

Crystal structure of a DNA decamer showing a novel pseudo four-way helix–helix junction

NEIL SPINK*[†], CHRISTINE M. NUNN*, JAROSLAV VOJTECHOVSKY[‡], HELEN M. BERMAN[‡], AND STEPHEN NEIDLE*[§]

*The Cancer Research Campaign Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey, SM2 5NG, United Kingdom; and [†]Department of Chemistry, Rutgers University, New Brunswick, NJ 08903

Communicated by Alexander Rich, Massachusetts Institute of Technology, Cambridge, MA, July 24, 1995

ABSTRACT The crystal structure of the decanucleotide d(CGCAATTGCG)₂ has been solved by a combination of molecular replacement and heavy-atom procedures and has been refined to an *R* factor of 20.2% at 2.7 Å. It is not a fully base-paired duplex but has a central core of eight Watson–Crick base pairs flanked by unpaired terminal guanines and cytosines. These participate in hydrogen-bonding arrangements with adjacent decamer duplexes in the crystal lattice. The unpaired guanines are bound in the G+C regions of duplex minor grooves. The cytosines have relatively high mobility, even though they are constrained to be in one region where they are involved in base-paired triplets with G·C base pairs. The 5'-AATT sequence in the duplex region has a narrow minor groove, providing further confirmation of the sequence-dependent nature of groove width.

The crystal structures of many B-DNA oligonucleotide duplexes have now been determined (1–3). These can be divided into two principal classes: (i) duplexes (the overwhelming majority) with Watson–Crick hydrogen-bonding arrangements for G·C and A·T base pairs and (ii) duplexes with base-pair mismatches such as A·G, G·G, and G·T. Such mismatches are invariably accommodated within double-helical structures (4, 5). Fully base-paired helices have also been observed to be the norm in protein–DNA complexes, even when the DNA is highly perturbed from a standard B form. The sole exceptions reported to date are (i) DNA structures with a looped-out region by virtue of insertion of nonpairing bases (6, 7) and (ii) the structure of the *Hha* I methyltransferase–d(TGATAGf¹⁵C-GCTATC), where the 5-fluorocytosine base is swung out of the DNA helix into the enzyme active site (8). A Z-DNA hexamer crystal structure has been reported (9) in which the first cytosine of each duplex is swung out, forming a Watson–Crick base pair with a guanine in a symmetry-related duplex.

The first crystal structure of a B-DNA oligonucleotide to be determined (10) was that of the self-complementary dodecanucleotide d(CGCGAATTCGCG)₂. This showed features of sequence-dependent structure such as a narrow minor groove in the 5'-AATT region and high propeller twist for the A·T base pairs. These have been observed in other dodecanucleotide crystal structures (11), such as that of d(CGTGAATTCACG)₂ (12, 13), and are unlikely to be due to the influence of crystal packing forces. The crystal structures of both of the decamer duplex sequences d(CCATAATGG)₂ (14) and d(CGATTAATCG)₂ (15) show narrow minor grooves and high propeller twists for most of the A·T base pairs; however, the alternating sequence d(CGATATATCG)₂ (16) does not have these features. We have determined the crystal structure of the self-complementary decamer d(CGCAATTGCG)₂, which has close sequence analogy to the original dodecanucleotide structure (10). This provides an independent structural assessment of the sequence-dependent features of short runs of A·T base

pairs. The decamer has B-DNA duplex geometry for its central eight base pairs, but with an unprecedented arrangement of unpaired bases at the 5' and 3' termini. Instead of being within the central helix of the duplex, these nonpaired bases interact with symmetry-related decamer molecules to form an unusual quasi four-stranded arrangement at the junctions between them.

METHODS

The self-complementary oligonucleotide d(CGCAATTGCG)₂ was purchased from the Oswell/DNA Service (Edinburgh). Yellow rhombic crystals were grown by vapor diffusion at 287 K over a period of 4–6 weeks from aqueous solution containing 30 mM sodium cacodylate buffer at pH 7.0, 2.0 M MgCl₂ (1 μl), 49% 2-methyl-2,4-pentanediol (MPD) (3 μl), 3 mM decamer (4 μl), and 15 mM minor-groove-binding drug berenil (2 μl), equilibrated against a 50% MPD reservoir. Intensity data were collected on a Siemens-Xentronics X-1000A multiwire area detector from one crystal (0.3 × 0.4 × 0.6 mm³) at 294 K with graphite-monochromated CuK_α radiation from a rotating anode x-ray generator (40 kV, 70 mA). Unit cell dimensions are *a* = 27.01 Å, *b* = 39.26 Å, and *c* = 54.01 Å in the space group *I*2₁2₁, with a single decamer strand in the asymmetric unit, as established by the subsequent analysis. Measurements of 10,180 reflections to 2.2 Å were processed with the XENGEN package and averaged to 1521 unique reflections [*R*_{sym}(*I*) = 10.2% for reflections with *I* > 0]. Analysis of subsequent refinement statistics showed that the data were of high quality [*R*_{sym}(*I*) = 5.1% for reflections with *I* > 0; 800 unique reflections] to 2.7-Å resolution only; 96% of the outermost shell of reflections were observed to this resolution.

An initial search model consisting of a decamer single strand in a B-DNA conformation yielded an acceptable structure using molecular replacement methods with the program package MERLOT (17). The validity of this model was corroborated by considerations of the restrictions placed on the formation of a base-paired duplex in this space group. This also ruled out the space group *I*222 since it resulted in highly implausible packing arrangements with parallel rather than antiparallel duplexes. With use of the X-PLOR program (18), the starting model was refined as a rigid body to an *R* factor of 47% in the resolution range 8–4 Å. Five other less plausible models (with one strand incorrectly rotated with respect to the other) did not refine to an *R* value of <50%. However, a serious deficiency of the model at this stage was that the terminal bases of adjacent duplexes in the crystal lattice still interpenetrated each other. The model was accordingly broken into various sets of nucleotide groups, each of which was refined with the X-PLOR program. The most successful refinement was one with the decamer split in the five groups C1/G2/C3A4A5T6T7G8/C9/G10. Resultant electron density showed the C1 residue to be flipped out of the duplex

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

[†]Deceased July 16, 1995.

[§]To whom reprint requests should be addressed at: Cancer Research Campaign Biomolecular Structure Unit, The Institute of Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, United Kingdom.

and coplanar with the G2:C9 base pair of an adjacent duplex in the crystal lattice. Nucleoside G10 was also observed to be flipped out from the duplex and positioned close to the same adjacent duplex in the lattice. Further rounds of simulated annealing, followed by positional and B-factor refinement with X-PLOR and manual model building into electron density, reduced the *R* factor to 29% for data in the range 8–2.5 Å. Forty-nine water molecules were located in difference electron density maps. These were included in further refinement. No density corresponding to a berenil drug molecule was observed either within the minor groove or elsewhere.

At this stage an iodinated cytosine derivative [d(CGio⁵CA-ATTGCG)] of the decamer was crystallized under the same conditions as above and utilized to confirm the molecular replacement solution. Intensity data were collected from one crystal (0.2 × 0.2 × 0.6 mm³) at 294 K. Unit cell dimensions are *a* = 27.52 Å, *b* = 38.64 Å, and *c* = 53.49 Å in the same space group (*I*₂₁₂₁) as the native structure, indicating isomorphism with the native structure. Crystal decay problems precluded the collection of a full data set on the sole useable crystal available; 8753 reflections to 2.2 Å were measured and averaged to 1271 unique reflections [*R*_{sym}(*I*) = 4.75% for reflections with *I* > 0].

The position of the iodine atom was found (i) from a difference Patterson map using (*F*_{iodine} - *F*_{native})² terms as coefficients, with coordinates of 0.457, 0.082, 0.398, and (ii) from a difference Fourier map using the iododecamer reflection data and the refined native decamer model, with coordinates of 0.463, 0.083, 0.400; a large 10σ difference density was observed in the difference map 2.3 Å from the carbon-5 atom of cytosine C3. Both determinations thus showed the iodine atom to be located adjacent to the cytosine C3 residue. This unequivocally confirmed that the overall position of the decamer in the unit cell, as found by the molecular replacement solution, was correct. The positions of individual residues, especially of the terminal residues cytosine C1 and guanosine G10, were separately confirmed by a series of omit maps omitting specific nucleotides (Fig. 1), which showed residues to be in positions seen in the earlier Fourier maps.

Final refinement was carried out in the resolution range 8–2.7 Å with the X-PLOR and NUCLSQ (19) programs and an updated set of parameters for DNA (20), resulting in an *R* factor of 20.2% for 800 observations on the basis of *F* > 2σ(*F*) with 202 DNA atoms and 49 solvent molecules; r.m.s. deviations are 0.016 Å in bond lengths and 2.4° in bond angles. Averaged temperature factors for bases, sugars, and phosphates are 8.9, 23.7, and 29.9 Å²; those for the C1 and G10 bases and sugars are 30.0 and 32.1, and 16.3 and 26.2 Å², respectively. Coordinates and structure factors have been deposited in the Brookhaven and Nucleic Acid data banks.

RESULTS

Overall Features. The d(CGCAATTGCG)₂ decamers in the crystal structure do not stack end-to-end, even though their helix axes are parallel to the crystallographic *c* axis. The length of the unit cell along *c* allows for only 16 base pairs per cell to be packed in this direction. This constraint together with the dictates of the space group means that the first and last base pair of each duplex cannot contribute to the helix. Instead, an antiparallel Watson–Crick base-paired duplex is formed by just 8 of the 10 nucleotides in the decanucleotide—i.e., from residues G2 through C9. The terminal cytosine C1 and guanosine G10 nucleosides and their equivalents on the complementary strand are not base-paired but interact with the adjacent duplexes along the *c* axis, forming a complicated and unusual twofold symmetrical helix–helix junction at this point (Figs. 2 and 3) involving all four strands of the two adjacent duplexes. The ends of the four strands are not equidistant at the junction. The 3' ends of the adjacent duplexes are close,

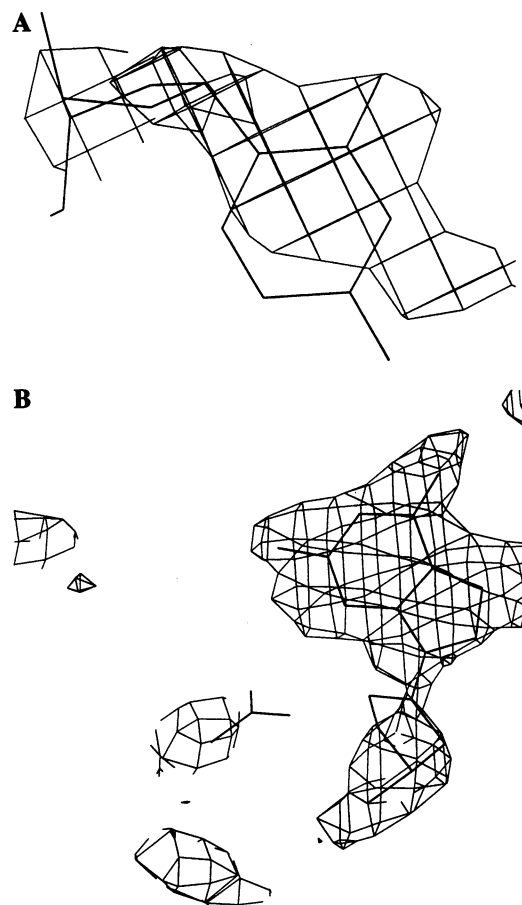


FIG. 1. Electron density in (*F*_o - *F*_c) "omit" maps (at a 1.5σ level), omitting the nucleotides C1 (A) and G10 (B). The final structures of the bases and sugars have been superimposed onto the densities. Note that the guanosine base is well-defined in the omitted electron density, whereas the final refined position for the cytosine is displaced from the density in accord with its significant mobility.

with the 3'-phosphorus atoms being 4.5 Å apart, whereas the 5' ends are 17.8 Å apart: the P ··· P separation for the 3' and 5' ends of each duplex is also 17.8 Å.

At the junction (Fig. 3), the two G2:C9 base pairs of adjacent duplexes are partially stacked on each other with a rise of 3.5 Å. The twofold symmetry relationship between them produces an arrangement distinctive from that in a continuous helix, with a pseudo helix twist angle between the two G2:C9 base pairs of 76°. The plane of the 5'-end cytosine base (C1) of one asymmetric unit single strand is parallel to each G2:C9 base pair of the neighboring duplex. C1 forms a hydrogen-bonded base triplet with it, so that the two adjacent C1·G2:C9 triplets involve bases from all four strands. The 3'-end guanine base (G10) and backbone lie in the minor groove of this next duplex, analogous to the adjacent-duplex packing of the 3'-terminal guanine observed in the B-DNA dodecamer crystal structures that crystallize in the space group *P*₂₁₂₁.

Guanine Minor Groove Binding. The G10 guanosine is oriented in the minor groove of the adjacent duplex, in its G+C terminal region, so that its (Watson–Crick) minor groove edge is inwards-facing (Fig. 4 *Top* and *Middle*). Atoms N2 and N3 are hydrogen-bonded to the edge of the guanine G2 of the adjacent duplex, forming a symmetric pair of N2 ··· N3 hydrogen bonds (3.1 Å). There are four more hydrogen-bonded contacts between the G10 nucleoside and other symmetry-related residues on both strands of this duplex in which G10 nests. These involve atoms from bases, phosphodiester backbone, and deoxyribose sugars, with two O4' deoxyribose sugar

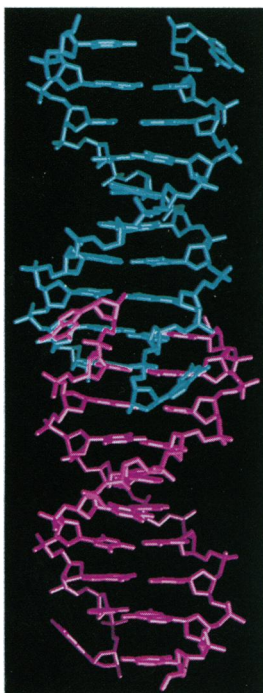


FIG. 2. Part of the continuous arrangement of $d(\text{CGCAATT-GCG})_2$ decamer duplexes along the c axis in the crystal structure, showing two duplexes, one of which is colored blue and the second is purple. The helix-helix junction region is shown in the center at the overlap of the two duplexes.

ring atoms acting as hydrogen-bond acceptors. That of cytosine C3 from this adjacent duplex plays an important role, hydrogen bonding to the inner-face N2 of G10 (3.0 Å). The other O4' participating in hydrogen bonding is from G10 itself, interacting with N2 of guanine G2 (3.1 Å), strengthening the overall interaction between G10 and G2. The O3' of G10 hydrogen bonds to N2 of guanine G8 from the opposite strand of this adjacent duplex (2.5 Å). Thus, the overall arrangement has each G10 interacting with elements of both G-C base pairs at the ends of the duplex helix.

The minor groove width in the region of the G10 minor groove insertion is 7.1 Å. This compares with 8.7 Å for the

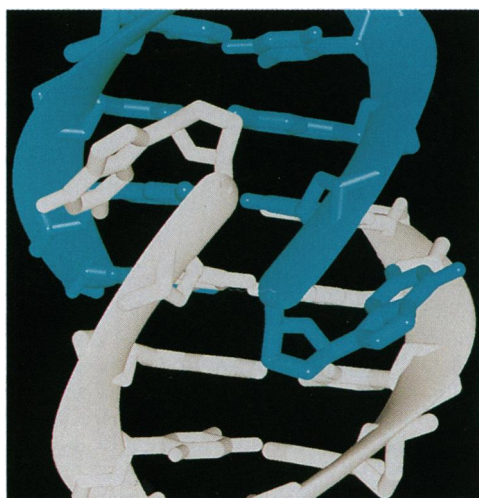


FIG. 3. View of the helix-helix junction, drawn with the GRASP program (21), looking into the minor groove formed by one of the duplexes (colored blue) at the junction. The two guanine G10 nucleosides, one from each decamer duplex, can be seen lying in the two minor grooves.

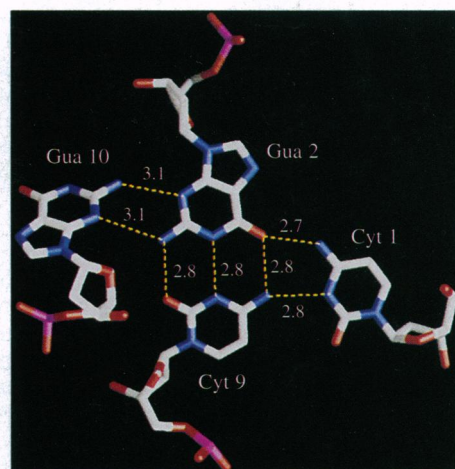
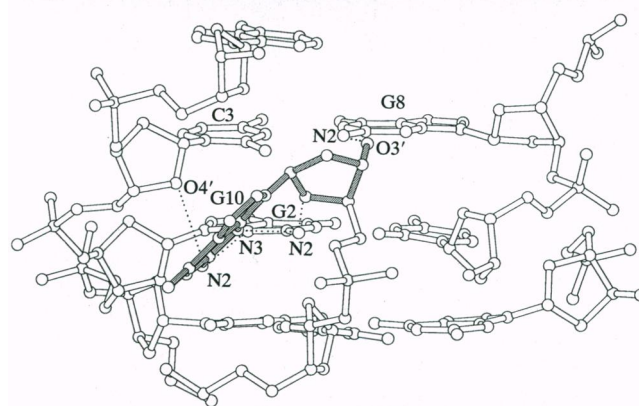
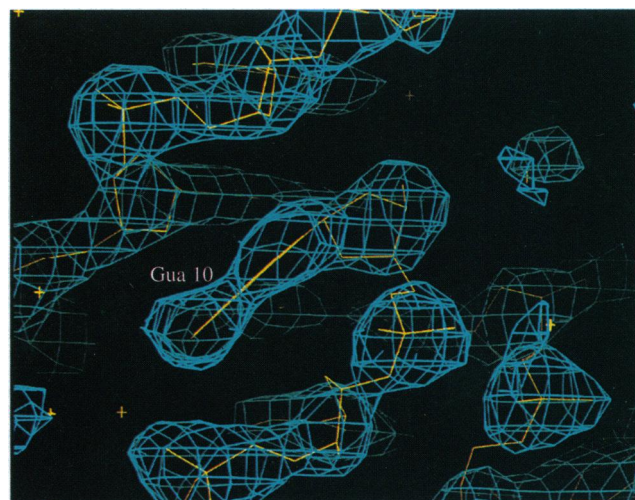


FIG. 4. (Top) $2F_o - F_c$ electron density map at a 1.5σ level, calculated over the helix-helix junction region of the $d(\text{CGCAATT-GCG})_2$ decamer, showing the 5'-terminal guanine G10, sandwiched between the phosphate backbone of its parent duplex and the 3'-terminal phosphodiester backbone of the symmetry-related duplex. (Middle) View of the hydrogen-bonded contacts involving guanosine G10. (Bottom) View of the C1-G2-C9 base triplet, together with guanosine G10.

minor groove at the 5' end of the dodecamer $d(\text{CGCGAAT-TCGCG})_2$ but is still sufficient to accommodate the guanine G10 base, sugar, and part of the connected backbone. Crystal packing of adjacent duplexes in all of the isomorphous dodecamers involves an analogous interaction of a (base-paired) guanine from the end of one duplex into the minor groove of another. The orientation of the G10 guanine in the present

Table 1. Backbone torsion angles in degrees for the asymmetric unit (one strand) of the d(CGCAATTGCG)₂ duplex

	α	β	γ	δ	ϵ	ζ	χ
Decamer							
Cyt-1	—	—	206	160	128	312	230
Gua-2	256	232	65	129	191	240	250
Cyt-3	34	182	302	152	159	265	270
Ade-4	298	186	65	156	175	245	268
Ade-5	305	196	36	130	122	300	259
Thy-6	94	164	249	156	180	279	246
Thy-7	285	187	54	145	221	231	263
Gua-8	308	163	18	161	285	150	292
Cyt-9	254	159	53	140	198	308	260
Gua-10	297	190	77	147	—	—	324
B-DNA*							
(fibers)	330	136	31	143	219	199	262
(crystals)	295	167	51	129	203	240	257

*B-DNA values (3) are averaged from single-crystal or fiber diffraction studies.

decamer is also similar to that observed for the 3'-terminal guanine base in the dodecamers, both being at an angle of $\approx 45^\circ$ to the average plane of the base pairs in the helical region of the duplexes.

The conformation of each of the nucleoside residues within the decamer is qualitatively similar to that found for averaged B-DNA from single crystal or fiber diffraction studies (Table 1). The torsion angles α and γ in residues C3 and T6 are in the g^+ , g^- rather than g^- , g^+ range, demonstrating the correlations often described for these two torsion angles. It is surprising that the values for the flipped-out nucleoside G10 are so normal. The other nonpaired residue, C1, does not adopt an extrahelical conformation and is parallel to other bases in the strand.

The C1-G2-C9 Base Triplet. The cytosine C1 base has somewhat higher mobility than other bases in the structure, with an average temperature factor of 30.0 \AA^2 . This mobility is somewhat restricted in that the cytosine must always be adjacent to both bases in the G2-C9 base pair of a symmetry-related duplex. Cytosine C1 occupies an unusually large amount of free space and is not stacked with any bases from adjacent molecules—it is held in position solely by hydrogen bonding. We suggest that this lack of stacking is a major contributor to its higher-than-average mobility. The electron density for cytosine C1 shows it to be in a coplanar position with the G2-C9 base pair and thus is always in a position to form a base triplet. We have ascertained that at least two other plausible hydrogen-bonding schemes are possible alternatives to the one described below, merely by altering the C1 backbone conformation. Even though significant density was not seen for these triplets, they may well be minor populations and thus also contribute to the mobility of C1.

In the major-occupancy position described here cytosine C1 forms a pair of hydrogen bonds with this base pair (Fig. 4 *Bottom*) involving $O6(G) \cdots N4(C)$ and $N4(C) \cdots N3(C)$ interactions—i.e., one each with the cytosine and guanine base in a reverse-Hoogsteen arrangement. The C1 cytosine has a standard *anti* conformation around the glycosidic bond. The 3' \rightarrow 5' backbone orientation for this cytosine is parallel to the cytosine C9 strand direction in the Watson-Crick base pair component of the triplet and thus is antiparallel to the strand orientation for the decamer containing guanosine G2.

Conserved Features in the B-DNA Duplex Region. The eight base pair region of the decamer forms a regular B-DNA helix. Analysis of the central AT tract shows a high conservation of the sequence-dependent structural features that were originally observed in the crystal structure of the dodecamer d(CGCGAATTCGCG)₂ (10), with high A-T base pair propeller twists and a narrow minor groove (Fig. 5) in both

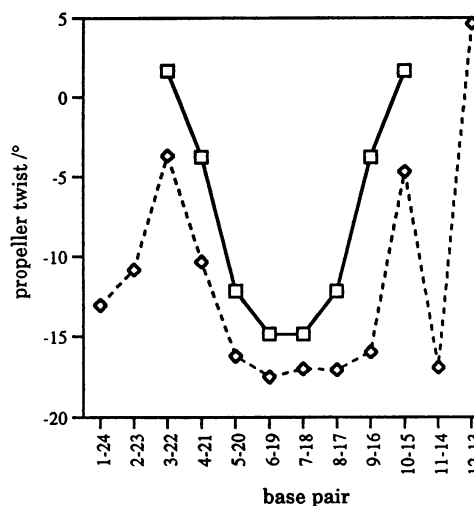
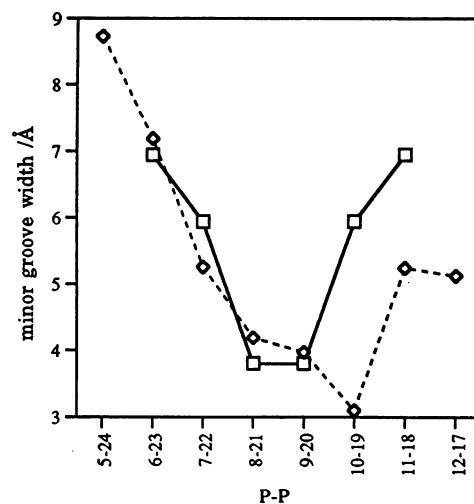


FIG. 5. (*Upper*) Plot of minor-groove widths (interstrand P...P distances in \AA — 5.8 \AA for the van der Waals radii of phosphate O atoms) in the present decamer (\square) compared to the dodecanucleotide d(CGCGAATTCGCG)₂, (\diamond). (*Lower*) Plot of propeller twist angles for the decamer duplex, calculated with NEWHEL92.

structures. The average propeller twist for the two crystallographically independent A-T base pairs in the decamer is 13° compared with 3° for the G-C ones. Beyond this A-T region, the minor groove width expands rapidly to accommodate the helix-helix junction. The particular features of the junction affect base pair parameters for only the two C-G base pairs at each end of the DNA sequence. Most change occurs in the C3-G8 base pair propeller twist, buckle, and roll parameters rather than those of the G2-C9 one.

DISCUSSION

This structure has two distinctive features: (*i*) a non-base-paired guanosine bound by a symmetrical pair of hydrogen bonds to a base-paired guanosine in the G-C region of the duplex minor groove, analogous to classic A-T-selective minor-groove drugs; and (*ii*) the first observation in a crystal structure to our knowledge of a C-G-C hydrogen-bonded base triplet in DNA.

The similarity of the position adopted by guanosine G10 to that of the terminal base-paired guanine in dodecamer crystal structures (even though they retain base-pairing at their termini) suggests that this is a generally favored arrangement for packing duplexes together that have the terminal sequence 5'-CGC... A similar arrangement has been observed (7) in

the crystal structure of the looped-out oligonucleotide d(CG-CAGAATTCGCG) for the packing of the terminal guanosine into the minor groove of the adjacent helix in the crystal lattice. The packing here presumably confers an energetic advantage in the crystal over a purely end-to-end continuous helix-stacking motif [as observed in several other decamer crystal structures (22)]. The mechanism for disrupting the terminal C-G base pair is presumably initiated by this packing of the G10 group, which provides sufficient hydrogen bonding and other interactions in the minor groove to separate a potential C1-G10 base pair. The C1 and G10 bases are then free to form other arrangements with adjacent molecules. The structure, in a distinct crystal packing arrangement compared to dodecanucleotides, thus provides unequivocal support for the concept of sequence-dependent DNA structural features that are independent of crystal lattice factors.

The interactions of drugs such as netropsin, Hoechst 33258, and pentamidine with stretches of A-T base pairs in oligonucleotides have been well characterized by a number of crystallographic studies (for example, refs. 23–26). These drug molecules typically have narrow cross sections with planar aromatic or heterocyclic rings, which are stabilized by close van der Waals contacts with the narrow minor groove in an A-T sequence. More recently, sequence-selective minor groove binding in (presumed to be wider) G-C regions has been found for series of peptide-like molecules that have been designed as dimers (27–29). The present structure suggests that appropriately designed nucleotide units could also act as G-C-selective minor-groove ligands.

The triplet arrangement observed here is very similar to the C-G-C triplet proposed for the intermediate in homologous recombination, with three DNA strands being in close proximity when exchange then occurs between them (30–32). This hydrogen-bonding scheme is of a type that has not been previously observed in DNA structures to our knowledge. [Intermolecular G-G-C triplets have been reported in an A-DNA decamer (33).] In contrast, the C⁺-G-C triplet postulated in classic inter- and intramolecular DNA triple helices (34–36) has the (protonated) cytosine in the pyrimidine third strand running antiparallel to the strand carrying the cytosines, and thus parallel to the purines (i.e., guanosines), and has hydrogen bonding via N3 and N4 solely to O6 and N7 of the guanine and not to the Watson-Crick cytosine; this triplet does not sit astride both bases of the G-C base pair as is the case in the triplet arrangement here.

This structure demonstrates that DNA duplexes with appropriate terminal sequences can pack end-to-end with the effective loss of two base pairs per duplex by forming favorable if unexpected extra-helical base...base arrangements. One possible consequence of such a junction at the site of double-stranded breaks, for example, would be that subsequent re-sealing of strands would not necessarily incorporate the extrahelical nucleotides. In any case, the bulged regions at the junctions could themselves be likely targets for nuclease attack.

We are grateful to Peter Collins for assistance with data collection and to David Brown for early work on this project. Neil Spink died on 16 July 1995. He was an outstanding student and colleague. We dedicate this paper to his memory. This work has been supported by the Cancer Research Campaign (program grant SP1384 to S.N. and a studentship to N.S.) and by the Haddow Fellowship and a Visiting Professorship at The Institute of Cancer Research (to H.M.B.). J.V. was supported by National Institutes of Health Grant GM21589 (to H.M.B.).

- Dickerson, R. E. (1992) *Methods Enzymol.* **211**, 67–111.
- Joshua-Tor, L. & Sussman, J. L. (1993) *Curr. Opin. Struct. Biol.* **3**, 323–335.
- Neidle, S. (1994) *DNA Structure and Recognition* (Oxford Univ. Press, Oxford).
- Brown, T. & Kennard, O. (1992) *Curr. Opin. Struct. Biol.* **2**, 354–360.
- Kennard, O. & Hunter, W. N. (1993) *J. Biol. Chem.* **268**, 10701–10704.
- Chattopadhyaya, R., Grzeskowiak, K. & Dickerson, R. E. (1990) *J. Mol. Biol.* **211**, 180–210.
- Joshua-tor, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D. & Sussman, J. L. (1992) *J. Mol. Biol.* **225**, 397–431.
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994) *Cell* **76**, 357–369.
- Malinina, L., Urpi, L., Salas, X., Huynh-Dinh, T. & Subirana, J. A. (1994) *J. Mol. Biol.* **243**, 484–493.
- Dickerson, R. E. & Drew, H. R. (1981) *J. Mol. Biol.* **149**, 761–786.
- Dickerson, R. E., Goodsell, D. S. & Neidle, S. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3579–3583.
- Narayana, N., Ginell, S. L., Russu, I. M. & Berman, H. M. (1991) *Biochemistry* **30**, 4449–4455.
- Larsen, T. A., Kopka, M. L. & Dickerson, R. E. (1991) *Biochemistry* **30**, 4443–4449.
- Goodsell, D. S., Kaczor-Grzeskowiak, M. & Dickerson, R. E. (1994) *J. Mol. Biol.* **239**, 79–96.
- Quintana, J. R., Grzeskowiak, K., Yanagi, K. & Dickerson, R. E. (1992) *J. Mol. Biol.* **225**, 379–395.
- Yuan, H., Quintana, J. & Dickerson, R. E. (1992) *Biochemistry* **31**, 8009–8021.
- Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr.* **21**, 273–288.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987) *Science* **235**, 458–460.
- Westhof, E., Dumas, P. & Moras, D. (1985) *J. Mol. Biol.* **184**, 119–145.
- Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. & Berman, H. M. (1996) *Acta Crystallogr.*, in press.
- Nicholls, A., Sharp, K. & Honig, B. (1991) *Proteins* **11**, 281–296.
- Privé, G. G., Yanagi, K. & Dickerson, R. E. (1991) *J. Mol. Biol.* **217**, 177–199.
- Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R. E. (1985) *J. Mol. Biol.* **183**, 553–563.
- Edwards, K. J., Jenkins, T. C. & Neidle, S. (1992) *Biochemistry* **31**, 7104–7109.
- Taberner, L., Verdagner, N., Coll, M., Fita, I., van der Marel, G., van Boom, J. H., Rich, A. & Aymami, J. (1993) *Biochemistry* **32**, 8403–8410.
- Spink, N., Brown, D. G., Skelly, J. V. & Neidle, S. (1994) *Nucleic Acids Res.* **22**, 1607–1612.
- Dwyer, T. J., Geiestranger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115**, 9900–9906.
- Mrksich, M., Parks, M. E. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116**, 7983–7988.
- Geiestranger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1994) *Science* **266**, 646–650.
- Camerini-Otero, R. D. & Hsieh, P. (1993) *Cell* **73**, 1–7.
- Zhurkin, V. B., Raghunathan, G., Ulyanov, N. B., Camerini-Otero, R. D. & Jernigan, R. L. (1994) *J. Mol. Biol.* **239**, 181–200.
- Kowalczykowski, S. C. & Eggleston, A. K. (1994) *Annu. Rev. Biochem.* **63**, 991–1043.
- Ramakrishnan, B. & Sundaralingam, M. (1993) *J. Mol. Biol.* **231**, 431–444.
- Felsenfeld, G., Davies, D. & Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024.
- Thuong, N. T. & Hélène, C. (1993) *Angew. Chem. Int. Ed.* **32**, 666–690.
- Radhakrishnan, I. & Patel, D. J. (1994) *Biochemistry* **33**, 11405–11416.