strange fate of Z-DNA raises interesting historical issues. The first concerns the relation between structure and function, a hallmark of molecular biology, in which the structure of macromolecules explains their function. What happens when the determination of the structure precedes the characterization of the potential functions it could fulfil? This new structure, and its potential roles, fitted quite well with the expectations of the numerous researchers working on the regulation of gene transcription in those years. But the expectations may have been an illusion, supported by an overly opportunistic vision of nature in which what is possible and what exists have a (too) strong tendency to coincide.

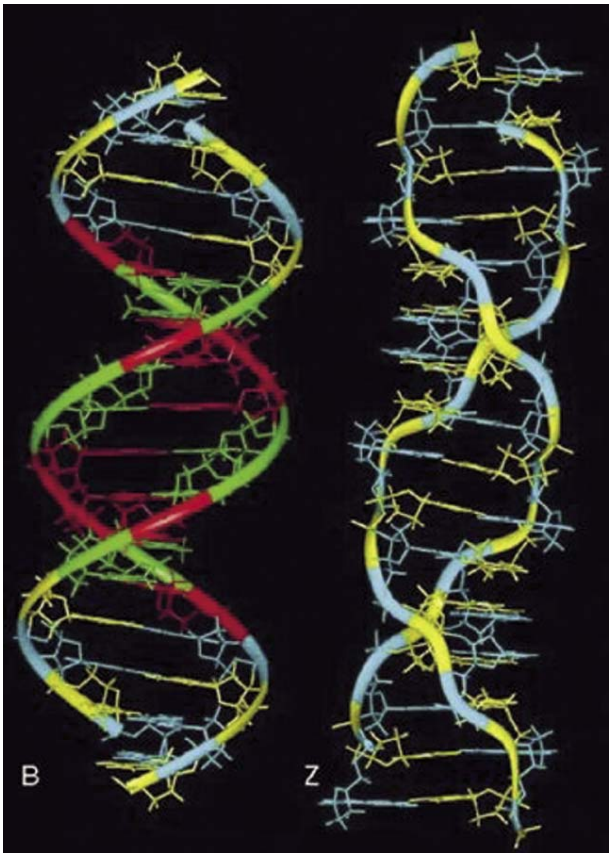

Figure 1. Line diagrams showing the B and Z forms of DNA. The base-pairing in both structures is of the normal Watson-Crick type. Nucleotides are colour-coded (cytosine in yellow, guanine in cyan, thymine in green and adenine in red), and the ribbon superposed on the backbones connects phosphorous atoms. B-DNA is a right-handed double helix with a pitch of 3.4 nm and 10 basepairs per turn. Z-DNA has a left-handed zig-zag form with a pitch of 4.5 nm and 12 base pairs per turn. (Courtesy Prof. Manju Bansal.)

Keywords. DNA conformation; Z-DNA

2. The characteristics of Z-DNA, and the context of its discovery

The Watson-Crick structure of DNA remained a model for more than 20 years: X-ray diffraction of DNA fibres provided a limited amount of data which could support the model, but not prove it. This began to change in 1973 with the crystallization of the short RNA oligomer ApU, and the 0.8 Å resolution of the diffraction pattern: the sugar-phosphate backbone had the form of a portion of a double helix (Rosenberg *et al* 1973). Jim Watson "had his first good night's sleep in 20 years" (Rich 2004), and the *News and Views* commentary in *Nature* called this discovery "the missing link". Decisive progress was made in the following years in the area of chemical synthesis, which opened the door to the preparation of the large amounts of oligonucleotides necessary for crystallization. These advances were also decisive for the rapid generalisation of DNA sequencing technology, and the development of PCR in the years ahead.

Crystallographic structures of short DNA molecules in the B form were obtained in 1980 (Wing *et al* 1980). But the paradox was that the first structure of a DNA fragment to be precisely determined, that of $d(CG)_3$, was the Z-structure (Wang *et al* 1979)! It had unusual characteristics: the repeating unit was formed of two nucleotides instead of one as in B-DNA; the helix contained twelve base pairs per turn; the dG residues were in the *syn* conformation, whereas the dC residues are in the *anti* conformation as in B-DNA; the position of the minor groove was reversed, and there was no major groove; etc. Not only did the article contain a detailed presentation of the high resolution structure of Z-DNA revealed by the diffraction study, but the authors also added many arguments in favour of the existence of this structure in solution. The form of DNA obtained by Fritz Pohl and Tom Jovin in 1972 (Pohl and Jovin 1972) in the presence of high concentrations of salt was retrospectively interpreted as being Z-DNA. The possibility of a transition between Z- and B-DNA, as well as that of a junction between these two forms along the DNA molecule was discussed. Arguments were provided that did not limit the occurrence of Z-DNA to alternating dCdG residues. The abundance of dCdG-rich sequences in the genome was noticed, as well as the association of these sequences with the initiation of replication and recombination. A link between methylation of cytosines at position 5 and the existence of Z-DNA was proposed.

3. A scientific context and early results suggestive of a biological function of Z-DNA

Many of the predictions made by Wang *et al* (1979) were subsequently confirmed. Two years later, Behe and Felsenfeld showed that cytosine methylation favoured the B to Z transition (Behe and Felsenfeld 1981). Two groups demonstrated that negative supercoiling stabilizes Z-DNA (Singleton *et al* 1982; Peck *et al* 1982). The sequences d(CA/GT)n, which had been shown to favour the formation of Z-DNA, were demonstrated to be widely distributed in eukaryotic genomes (Hamada *et al* 1982; Hamada and Kakunaga 1982). Other repeated sequences were also shown to be able to adopt the Z conformation.

In the 1970s and at the beginning of the 1980s, the search for the mechanisms regulating gene expression in eukaryotes was very active. The new genetic engineering tools allowed the characterization of sequences found upstream of genes which controlled their expression, and paved the way to the isolation of the protein factors binding to these sequences. In addition to the extension to eukaryotes of the strategy and models which had been successful in prokaryotes, other lines of research were explored, and new models of regulation proposed. Since the 1960s, the role of repeated sequences had been emphasized, and incorporated in one of the first models of eukaryotic gene transcription, the Britten-Davidson model of 1969 (Britten and Davidson 1969). Epigenetic modifications were also actively studied: DNA methylation on CpG was shown to inhibit transcription in certain cases (for a review, see Doerfler 1983). The existence of DNA supercoiling, and the enzymes involved in its control, were described (Champoux 1978). Their role in the control of gene transcription was demonstrated in prokaryotes (Yang *et al* 1979). In eukaryotes, nucleosomes, the basic components of chromatin, were responsible for DNA supercoiling, making an obvious link between DNA supercoiling and chromatin modification. The structure of the chromatin state of actively transcribed genes showed the existence of nuclease-sensitive and -hypersensitive sequences in close proximity to promoters and enhancers, but their precise nature remained unknown (Igo-Kemenes *et al* 1982). And a possible role of spermine and other polyamines in controlling transcription was not excluded, despite controversial data (Tabor and Tabor 1976, 1984; Heby 1981).

In this complex and shifting landscape, Z-DNA could appear as one of the pieces of the jigsaw puzzle which so far had been missing and was helpful for building a coherent picture of these different observations. Z-DNA formation was favoured by DNA supercoiling, cytosine methylation, the addition of spermine, and was probably present in some of the repeated sequences dispersed in the genome. From its structure, Z-DNA and its junction with B-DNA were proposed to be more sensitive to the action of nucleases than the other forms of DNA. Therefore, there was a convergence (or coalescence) of lines of research pointing to the importance of Z-DNA in transcription.

Experimental arguments rapidly accumulated in favour of such a role, thanks to production of polyclonal and

monoclonal antibodies against Z-DNA (Lafer *et al* 1981; Malfoy and Leng 1981; Moller *et al* 1982). The antibodies could also be isolated from the sera of mice and people suffering from *lupus erythematosus*, an autoimmune disease (Lafer *et al* 1983). The immunogenicity of Z-DNA, an observation not in favour of the existence of this structure *in vivo*, was explained by its high sensitivity to nucleases, which prevented its presentation to the immune system in normal conditions, as well as by the disappearance of Z-DNA as soon as supercoiling was abolished. These antibodies were used to demonstrate the presence of Z-DNA in supercoiled plasmids (Nordheim *et al* 1982) and in the enhancer of the oncogenic SV40 virus (Nordheim and Rich 1983), in polytene chromosomes of *Drosophila* (Nordheim *et al* 1981), as well as in the macronucleus of ciliates (Lipps *et al* 1983). In *Drosophila*, Z-DNA was present in interbands and puffs, sites of active transcription; and in ciliates, it was absent from the micronucleus, which is inactive in transcription unlike the macronucleus. Both observations were consistent with a selective association of Z-DNA with active transcription, although the dependence of the results on the methods used for the preparation of the samples was hotly debated (Hill and Stollar 1983).

But Z-DNA was more than a single piece in the mechanisms invoked to explain gene regulation in eukaryotes. Rather, it was thought to usher in a new dynamic and conformational vision of DNA (Rich 1983a; Dickerson 1983). The existence of left-handed structures of DNA had been previously proposed (Rodley *et al* 1976; Sasisekharan *et al* 1976), but with the crystallisation of Z-DNA these alternative forms received more attention. DNA was no longer a rigid molecule to which proteins bind by direct recognition of the nucleotides. DNA was a dynamic structure, oscillating between different conformations, the balance between which was controlled by proteins. The latter were able to directly recognize DNA conformations, and not only the base sequences favouring them. Therefore, the information contained in DNA was both sequential and conformational (Rich 1983b). The former determined the amino acid sequence of the proteins which are produced from the genome, whereas the latter controlled their level of expression. Conformational information could be transmitted over an appreciable length along the DNA molecule, which explained why enhancers were relatively distant from the site of transcription initiation, both in prokaryotes and eukaryotes (Nordheim and Rich 1983). This global dynamic vision of DNA functioning was in sharp contrast to the traditional reductionist approach of molecular biology.

4. The interest in Z-DNA declines

The existence of different lines of research converging on Z-DNA and the rise of a new vision of the informational role of DNA might have heralded a revolution in molecular biology, and a change of paradigm similar to what was simultaneously happening in the field of cancer with the convergence of studies on a small family of genes, the oncogenes (Morange 1993). Yet such was not the case. Overwhelmingly, the phenomena which had progressively emerged in the previous years – regulation over long distances in transcription, epigenetic modification of DNA by methylation, alternative DNA conformations – were incorporated into the "classical" models of protein-DNA interactions and did not require any radical rethink. Regulation over long distances could be easily explained by the formation of loops, methylated bases were recognized by specific proteins which favoured the formation of different chromatin states, and the recognition of DNA by proteins was shown, for most of the new factors isolated in these years, to result from a precise interaction between amino acids of the proteins and bases in the target sequences, not from a recognition of the form of DNA as such. Activation resulted from a direct protein-protein interaction with RNA polymerase, not from an indirect effect through DNA conformation. In addition, the association between supercoiling, DNA methylation, an unusual sensitivity to nucleases and the occurrence of Z-DNA was not always valid (Nickol and Felsenfeld 1983). No clear confirmation of the presence of Z-DNA *in vivo* (Peck and Wang 1985) or of its role in the control of transcription was obtained, and a general scepticism became widespread (Marx 1985; Rich and Zhang 2003). As early as 1981, Charles Cantor had put forward two extreme hypotheses (Cantor 1981): alternative DNA structures "may just be physical chemical noise, a way of removing torsional stress but otherwise unnoticed except for the sequences contained within them" and "such unique structures will have been selected to serve central or unique functions that go beyond normal sequence-specific recognition". The absence of convincing data tilted the balance towards the first hypothesis, and the numerous speculations that had accumulated on the potential role of Z-DNA (Rich 1983b, 1994) remained unsubstantiated. It appeared that Z-DNA was of interest for physical chemists, not for biologists. It is not my contention that transitions between DNA structures have no role at all. Rather, my point is that a confirmational vision of DNA oscillating between different structures – including Z-DNA – did not replace the more traditional vision. I do not mean to argue against the growing evidence that remarkable structures, such as the G-Quadruplex, have major roles to play for the physiology of the cell.

5. Alexander Rich's persistence

Trained in Linus Pauling's laboratory, Rich had made many major contributions to the characterization of macromolecules

by physicochemical techniques before his discovery of Z-DNA (Rich 2004). In parallel with the group of G N Ramachandran, together with Francis Crick he participated in the discovery of the structure of collagen; he was the first to show the formation of a hybrid between RNA and DNA; he characterized the structure of tRNA; discovered polysomes, etc. In all these cases, the characterization of the structure provided essential functional information. It was reasonable for him to expect an equally rich harvest from the discovery of Z-DNA, and the initial difficulties did not discourage him. He – and some other researchers – patiently collected data in favour of a role for Z-DNA in the control of transcription, even when the general interest in this structure decreased.

Nevertheless, Rich was right when he realized that correlations were not sufficient to convince the community of biologists. The only way to do so was to isolate proteins with an affinity for Z-DNA, and to demonstrate that the functions of these proteins were essential. A first approach allowed the isolation of a yeast protein, zuotin, but did not provide any insight into its function (Zhang *et al* 1992). A new method – the gel shift assay – was used (Herbert and Rich 1993), and two years later led to the isolation of a Z-DNA-binding nuclear-RNA editing enzyme (Herbert *et al* 1995). Still convinced that the characterization of structure was the way to understand function, Rich's group isolated the Z-DNA-binding domain of this protein (Herbert *et al* 1997), and determined its structure by X-ray diffraction (Schwartz *et al* 1999). From this knowledge, other proteins harbouring Z-DNA-binding domains were identified in the database. Although it was shown that these Z-DNA-binding domains could affect gene transcription (Oh *et al* 2002), precise biological roles for them remained elusive. This goal may have been reached recently. The E3L Z-DNA-binding protein of vaccinia virus has been characterized, its Z-DNA-binding domain has been shown to play an essential role in viral pathogenicity (Kim *et al* 2003, 2004), and there is evidence that the protein has an anti-apoptotic and gene transactivation role (Kwon and Rich 2005). However, the direct action of E3L on the control of transcription remains to be demonstrated.

Despite these recent advances, it is clear that the road to the biological function of Z-DNA is turning out to be longer and more difficult than expected (Rich and Zhang 2003). Not long ago, Rich looked back at the main discoveries to which he had contributed (Rich 2004). In that survey he accorded Z-DNA a minor place, even though he had spent about half of his scientific career looking for its function.

6. Conclusion: Serendipity and opportunism of nature, two ideas requiring caution

Since its introduction into the English language by Horace Walpole in 1754, serendipity, the art of making happy and unexpected discoveries by accident, has been considered to play an important part in scientific progress (Merton and Barber 2004). Rich himself used this term to designate the possibility of designing anti-viral drugs from his research on Z-DNA-binding proteins. "Chance favours only the prepared mind" said Pasteur (Vallery-Radot 1901). Z-DNA is an example of a discovery made by accident, where, however, belief in serendipity has so far led those who adopted it to a dead end. Despite all the evidence in its favour when it was discovered, Z-DNA has not been the cornerstone in understanding gene regulation in higher organisms. Once or twice in their scientific careers, all scientists are transiently convinced that a particular observation that they have made by accident fits all previous results and raises new prospects, "makes sense". Later, it often happens that they are obliged to renounce their hopes. Renouncing serendipity can be painful.

The so-called opportunism of nature is a second belief which is very popular among biologists. Organismal-level opportunism has an important place in the explanations provided by evolutionists within the Darwinian framework. Z-DNA can exist; therefore, one might think, the opportunities made available by it must be made use of. It must offer organisms new ways of controlling gene expression. The wealth of hypotheses proposed was incredible (Rich 1983b, 1994): Z-DNA might prevent the formation of odd DNA structures by forming an energy sink, regulate the movements of RNA polymerase, prevent trans-splicing, control the formation of loops and the action of enhancers, play an essential role in recombination, etc. So far, none of these possibilities has been shown to have been exploited by organisms, which have been far less inventive than scientists. (The exception seems to be the vaccinia virus: it is not a surprise to discover that the most opportunistic organisms are viruses.) The dictum that "what can happen will happen" is not always true (Cantor 1981). Life is the art of the possible (Jacob 1982), but not all possibilities have been made actual by organisms.

References

Behe M and Felsenfeld G 1981 Effects of methylation on a synthetic polynucleotide: the B-Z transition in poly(dG-m^{5}dC).poly(dG-m^{5}dC); *Proc. Natl. Acad. Sci. USA* **78** 1619–1623

Britten R J and Davidson E H 1969 Gene regulation for higher cells: A theory; *Science* **165** 349–357

Cantor C R 1981 DNA choreography; *Cell* **25** 293–295

Champoux J J 1978 Proteins that affect DNA conformation; *Annu. Rev. Biochem.* **47** 449–479

Dickerson R E 1983 *Sci. Am.* **249** 94–111

Doerfler W 1983 DNA methylation and gene activity; *Annu. Rev. Biochem.* **52** 93–124

Hamada H and Kakunaga T 1982 Potential Z-DNA forming sequences are highly dispersed in the human genome; *Nature* **298** 396–398

Hamada H, Petrino M G and Kakunaga T 1982 A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes; *Proc. Natl. Acad. Sci. USA* **79** 6465–6469

Heby O 1981 Role of polyamines in the control of cell proliferation and differentiation; *Differentiation* **19** 1–20

Herbert A, Lowenhaupt K, Spitzner J and Rich A 1995 Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA; *Proc. Natl. Acad. Sci. USA* **92** 7550–7554

Hill R J and Stollar B D 1983 Dependence of Z-DNA antibody binding to polytene chromosomes on acid fixation and DNA torsional strain; *Nature (London)* **305** 338–340

Igo-Kemenes T, Hörz W and Zachau H G 1982 Chromatin; *Annu. Rev. Biochem.* **51** 89–121

Jacob F 1982 *The possible and the actual* (Seattle: University of Washington Press)

Kim Y G, Muralinath M, Brandt T *et al* 2003 A role for Z-DNA binding in vaccinia virus pathogenesis; *Proc. Natl. Acad. Sci. USA* **100** 6974–6979

Kim Y G, Lowenhaupt K, Oh D B *et al* 2004 Evidence that vaccinia virulence factor E3L binds to Z-DNA *in vivo*: Implications for development of a therapy for poxvirus infection; *Proc. Natl. Acad. Sci. USA* **101** 1514–1518

Kwon J A and Rich A 2005 Biological function of the vaccinia virus Z-DNA-binding protein E3L: Gene transactivation and anti-apoptotic activity in HeLa cells; *Proc. Natl. Acad. Sci. USA* **102** 12759–12764

Lafer E M, Möller A, Nordheim A *et al* 1981 Antibodies specific for left-handed Z-DNA; *Proc. Natl. Acad. Sci. USA* **78** 3546–3550

Lafer E M, Valle R P C, Möller A *et al* 1983 Z-DNA-specific antibodies in human systemic lupus erythematosus : *J. Clin. Invest.* **71** 314–321

Lipps H J, Nordheim A, Lafer E M *et al* 1983 Antibodies against Z DNA react with the macronucleus but not the micronucleus of the hypotrichous ciliate *Stylonychia mytilus; Cell* **32** 435–441

Malfoy B and Leng M 1981 Antiserum to Z-DNA; *FEBS Lett.* **132** 45–48

Marx J L 1985 Z-DNA: Still searching for a function; *Science* **230** 794–796

Merton R K and Barber E 2004 *The travels and adventures of serendipity* (Princeton: Princeton University Press)

Möller A, Gabriels J E, Lafer E M *et al* 1982 Monoclonal antibodies recognize different parts of Z-DNA; *J. Biol. Chem.* **257** 12081–12085

Morange M 1993 The discovery of cellular oncogenes; *Hist. Philos. Life Sci.* **15** 45–59

Nickol J M and Felsenfeld G 1983 DNA conformation at the 5' end of the chicken adult *β*-globin gene; *Cell* **35** 467–477

Nordheim A, Pardue M L, Lafer E M *et al* 1981 Antibodies to left-handed Z-DNA bind to interband regions of *Drosophila* polytene chromosomes; *Nature (London)* **294** 417–422

Nordheim A, Lafer E M, Peck L J *et al* 1982 Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding; *Cell* **31** 309–318

Nordheim A and Rich A 1983 Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences; *Nature (London)* **303** 674–679

Oh D B, Kim Y G and Rich A 2002 Z-DNA-binding proteins can act as potent effectors of gene expression *in vivo; Proc. Natl. Acad. Sci. USA* **99** 16666–16671

Peck L J, Nordheim A, Rich A *et al* 1982 Flipping of cloned d(CG)n.d(CG)n DNA sequences from a right to a left-handed helical structure by salt, Co(III), or negative supercoiling; *Proc. Natl. Acad. Sci. USA* **79** 4560–4564

Peck L J and Wang J C 1985 Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence; *Cell* **40** 129–137

Pohl F M and Jovin T M 1972 Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly(dG-dC); *J. Mol. Biol.***67** 375–396

Rich A 1983a Left-handed Z-DNA; *Harvey Lect.* **78** 105–135

Rich A 1983b Right-handed and left-handed DNA: Conformational information in genetic material; *Cold Spring Harbor Symp. Quant. Biol.* **47** 1–12

Rich A 1994 Speculation on the biological roles of left-handed Z-DNA; *Ann. NY Acad. Sci.* **726** 1–16

Rich A 2004 The excitement of discovery; *Annu. Rev. Biochem.* **73** 1–37

Rich A, Nordheim A and Wang A H J 1984 The chemistry and biology of left-handed Z-DNA; *Annu. Rev. Biochem.* **53** 791–846

Rich A and Zhang S 2003 Z-DNA: The long road to biological function; *Nature Rev/genetics* **4** 566–572

Rodley G A, Scobie R S, Bates R H T and Levitt R M 1976 A possible confirmation for double-stranded polynucleotides; *Proc. Natl. Acad. Sci. USA* **73** 2959–2963

Rosenberg J M, Seeman N C, Kim J J P *et al* 1973 Double helix at atomic resolution; *Nature (London)* **243** 150–154

Sasisekharan V, Pattabiraman N and Gupta G 1978 Some implications of an alternative structure of DNA; *Proc. Natl. Acad. Sci. USA* **75** 4092–4096

Schwartz T, Rould M A, Lowenhaupt K *et al* 1999 Crystal structure of the *Zα* domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA; *Science* **284** 1841–1845

Singleton C K, Klysik J, Stirdivant S M and Wells R D 1982 Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions; *Nature (London)* **299** 312–316

Tabor C W and Tabor H 1976 1,4-Diaminobutane (putrescine), spermidine and spermine; *Annu. Rev. Biochem.* **45** 285–306

Tabor C W and Tabor H 1984 Polyamines; *Annu. Rev. Biochem.* **53** 749–790

Vallery-Radot R 1901 *The life of Pasteur* (London: Constable)

Wang A H J, Quigley G J, Kolpak F J *et al* 1979 Molecular structure of a left-handed double helical DNA fragment at atomic resolution; *Nature (London)* **282** 680–686

Wing R, Drew H, Takano T *et al* 1980 Crystal structure analysis of a complete turn of B-DNA; *Nature (London)* **287** 755–758

Yang H L, Heller K, Gellert M and Zubay G 1979 Differential sensitivity of gene expression *in vitro* to inhibitors of DNA gyrase; *Proc. Natl. Acad. Sci. USA* **76** 3304–3308

Zhang S, Lockshin C, Herbet A *et al* 1992 Zuotin, a putative Z-DNA binding protein in *Saccharomyces cerevisiae; EMBO J.* **11** 3787–3796

*e*Publication: 15 May 2007