Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of d(ACGTAGCTACGT)$_2$:[actinomycin D, (echinomycin)$_2$] and d(ACGTAGCTACGT)$_2$:[actinomycin D, (triostin A)$_2$] complexes

Hana L. Takusagawa† and Fusao Takusagawa*

Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045-2106, USA

† Presently at Harvard University.

Correspondence e-mail: xraymain@ukans.edu

Received 16 September 1999
Accepted 22 December 1999

A DNA–multiple drug complex, d(ACGTAGCTACGT)$_2$:[actinomycin D, (echinomycin)$_2$] has been crystallized. The crystals belong to the monoclinic space group $C2$, with unit-cell parameters a = 85.6, b = 72.8, c = 56.6 Å, β = 101.5° at 93 K and Z = 8. The crystal diffracted to 3.0 Å resolution along the DNA fiber axis and to 3.5 Å resolution in other directions. The Patterson maps indicate that all complexes in the crystal are oriented along their helical axes in the [10$\bar{1}$] direction.

1. Introduction

The human genome project (HGP) will be completed before 2003. It is expected that abnormal genes related to human diseases will be discovered by referring to the HGP database. A therapy which could restore these abnormal genes to normal genes would be one of the ultimate goals of biomedical research. However, this would be very difficult. An alternative is to suppress the abnormal genes. The replication and/or transcription of specific genes can be inhibited in either non-DNA-directed or DNA-directed manners. The non-DNA-directed manner involves inhibiting the activities of the specific proteins that are participating in the abnormal DNA replication and/or transcription. To accomplish this, it is necessary to know what proteins control the replication or transcription of the specific genes and how to inhibit them specifically. The DNA-directed manner involves disrupting the DNA and/or RNA polymerase activities with small molecules that bind to the specific sequence of abnormal DNA. To accomplish this, it is necessary to have small molecules that bind tightly to the specific DNA sequences only observed in the abnormal genes.

Our goal is to develop compounds that recognize DNA sequences of longer than ten base pairs. While protein variability is achieved *via* 20 different amino acids, DNA has only four different bases. This means that it takes a span of about 16 bases to define a unique DNA sequence in humans *via* base specificity, *i.e.* $4^{16} = 4.3 \times 10^9$, which is slightly larger than the size of human DNA, $\sim 2.9 \times 10^9$ base pairs. The main molecules that are able to read DNA with high specificity are large proteins (Choo & Klug, 1997) and the complementary strands of DNA or RNA. Smaller molecules with less complex shapes are able to interact with DNA, but these cannot recognize relatively long DNA sequences, *i.e.* they are much less specific.

Significant progress has been made with the polyamide minor-groove binders (reviewed by Walker *et al.*, 1997) and current compounds can discriminate among all four Watson–Crick base pairs in the minor groove, *i.e.* py–Hy, Hy–py, Im–py and py–Im ring pairs recognize A–T, T–A, G–C and C–G base pairs, respectively (Kielkopf, Baird *et al.*, 1998; Keilkopf, White *et al.*, 1998). However, polyamide–DNA structures reveal a slight mismatch between the geometry of the DNA minor groove and the geometry of the polyamide, *i.e.* the amide-pyrrole subunit is slightly longer than the base-to-base repeat length along the minor groove of DNA. For longer polyamides, therefore, the amide groups become out of phase with the DNA bases. This phasing incompatibility limits the number of repeating units of polyamide to about four. To extend the sequence specificity, two or more binding agents were connected with tethers (for example, Khorlin *et al.*, 1980; Gursky *et al.*, 1983; Youngquist & Dervan, 1985; Lown *et al.*, 1986, 1989; Youngquist & Dervan, 1987; Grokhovsky & Zubarev, 1990; Singh *et al.*, 1994; Mrksich *et al.*, 1994; Nikolaev *et al.*, 1996; Parks *et al.*, 1996; White *et al.*, 1997; Grokhovsky *et al.*, 1998). Although most of these bis- and tri-linked polyamides did not bind to the specific DNA sequences, it has been reported that a few bis-linked netropsin derivatives inhibit selectively activities of topoisomerases I and II (Beerman *et al.*, 1991; Burckhardt *et al.*, 1997) and HIV-I reverse transcriptase (Filipowsky *et al.*, 1996).

We have postulated that long sequence-specific DNA binders can be designed from the crystal structure of a DNA:(INT, AMD, INT) complex, where INT and AMD are an intercalator and actinomycin D, respectively. If the

three DNA intercalators in a DNA complex are connected to each other without changing their structures, relative orientations and positions, the resultant hybrid molecule will bind tightly to a specific site on the DNA minor groove and will cover more than ten base pairs. As of May 15, 1999, no coordinates of DNA complexed with more than two different drugs had been deposited in the PDB and, as far as we are aware, no crystal structures of DNA containing more than two different drugs have been published. As the first step in testing our hypothesis, we crystallized DNA:(AMD, ECM_2) and DNA:(AMD, TOA_2) complexes for an X-ray diffraction study. Actinomycin D (AMD; Fig. 1) was selected as the central intercalator for the following reasons. (i) The fifth amino-acid residues (N-MeVal) of the depsipentapeptide rings of AMD are excellent candidates to which to attach another antibiotic (Takusagawa *et al.*, 1996). (ii) The crystal structures of DNA–AMD complexes indicate that the depsipentapeptide rings are mobile around the N—C bond of Thr and are also relatively flexible (Kamitori & Takusagawa, 1994). The mobile and flexible rings can absorb the small distortion gap generated by connecting AMD to the other intercalators. Triostin A (TOA; Fig. 1) and echinomycin (ECM; Fig. 1) were selected as the other intercalators, since their structures and DNA-binding characteristics have been extensively studied. TOA contains two planar aromatic quinoxaline rings which are covalently attached to a cyclic depsiocta-peptide ring. The two Cys residues form a disulfide bond. ECM has a chemical structure similar to TOA, except it contains a thioacetal bridge instead of a disulfide bridge. Both antibiotics bind sequence specifically to DNA as bis-intercalators around the CG step and the peptide rings of both drugs bind in the minor groove of DNA (Wang *et al.*, 1984, 1986; Ughetto *et al.*, 1985; Quigley *et al.*, 1986; Addess & Feigon, 1994).

Chaires and co-workers realised that a 2:1 stoichiometry of drug to DNA duplex was observed in all of the structures reported for non-covalent complexes of anthracyclines bound to DNA (Chaires *et al.*, 1987). They linked the two 3′-NH_2 groups from the daunosamine moieties without disturbing the DNA–drug interactions observed in the crystal structures. WP631, a newly synthesized bisanthracycline, has exhibited very promising anticancer activities (Chaires *et al.*, 1987). This is a good example of the importance of crystal structures, which provide not only the detailed information of drug–DNA interactions, but also the relative orientation of drugs in DNA–drug complexes. The crystal structure of $d(CGTACG)_2$–WP631 confirmed the binding mode of this new antibiotic (Hu *et al.*, 1997). On the other hand, without having crystal structures, two or more DNA binding agents have been connected to improve their biological properties. For example, Krivtsova *et al.* (1984) and Dervan (1986) prepared netropsin and distamycin analogues attached to the phenoxazone ring of AMD. Krivtsova and co-workers attached from one to three pyrrolecarboxamides of netropsin directly to the 1,9-position of the phenoxazone chromophore of AMD. Dervan linked the distamycin *via* a glycine tether to the phenoxazone chromophore of AMD and also replaced the two cyclic pentapeptides of AMD with the tripeptide of distamycin, preparing a bis(distamycin)-phenoxazone. Footprints of Dervan's bis(distamycin)phenoxazone revealed a 20 base-pair $(AT)_4(GC)_2(TA)_4$ site consistent with a groove binder–intercalator–groove binder mode of recognition. However, two other footprint sites indicated incomplete coverage by the entire molecule, probably with only one distamycin binding or distamycin–phenoxazone binding. To our knowledge, these hybrid binding compounds have not shown any promising biological activities *in vivo*.

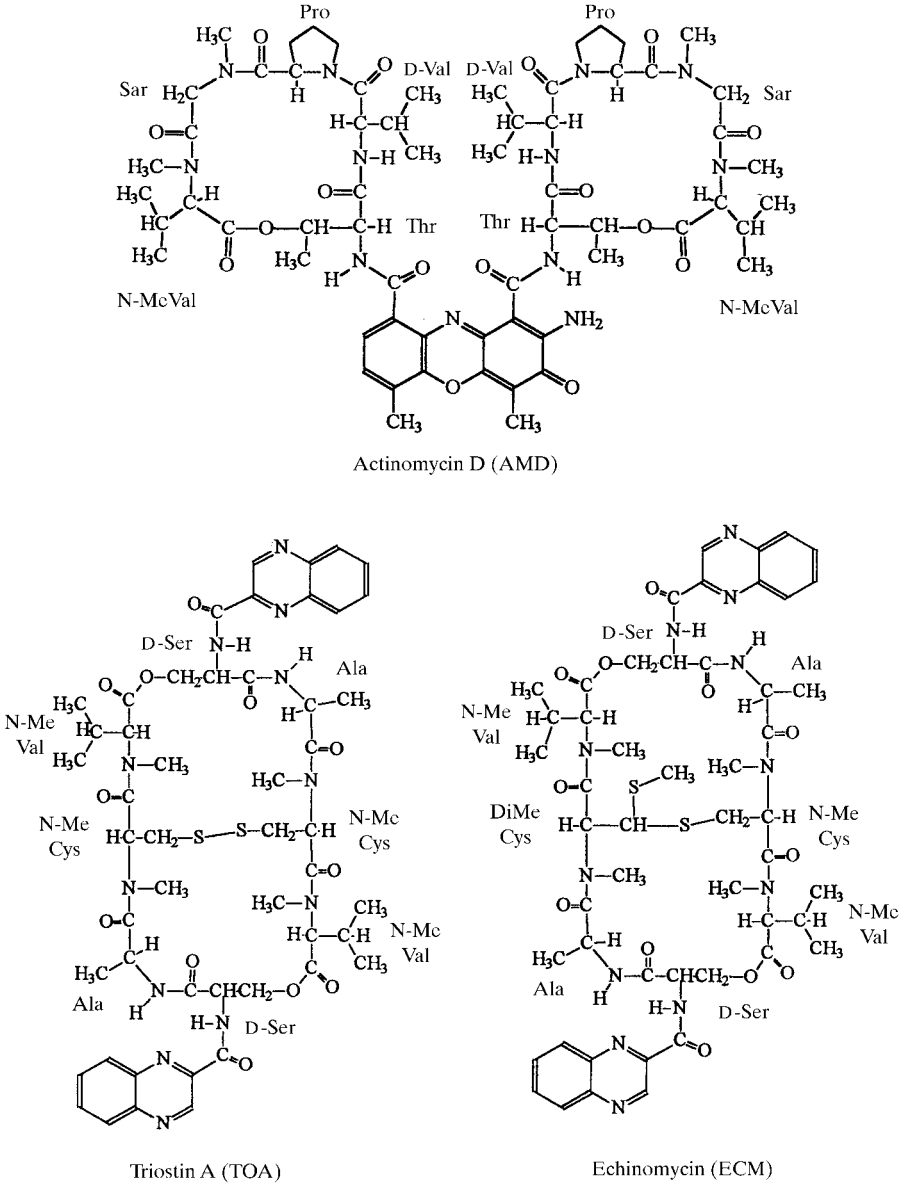

Figure 1
Chemical formulae of actinomycin D, triostin A and echinomycin.

Table 1
Crystallographic data of two DNA–multiple drug complexes.

	DNA:[AMD, $(ECM)_2$]†	DNA:[AMD, $(TOA)_2$]†
Space group	*C*2	*C*2
Unit-cell parameters (Å, °)	a = 85.63, b = 72.77, c = 56.58, β = 101.5	a = 85.33, b = 72.98, c = 55.90, β = 101.2
No. of subunits in asymmetric unit	8	8
Resolution (Å)	8.0–3.0	8.0–3.0
No. of reflections measured	16 228	15 773
No. of independent reflections used	6068 (all data)	5603 (all data)
Completeness (%)	81.6	75.3
R_{merge}‡	0.068	0.071

† DNA = d(ACGTAGCTACGT)$_2$. ‡ $R_{merge} = \sum |I - \langle I \rangle| \sum |I| = 0.068$.

2. Materials and methods

2.1. Crystallization

A self-complemental 12-mer DNA, A**CG**T-A**GC**T-A**CG**T, which has an AMD binding site (GC sequence) in the middle and TOA/ECM binding sites (CG sequences) on both ends was synthesized by a DNA synthesizer. The crude DNA was purified by HPLC using a reverse-phase C_{18} column. The AMD and ECM were purchased from Sigma; the TOA was a gift from Shionogi Pharmaceutical Company, Japan. The concentrations of DNA and drugs were determined from the absorbance at 260 nm with the extinction coefficient $2.932 \times 10^5\ M^{-1}\ cm^{-1}$ for the DNA, at 440 nm with an extinction coefficient of $2.45 \times 10^4\ M^{-1}\ cm^{-1}$ for AMD and at 325 nm with an extinction coefficient of $1.15 \times 10^4\ M^{-1}\ cm^{-1}$ for TOA and ECM (Low *et al.*, 1984). The following stock solutions were prepared: (i) 2.5 m*M* DNA (double stranded) in water, (ii) 5 m*M* AMD in water, (iii) 5 m*M* TOA/ECM in 1:1 methanol and chloroform, (iv) 100 m*M* spermine tetrachloride in water, (v) 100 m*M* $MgCl_2$ in water, (vi) 100 m*M* cacodylate/HCl pH 7.0. These solutions were mixed in a 1:1:2:2:5:10 mole ratio. The mixtures were evaporated in a Speedvac and dissolved in water. This process was repeated several times. DNA–drug solutions containing 2 m*M* DNA, 2 m*M* AMD, 4 m*M* TOA/ECM, 4 m*M* spermine, 10 m*M* $MgCl_2$ and 20 m*M* cacodylate (pH 7.0) were prepared for crystallization screening.

The crystallization conditions were initially screened with the 48 screening solutions formulated by Hampton Research. However, no crystals grew. Therefore, new crystallization conditions were investigated using 2-methyl-2,4-pentanediol (MPD) and PEG 400. Relatively large crystals of $(ACGTAGCTACGT)_2$:[AMD, $(TOA)_2$] (~0.2 × 0.1 × 0.1 mm) and $(ACGTAGCTACGT)_2$:[AMD, $(ECM)_2$] (~0.5 × 0.2 × 0.1 mm) were grown in a solution containing 2 m*M* DNA, 2 m*M* AMD, 4 m*M* TOA/ECM, 4 m*M* spermine, 10 m*M* $MgCl_2$, 20 m*M* cacodylate (pH 7.0) and 10% MPD at 277 K (Fig. 2). In order to confirm the content of these crystals, several were scooped out from the hanging drop, washed in a 15% MPD solution and dissolved in a UV-suitable cuvette with 20 m*M* cacodylate buffer (pH 7). The UV absorption spectrum of the crystal solution was measured from 500 to 240 nm. The absorption spectra matched the spectra of the DNA–drug solution well, indicating that the crystals contain the DNA:[AMD, $(TOA)_2$] and the DNA:[AMD, $(ECM)_2$] complexes, respectively.

2.2. X-ray diffraction study of DNA:[AMD, $(TOA)_2$] and DNA:[AMD, $(ECM)_2$] complexes

A crystal (~0.3 × 0.1 × 0.1 mm) of DNA:[AMD, $(ECM)_2$] was scooped out of a hanging drop using a nylon loop and dipped into a cryoprotectant solution containing 20% glycerol and 10% MPD for 15 s. Immediately afterwards, the crystal was flash-frozen in cold nitrogen gas (93 K) on a

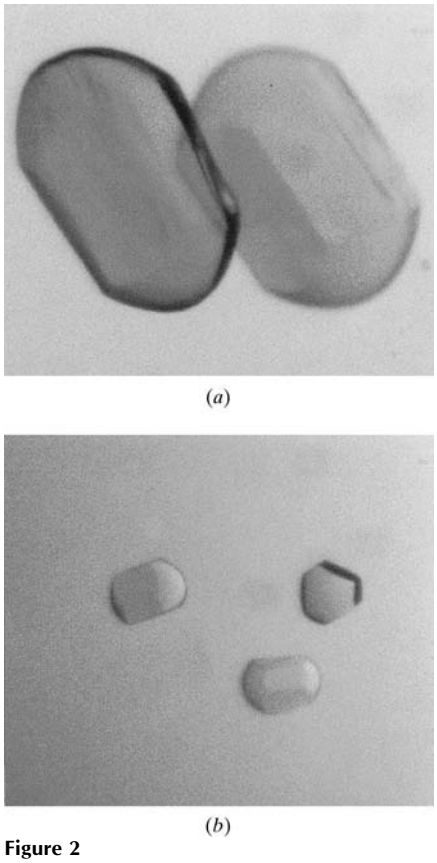

Figure 2
Crystals of DNA:[AMD, $(INT)_2$]. (*a*) DNA:[AMD, $(ECM)_2$] (*b*) DNA:[AMD, $(TOA)_2$].

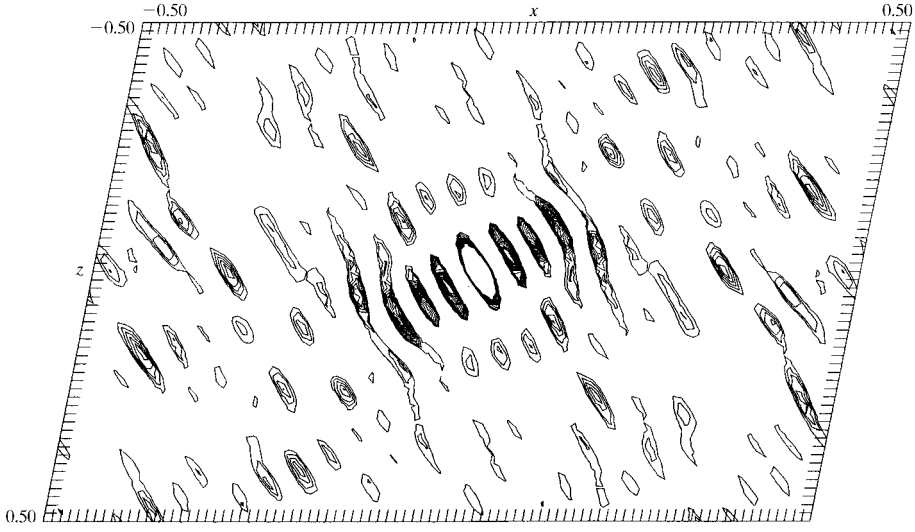

Figure 3
V = 0 section of the Patterson map of DNA:[AMD, $(ECM)_2$].

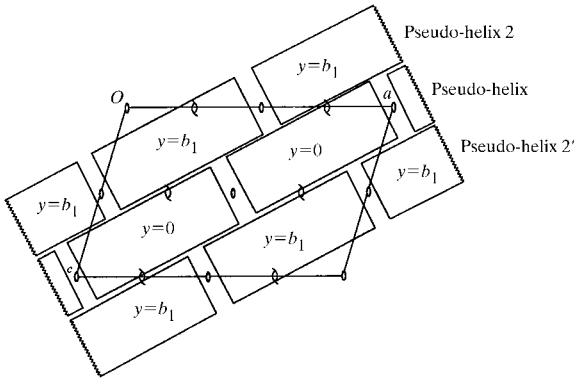

Figure 4
Complex packing in the crystal structure of DNA:[AMD, $(ECM)_2$] deduced from Patterson map. Each rectangle represents a DNA:[AMD, $(ECM)_2$] complex.

Rigaku R-AXIS IIc imaging-plate X-ray diffractometer with a rotating-anode X-ray generator (Cu $K\alpha$ radiation, operated at 50 kV and 100 mA). The still photographs showed a typical DNA diffraction pattern and indicated that the crystal diffracted to 3.0 Å resolution along the DNA fiber axis and to 3.5 Å resolution in other directions. The diffraction data were measured to 3.0 Å resolution. The data were processed with the program *DENZO* (Otwinowski & Minor, 1997). The crystallographic data are listed in Table 1.

The diffraction data of DNA:[AMD, $(TOA)_2$] were measured using the same procedures as described above; the crystallographic data are listed in Table 1. The unit-cell parameters and space group indicate that the crystal of DNA:[AMD, $(TOA)_2$] is isomorphous to the crystal of DNA:[AMD, $(ECM)_2$]. Indeed, as described below, the Patterson maps of DNA:[AMD, $(ECM)_2$] and DNA:[AMD, $(TOA)_2$] are quite similar. Since the DNA:[AMD, $(ECM)_2$] complex has a slightly larger number of independent reflections, the preliminary analysis of this complex will be described below.

2.3. Patterson map interpretation

The calculated unit-cell volume of 345 488 $Å^3$ suggests that there are two DNA–drug complexes in the asymmetric unit, with a V_m of 4.1 $Å^3\,Da^{-1}$, corresponding to a solvent content of 70% (Matthews, 1968). Although the V_m value is higher than those of protein structures, a similar V_m value (4.0 $Å^3\,Da^{-1}$) is found in the DNA–N8AMD and DNA–F8AMD crystal structures (Shinomiya *et al.*, 1995; Takusagawa *et al.*, 1997).

The Patterson map calculated with 10–3 Å resolution data clearly shows the stacking direction of the planar bases (helical axis direction) along the $[10\bar{1}]$ direction (Fig. 3). The map indicates that all DNA–drug complexes in the crystal are oriented along their helical axis in the $[10\bar{1}]$ direction. The base pairs and chromophores are nearly parallel to each until the fourth base pair, after which they are no longer parallel.

It is noted that the diagonal distance of the unit cell along the $[10\bar{1}]$ direction is 112 Å ($= 85.63^2 + 56.58^2 - 2 \times 85.63 \times 56.58 \times \cos 101.5^\circ$). The length of the complex along the helical axis is 56 Å [= 3.3 Å × (12 + 1 + 4)], where 3.3 Å is the base pair/base pair or the base pair/chromophore spacing and 12, 1 and 4 are the numbers of base pairs in the 12-mer DNA, chromophores of AMD and chromophores of the two ECMs, respectively. The complex length and the unit-cell parameters indicate that the two complexes are stacked upon each other along the $[10\bar{1}]$ direction. Thus, the complexes form a pseudo-continuous helix as observed in many DNA structures (Fig. 4). This characteristic feature of the Patterson map and the unit-cell parameters also indicates that the complex is an dodecamer duplex with one AMD and two ECMs intercalated into it. The second pseudo-helix passes at $y \simeq 0.15$; the $y = 0.15$ section of the Patterson map is relatively clouded.

References

Addess, K. J. & Feigon, J. (1994). *Biochemistry*, **33**, 12386–12396.
Beerman, T. A., Woynarowski, J. M., Sigmund, R. D., Gawron, L. S., Rao, K. E. & Lown, J. W. (1991). *Biochem. Biophys. Acta*, **1090**, 52–60.
Burckhardt, G., Simon, H., Storl, K., Triebel, H., Walter, A., Lown, J. W. & Zimmer, Ch. (1997). *J. Biomol. Struct. Dyn.* **15**, 81–95.
Chaires, J. B., Fox, K. R. F., Herrera, J. E., Britt, M. & Waring, M. J. (1987). *Biochemistry*, **26**, 8227–8236.
Choo, Y. & Klug, A. (1997). *Curr. Opin. Struct. Biol.* **7**, 117–125.
Dervan, P. B. (1986). *Science*, **232**, 464–471.
Filipowsky, M. E., Kopeka, M. L., Brazil-Zison, M., Lown, J. W. & Dickerson, R. E. (1996). *Biochemistry*, **35**, 15397–15410.
Grokhovsky, S. L., Surovaya, A. N., Burckhardt, G., Pismensky, V. F., Chernov, B. K., Zimmer, Ch. & Gursky, G. V. (1998). *FEBS Lett.* **439**, 346–350.
Grokhovsky, S. L. & Zubarev, V. E. (1990). *Nucleic Acids Res.* **19**, 257–264.
Gursky, G. V., Zasedatelev, A. S., Zhuze, A. L., Khorlin, A. A., Grokhovsky, S. L., Streltsov, S. A., Surovaya, A. N., Nikitin, S. M., Krylov, A. S., Retchinsky, V. O., Mikhailov, M. V., Beabealaschvilli, R. Sh. & Gottikh, B. P. (1983). *Cold Spring Habor Symp. Quant. Biol.* **47**, 367–378.
Hu, G. G., Shui, X., Leng, F., Priebe, W., Chaires, J. B. & Williams, L. D. (1997). *Biochemistry*, **36**, 5940–5946.
Kamitori, S. & Takusagawa, F. (1994). *J. Am. Chem. Soc.* **116**, 4154–4165.
Khorlin, A. A., Krylov, A. S., Grokhovsky, S. L., Zhuze, A. L., Zasedatelev, A. S., Gursky, G. V. & Gottikh, B. P. (1980). *FEBS Lett.* **118**, 311–314.
Kielkopf, C. L., Baird, E. E., Dervan, P. B. & Rees, D. C. (1998). *Nature Struct. Biol.* **5**, 104–109.
Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B. & Rees, D. C. (1998). *Science*, **282**, 111–115.
Krivtsova, M. A., Moroshkina, E. B., Glibin, E. N. & Frisman, E. V. (1984). *Mol. Biol. (USSR)*, **18**, 950–956.
Low, C. M. L., Olsen, R. K. & Waring, M. J. (1984). *FEBS Lett.* **176**, 414–419.
Lown, J. W., Krowicki, K., Balzarini, J., Newman, R. A. & De Clerk, E. (1989). *J. Med. Chem.* **32**, 2368–2375.
Lown, J. W., Krowicki, K., Bhat, G. U., Skorobogaty, A., Ward, B. & Dabrowiak, J. C. (1986). *Biochemistry*, **25**, 7408–7416.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mrksich, M., Parks, M. E. & Dervan, P. B. (1994). *J. Am. Chem. Soc.* **116**, 7983–7988.
Nikolaev, V. A., Grokhovsky, S. L., Surovaya, A. N., Leinsoo, T. A., Sidorova, N. Yu., Zasedatelev, A. S., Zhuze, A. L., Strachan, G. A., Shafer, R. H. & Gursky, G. V. (1996). *J. Biomol. Struct. Dyn.* **14**, 31–47.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Parks, M. E., Baird, E. E. & Dervan, P. B. (1996). *J. Am. Chem. Soc.* **118**, 6141–6146.
Quigley, G. J., Ughetto, G. A., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J. & Rich, A. (1986). *Science*, **232**, 1255–1258.
Shinomiya, M., Chu, W., Carlson, R. G., Weaver, R. F. & Takusagawa, F. (1995). *Biochemistry*, **34**, 8481–8491.
Singh, M. P., Plouvier, B., Hill, G. C., Gueck, J., Pon, R. T. & Lown, J. W. (1994). *J. Am. Chem. Soc.* **116**, 7006–7020.
Takusagawa, F., Takusagawa, K. T. & Weaver, R. F. (1997). *Bioorg. Med. Chem.* **5**, 1197–1207.
Takusagawa, F., Wen, L., Li, Q., Chu, W., Takusagawa, K. T., Carlson, R. G. & Weaver, R. F. (1996). *Biochemistry*, **35**, 13240–13249.
Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1985). *Nucleic Acids Res.* **13**, 2305–2323.
Walker, W. L., Kopka, M. L. & Goodsell, D. S. (1997). *Biopolymers*, **44**, 323–334.
Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H. & Rich, A. (1984). *Science*, **225**, 1115–1121.
Wang, A. H.-J., Ughetto, G., Quigley, G. J. & Rich, A. (1986). *J. Biomol. Struct. Dyn.* **4**, 319–342.
White, S., Baird, E. E. & Dervan, P. B. (1997). *Chem. Biol.* **4**, 569–578.
Youngquist, R. S. & Dervan, P. B. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 2565–2569.
Youngquist, R. S. & Dervan, P. B. (1987). *J. Am. Chem. Soc.* **109**, 7564–7566.