Cyclic diguanylic acid behaves as a host molecule for planar intercalators

Yen-Chywan Liaw[1], Yi-Gui Gao[1], Howard Robinson[1], George M. Sheldrick[2], L.A.J.M. Sliedregt[3], Gijs A. van der Marel[3], Jacques H. van Boom[3] and Andrew H.-J. Wang[1]

[1]*Department of Physiology and Biophysics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA,* [2]*Institut für Anorganische Chemie der Universität, Tammannstrasse 4, D-3400 Göttingen, FRG* and [3]*Gorlaeus Laboratories, Leiden State University, 2300 RA Leiden, The Netherlands*

Received 2 March 1990

Cyclic ribodiguanylic acid, c-(GpGp), is the endogenous effector regulator of cellulose synthase. Its three-dimensional structure from two different crystal forms (tetragonal and trigonal) has been determined by *X*-ray diffraction analysis at 1 Å resolution. In both crystal forms, two independent c-(GpGp) molecules associate with each other to form a self-intercalated dimer. A hydrated cobalt ion is found to coordinate to two N7 atoms of adjacent guanines, forcing these two guanines to destack with a large dihedral angle (32°), in the dimer of the tetragonal form. This metal coordination mechanism may be relevant to that of the anticancer drug cisplatin. Moreover, c-(GpGp) exhibits unusual spectral properties not seen in any other cyclic dinucleotide. It interacts with planar organic intercalator molecules in ways similar to double helical DNA. We propose a cage-like model consisting of a tetrameric c-(GpGp) aggregate in which a large cavity ('host') is generated to afford a binding site for certain planar intercalators ('guests').

Guest-host chemistry; DNA, cyclic; X-ray diffraction; Metal-DNA interaction; Drug design; DNA conformation

1. INTRODUCTION

Cyclic ribodiguanylic acid (cyclic-(GpGp) or c-(GpGp)) is an endogenous regulator of the biological synthesis of cellulose in the Gram-negative *Acetobacter xylinum* [1]. The biological function of this small dinucleotide is likely associated with its ability to bind to the membrane-bound cellulose synthase and thereby induce a conformational change in the enzyme with subsequent activation of the enzyme. To understand better how small cyclic nucleotides can exert their biological functions, we need to know the variety of conformations that these molecules can adopt. We have recently determined the molecular structure of a related cyclic dinucleotide, c-(dApdAp), and showed that it adopts a conformation in which the 12-membered circular sugar-phosphate backbone provides a rigid framework to hold the two adenines 6.8 Å apart in parallel planes [2]. To find out whether this interesting conformation is a common feature shared by all cyclic dinucleotides, including the biologically relevant c-(GpGp), we have determined the molecular structure of the latter molecule at high resolution.

2. EXPERIMENTAL

A number of different cyclic dinucleotides were synthesized according to the published procedure [3]. Two different crystal forms (tetragonal and trigonal) were obtained for c-(GpGp). The tetragonal form was crystallized from a solution containing 2 mM of dinucleotide, 11 mM $CoCl_2$, 20 mM glycine-HCl (pH 2.1) and 7% in 2-methyl-2,4-pentanediol (2-MPD). The solution was equilibrated with 50% 2-MPD at 25°C by vapor diffusion. The trigonal form was crystallized from a similar condition except the $CoCl_2$ was replaced with 20 mM $MgCl_2$. Crystals in the form of large tetragonal bipyramids and hexagonal plates, respectively, for the two crystal forms appeared after a week. They were in the space group $I4_1$ and $P3_121$ with respective unit cell dimensions $a=b=19.656(6)$ and $c=40.00(1)$ Å and $a=b=13.523(5)$ and $c=76.24(2)$ Å. The crystal was mounted in a sealed glass capillary with a droplet of mother liquor for data collection on a Rigaku AFC-5R rotating anode X-ray diffractometer at 25°C using the ω-scan mode with CuKα radiation. Data were collected to a resolution of 0.9 Å, but data beyond 1.1 Å were weak. There were 4289 (tetragonal) and 3207 (trigonal) independent reflections observed at the 3.0 σ(F) level above the background and they were used in the refinement.

The structure was solved with considerable difficulty by the direct methods; eventually the program SHELXS-86 and SHELXS-90 [4] proved successful. It was refined by block-diagonal least squares using the NRCVAX package [5]. In the tetragonal crystal structure, there are two independent c-(GpGp) molecules, two hydrated cobalt ions (one on two-fold axis), and 22 water molecules (2 on 2-fold axis) in the asymmetric unit. In the trigonal crystal structure, two independent c-(GpGp) molecules, one hydrated magnesium ion, one disordered glycine and 21 water molecules were found. Some water molecules were disordered. All 116 (tetragonal) and 123 (trigonal) non-hydrogen atoms were refined anisotropically. Forty-four hydrogen atoms were also included in the refinement with fixed positions and isotropic temperature factors in both structures. The final R-factors are 11.2% (tetragonal) and 11.8% (trigonal).

UV spectra of the complexes between various intercalators and cyclic dinucleotides and d(CGTACG) were recorded on a Hewlett-Packard 8452A diode array UV spectrophotometer at room temperature. The concentrations of the intercalators were about 30 μM in the buffered solution (100 nM cacodylate, pH 7.0). The high salt experiments were done in the same buffer except that 1.0 NaCl was added.

Correspondence address: A.H.-J. Wang, Department of Physiology and Biophysics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

3. RESULTS AND DISCUSSION

The two independent c-(GpGp) molecules in both crystal forms adopt very similar conformations with a rigid 12-membered ribose-phosphate ring which holds the two guanines in parallel planes 6.8 Å apart (Fig.1). The root mean square (rms) deviation of a least-squares fit of the backbones of the two molecules in 0.096 Å (tetragonal) and 0.178 Å (trigonal). Each molecule posseses non-crystallographic molecular two-fold symmetry. All 8 riboses are in the C3′ endo conformation with pseudorotation angles ranging from −7° to 15°. The two independent molecules are associated together into a dimeric aggregate, each using one of its two guanines to intercalate between the two guanines of the other molecule (Fig. 1A and B).

In the tetragonal form, the mutually self-intercalated dimer of c-(GpGp) is held together by 4 hydrogen bonds (Fig. 1A). In addition, a cobalt ion is situated between the middle two guanines (G1A and G2A) to form a bidentate coordination to the adjacent N7 atoms, inducing a large dihedral angle (32°) between G1A and G2A bases. This may have important implications in the role of transition metal ions, such as cobalt ion, in biological systems as discussed later. In the trigonal form, the mutually self-intercalated dimer of c-(GpGp) is also held together by 4 hydrogen bonds (Fig. 1B).

The two c-(pGpG) dimers and another self-intercalated dimer of c-(dApdAp), whose structure was previously determined [2], are quite similar despite the fact that they differ in the ways that they associate. All 3 molecules have C3′ endo sugar conformations in their backbone, even though one is a DNA fragment while the other two are RNA fragments. This suggests that the sugar conformation is dictated by the cyclic nature of the dinucleotide. The glycosyl torsion angles of all nucleotides adopt an *anti* conformation (χ ranges from −153.3° to 175.7°). The dinucleotides associate into a dimer maximizing the possible number of hydrogen bonds and stacking interactions between the two molecules. The existence of the dimer structure of c-(dApdAp) in solution has been demonstrated by NMR studies [6].

The recurrence of the formation of self-intercalated dimers in the cyclic dinucleotides reinforces our contention that the cyclic dinucleotides may bind other planar aromatic compounds (e.g. intercalators) between the two bases. We have tested this hypothesis by examining the change of the absorption spectra of various intercalators effected by cyclic dinucleotides. Several cyclic dinucleotides (including c-(GpGp), c-(dGpGp), c-(XpXp), c-(GpXp) and c-(dApdTp); X is xanthosine) and different intercalators (including 9-aminoacridine (9-AA), proflavine, ethidium and daunomycin) were used in these experiments. Under the experimental conditions, only c-(GpGp) and c-(dGpGp) showed significant effects on the absorption spectra of simple intercalators. Fig.2 compares the spectra of the complex between intercalators (9-AA and ethidium) and c-(GpGp) versus those between the same intercalators and d(CGTACG), which is used to represent a typical double helical DNA. In the case of the interaction between 9-AA and c-(GpGp), the resulting spectrum (Fig.2A) is remarkably similar to that of the 9-AA/d(CGTACG) system (Fig. 2B). This suggests that in the former case, 9-AA resides in an environment very similar to that of an intercalation site in a DNA double helix. In contrast, c-(GpGp) has a lesser effect than does d(CGTACG) on the spectra of ethidium (Fig. 2C and D, respectively).

These findings raise several interesting questions. Why is it that only c-(GpGp) and c-(dGpGp) have strong effects, but not any of the other cyclic dinucleotides investigated? When compared to a duplex oligodeoxynucleotide, why do these two cyclic dinucleotides have a more pronounced effect on smaller planar compounds such as 9-AA and proflavine (data not shown) than on other larger compounds like ethidium and daunomycin? We suggest that these observations are related to the unique bis-intercalator-like conformation that the cyclic dinucleotide can adopt and the unique associative properties of guanine bases as explained below.

Poly(G) is known to aggregate into a large structure, a 4-stranded polymer in which 4 guanine bases are hydrogen-bonded together in a head-to-tail manner with a 4-fold symmetry [7]. This self-associated 4-stranded structure involving stretches of guanines has been suggested to play an important role in chromosomes [8–10]. We propose a model (Fig. 3) in which 4 c-(GpGp) molecules utilize the similar 4-fold symmetrical guanine-guanine base pairings to explain the unusual spectroscopic phenomena. It is interesting to note that a similar but not identical type of tetrameric aggregate of small guanine-containing oligonucleotides has recently been seen in the crystal structure of the adduct of cisplatin-d(pGpG) [11].

This model is consistent with the observed data. First, only cyclic dinucleotides with *two* guanines can form this circular tetrameric aggregate because of the hydrogen bonding requirements. Second, the internal space created by this cage-like structure can accommodate planar compounds like 9-AA (shown in Fig. 3) and proflavine easily. Any bulky side chain on the aromatic ring, however, will inhibit entry into the cavity of the cage structure, consistent with the fact that daunomycin shows no appreciable spectral change with c(GpGp). Third, the equilibrium between free 9-AA and 9-AA bound in the cavity of the cage structure requires the dissociation-association of the tetramer complex; this is expected to be a slow process as it involves the breakage and formation of many hydrogen bonds. We have carried out a time-dependent study of the binding of 9-AA with c-(GpGp) (data shown). The absorbance of the mixtures (3 different ratios of c-(GpGp) to

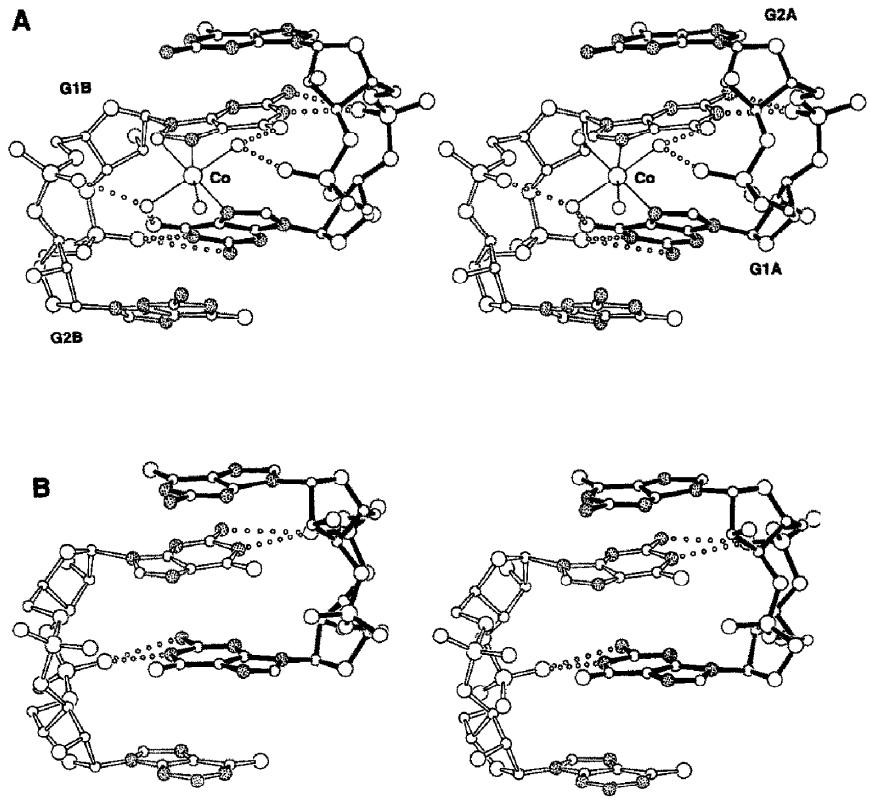

Fig. 1. Stereoscopic diagrams of the self-intercalated c-(pGpG) dinucleotides. The molecules with filled bonds are denoted molecule A and the molecules with open bonds are denoted molecule B. (A) Tetragonal form of c-(GpGp). A cobalt ion crosslinks and destacks two adjacent guanines (dihedral angle 32°) using bidentate coordination to the two N7 atoms. (B) Trigonal form of c-(GpGp).

9-AA) decreases with a kinetic rate that can be best fitted with two rate constants; one is fast (in the 1–3 min range), and the other is slower (in the 2 h range). Finally, the tetramer aggregate is highly negatively charged; thus, one would expect that salt has a strong effect on aggregate formation. When we compare the spectra of the 9-AA/c-(GpGp) complex in low (0.1 M sodium cacodylate buffer) and high salt (same buffer plus 1.0 M NaCl) solution, it is evident that the absorbance is further decreased in high salt solution (data not shown). From these arguments, we believe that the tetrameric cage model of c-(GpGp) molecules is a plausible one for its structure in solution.

Are the conformations seen in the present crystal structure and the unusual spectral properties of c-(GpGp) relevant to its biological function as the endogenous effector for cellulose synthase? We are not certain at this time. However, it is conceivable that the effector binding site in cellulose synthase has hydrophilic amino acids that can form specific hydrogen bonds with guanines and aromatic amino acids that can intercalate between two guanines, similar to the way we have seen in the crystal structure. The mimicry of an intercalator DNA binding pocket by c-(GpGp) is an even greater mystery. Is it possible that c-(GpGp) acts like a 'carrier' or 'host' for certain small aromatic compounds ('guests') in cells? Whether this property plays any role in the biological function of cellulose synthase or other proteins remains to be explored.

The observation of a crosslinking cobalt ion merits some comments. Many transition metal ions such as cobalt(II), copper(II) are known to coordinate to nucleic acid bases, in particular the N7 atom of guanine [12]. Recently, it has been shown that Co can cause substantial DNA bending as judged by electron microscopy studies [13]. This may be related to the fact that cobalt ion can crosslink two adjacent guanines and induce a significant dihedral angle between them, as observed in the present structure. Cisplatin forms a stable adduct with DNA by forming two Pt-N7 bidentate bonds to two consecutive guanines on the same strand. X-Ray structure of *cis*-[$Pt(NH_3)_2${d(pGpG)}] showed that the interplanar angle of the two guanines ranges between 75° and 90° [13]. However, gel electrophoresis experiments indicated that DNA bending induced by the cisplatin crosslink to d(GpG) is only about $40 \pm 5°$ [14], a value very close to the value seen here. Our results therefore suggest a possible mechanism to construct a model for the metal ion (Pt or Co) crosslinked GpG site in DNA double helix.

Finally, we have pointed out earlier that this type of cyclic dinucleotide molecules resembles other bis-intercalators such as echinomycin [15,16], except that it is suited for nearest-neighbor intercalation, i.e. to

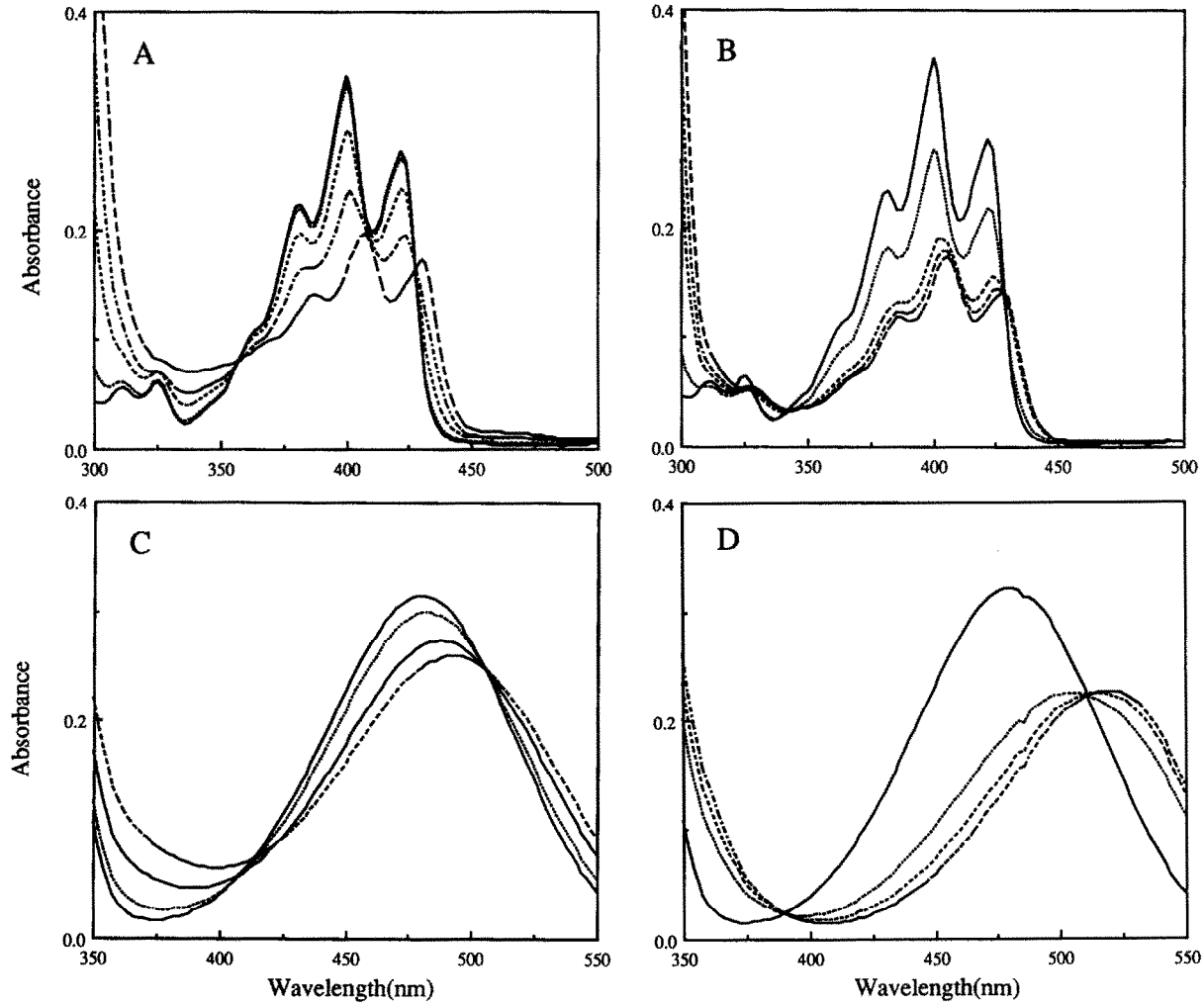

Fig. 2. UV absorption spectra of the solutions of the mixture of intercalators and oligonucleotides. (A) 9-AA + c-(GpGp). RNA-to-9-AA ratio (0:1, 1:1, 6:1, 10:1, 20:1) for the curves going from high to low absorbance at $\lambda = 400$ nm. (B) 9-AA + d(CGTACG). DNA duplex-to-9-AA ratio: (0:1, 0.7:1, 4:1, 7:1, 14:1). (C) Ethidium + c-(GpGp). RNA-to-ethidium ratio (0:1, 3:1, 10:1, 20:1) for the curves going from high to low absorbance at $\lambda = 500$ nm. (D) Ethidium + d(CGTACG). DNA duplex-to-ethidium ratio (0:1, 1:1, 3:1, 10:1). All measurements were performed at time 15 min after the mixing of RNA/DNA and intercalators to ensure the mixture solution had reached the equilibrium state. Note the similarity between (A) and (B); both show well-defined isosbestic points. The curves in (C) show some hypochromic red-shifts, but to a much lesser extent compared to those in (D), suggesting that the interactions between ethidium and c-(GpGp) and d(CGTACG) are not similar.

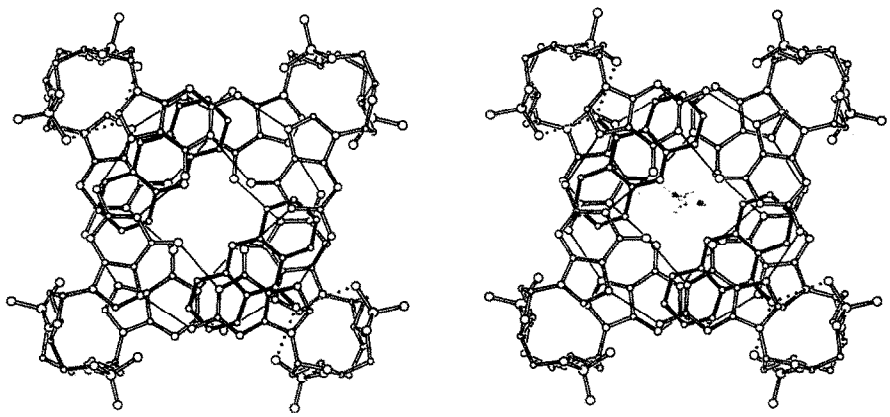

Fig. 3. A cage-like model constructed by taking 4 c-(GpGp) molecules with the conformation seen in the crystal and connecting them together using the hydrogen bonding scheme seen in the 4-stranded poly(G) structure [7]. A 4-fold symmetry is imposed and refined to remove any possible bad contacts. It is surprising that very little adjustment in the c-(GpGp) conformation is required to form such a tetramer aggregate. This hypothetical structure is likely to be very stable, as it has a total of 16 intra-aggregate guanine-to-guanine hydrogen bonds. It can be further stabilized by the inclusion of aromatic compounds (two 9-AA molecules shown in the figure) to fill the cavity. It can be seen that no more than two 9-AA molecules can be fitted inside the hydrophobic cavity. It is possible that this structure can encapsulate a single larger aromatic compound such as benzpyrene.

bracket only one base pair. Our work here suggests that unmodified cyclic dinucleotides (e.g. c-(dApdAp)) are not adequate to serve such a purpose, as they do not bind intercalators very tightly (data not shown). This is likely due to the insufficient aromatic surface area for stacking interactions. By enlarging the aromatic rings in the molecule (e.g. changing adenine to ethenoadenine), it should be possible to improve the binding affinity of such cyclic dinucleotides toward intercalators. By further molecular engineering of this type of molecules (e.g. converting to a positively charged molecule), we should be able to synthesize some compounds that have the property of a nearest-neighbor bis-intercalator [17] with unique binding affinity and specificity toward DNA.

Acknowledgements: This work was supported by grants from NSF and NIH (A.H.-J. W.). G.A.vdM. and J.H.vB. were supported by the Netherlands Organization for the Advancement of Pure Research (ZWO).

REFERENCES

[1] Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., De Vroom, E., Van der Marel, G.A., Van Boom, J.H. and Benziman, M. (1987) Nature (Lond.) 325, 279–281.

[2] Frederick, C.A., Coll, M., Van der Marel, G.A., Van Boom, J.H. and Wang, A.H.-J. (1988) Biochemistry 27, 8350–8361.

[3] De Vroom, E., Broxterman, H.J.G., Sliedregt, L.A.J.M., Van der Marel, G.A. and Van Boom, J.H. (1988) Nucleic Acids Res. 16, 4607–4620.

[4] Sheldrick, G.M. (1989) Acta Crystallogr., submitted.

[5] Gabe, E.J., Lee, F.L. and Le Page, Y. in: Crystallographic Computing, vol. 3 (Sheldrick, G.M., Kruger, C. and Goddard, R. eds) Clarendon Press, Oxford, pp. 167–174.

[6] Blommers, M.J.J., Haasnoot, C.A.G., Walters, J.A.L.I., Van der Marel, G.A., Van Boom, J.H. and Hilbers, C.W. (1988) Biochemistry 27, 8361–8369.

[7] Arnott, S., Chandrasekaran, R. and Marttila, C.M. (1974) Biochem. J. 141, 537–543.

[8] Sen, D. and Gilbert, W. (1988) Nature (Lond.) 334, 364–366.

[9] Williamson, J.R., Raghuraman, M.K. and Cech, T.R. (1989) Cell 59, 871–880.

[10] Sundquist, W.I. and Klug, A. (1989) Nature 342, 825–829.

[11] Sherman, S.E., Gibson, D., Wang, A.H.-J. and Lippard, S.J. (1988) J. Am. Chem. Soc. 110, 7368–7381.

[12] Sigel, H. (1989) in: Metal-DNA Chemistry (Tullius, T.D. ed.) p. 159, Am. Chem. Society.

[13] Laundon, C.H. and Griffith, J.D. (1987) Biochemistry 26, 3759–3762.

[14] Rice, J.A., Crothers, D.M., Pinto, A.L. and Lippard, S.L. (1988) Proc. Natl. Acad. Sci. USA 85, 4158–4161.

[15] Wang, A.H.-J., Ughetto, G., Quigley, G.J., Hakoshima, T., Van der Marel, G.A., Van Boom, J.H. and Rich, A. (1984) Science 225, 1115–1121.

[16] Wakelin, L.P.G. (1986) Med. Res. Rev. 6, 275–340.

[17] Atwell, G.J., Stewart, G.M., Leupin, W. and Denny, W.A. (1985) J. Am. Chem. Soc. 107, 4335–4337.