Volume 298, number 1, 97–99 FEBS 10716 February 1992

Minor-groove width and accessibility in B-DNA drug and protein complexes

Stephen Neidle

Cancer Research Campaign Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK

Received 21 November 1991; revised version received 23 December 1991

A new definition is presented for minor-groove width in double-helical B-DNA structures. This uses interstrand H4'...H5' rather than P...P distances. It is shown by examination of various oligonucleotide crystal structures that these H4'...H5' distances are a sensitive measure of minor-groove drug and protein binding, since these hydrogen atoms are in direct non-bonded contact with such bound ligands.

DNA; Minor-groove width; Ligand binding

1. INTRODUCTION

It is now well established that a number of ligands and drugs, such as distamycin, netropsin, Hoechst 33258 and berenil, bind non-covalently to AT regions of double-helical B-DNA [1–3]. Crystal-structure analyses and NMR studies have shown that they interact in the minor groove of these AT regions (see for example [4–8]). This groove can also accommodate sequence-specific covalently binding drugs, typified by anthramycin and CC-1065 [9,10], which have highly cytotoxic properties that may be related to their lack of repairability, which in turn may be a consequence of the minimal perturbations in DNA structure produced as a result of drug binding in the minor groove.

The minor groove is also used by a variety of nucleic acid binding proteins such as DNase I [11] and histone H1 [12,14]. The N-terminal arm of the helix-turn-helix engrailed homeodomain protein has been found by X-ray crystallography [15] and NMR studies [16] to fit into the minor groove of its DNA complex, with arginine residues interacting with thymine O2 oxygen atoms. An analogous arrangement has been proposed for the N-terminus of the DNA binding domain of the Hin recombinase enzyme [17,18].

The preference of both drugs and amino acid residues for AT minor-groove sequences has been attributed to several factors: specific hydrogen-bonding to adenine and/or thymine bases, the greater negative electrostatic potential in such sequences [19], the narrowing of the minor groove in AT regions and oligo(dA) tracts [20], and the complementarity between groove and ligand curvature [3,21].

Groove width is conventionally defined in terms of the shortest interstrand phosphorus...phosphorus distances, taking into account the van der Waals radius of a phosphate group (5.8 Å) [22]. However, phosphate groups are generally not oriented towards the minor groove, and they do not directly contact molecules bound to it, nor indeed to O4' atoms, which have been recently used to define groove width [6]. Rather, the hydrophobic nature of the walls of the minor-groove surface arise from the array of hydrogen atoms attached to deoxyribose C4' and C5' atoms. Crystallographic analyses of a number of drug–oligonucleotide structures in this laboratory have suggested that these hydrogen atoms play an important role in drug binding, by forming close contacts with the hydrophobic groups of the drug molecules. Hydrogen atoms cannot be directly located in these studies (with resolutions in the range 2.0–2.5 Å), and thus their positions have been calculated by standard geometric considerations. Groove widths based on inter-strand H4' to H5' distances, are presented here as a direct and relevant measure of both intrinsic sequence and minor-groove ligand-interaction effects.

2. MATERIALS AND METHODS

Coordinates for the following oligonucleotide and berenil–oligonucleotide complexes from crystallographic analyses were used in this study: $d(CGCGAATTCGCG)_2$ [23] (obtained from the Brookhaven Data Bank); $d(CGCGAATTCGCG)_2$ + berenil [7]; $d(CGCAAATTTGCG)_2$ (Edwards, K.J., Brown, D.G., Spink, N. and Neidle, S., to be published); $d(CGCAAATTTGCG)_2$ + berenil (Brown, D.G., Sanderson, M.R., Garman, E. and Neidle, S., to be published). Coordinates for a canonical B-DNA structure were generated from fibre-diffraction helical values [24]. Coordinates for the homeodomain–DNA complex [15] were kindly provided by C.O. Pabo.

Correspondence address: S. Neidle, Cancer Research Campaign Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK.

Structures were examined by means of the interactive molecular graphics program GEMINI [25] running on a Silicon Graphics IRIS 4D/20 workstation. Positions of hydrogen atoms attached to deoxyribose C4' and C5' atoms were generated by GEMINI, which was also used to calculate inter-strand distances between H4' and H5' atoms (Fig. 1), as well as between phosphorus atoms on the two DNA strands in each structure. Two H4'...H5' distances were measured for each nucleotide pair, making a total of 18 distances for each dodecanucleotide duplex. Each H4' or H5' atom was paired up with a H5' or H4' atom on the opposite strand, $(n + 3)$ nucleotides along the 3' direction. Eight interstrand phosphorus...phosphorus distances were measured for each of these four structures, in accord with the conventional definition of minor-groove width [22].

A sequence 5'-GTAATTAC, was selected from the homeodomain-DNA structure [15]. Fourteen H4'...H5' and seven P...P distances were calculated in order to describe its minor-groove width.

3. RESULTS AND DISCUSSION

The plot of groove widths in terms of H4'...H5' distances is shown in Fig. 2 for the two dodecanucleotides and their berenil minor-groove complexes. There is a smooth decrease in groove width from ~9.7 Å at the 5' ends, reaching a minimum of 5.5 Å towards the centre of the sequences. All four crystal structures show these features, although there are some differences in detail, notably for the native d(CGCGAATTCGCG)$_2$. In each case, the central shallow minimum groove width extends over between three and four base pairs. The average width in both d(CGCAAATTTGCG)$_2$ and its berenil complex, for the inner, constant three base-pair region is 5.50 Å. The native d(CGCGAATTCGCG)$_2$ structure has a significantly narrower region, of average width 4.9 Å, although its berenil complex is substantially wider in this region (average 5.58 Å). The interstrand H4'...H5' distance in canonical B-DNA is 7.01

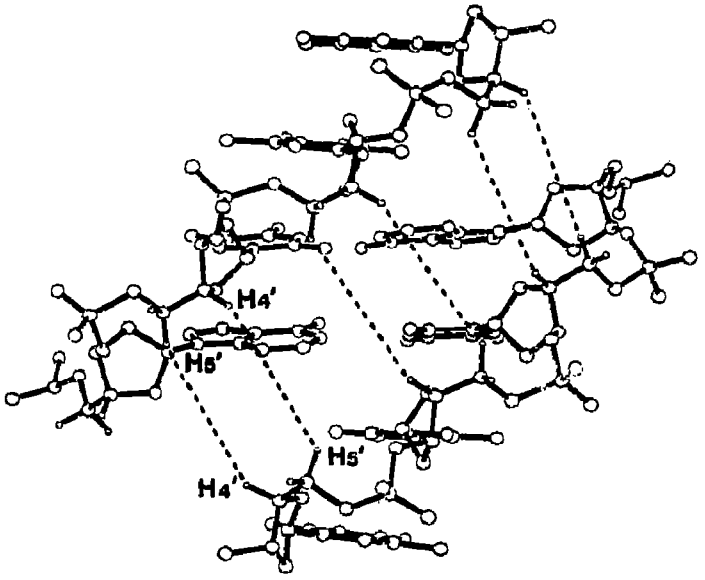

Fig. 1. The molecular structure of the central AT region in d(CGCAAATTTGCG), (Edwards et al., to be published), showing H4' and H5' hydrogen atoms at generated positions. Dashed lines indicate calculated interstrand H4'...H5' distances (see text).

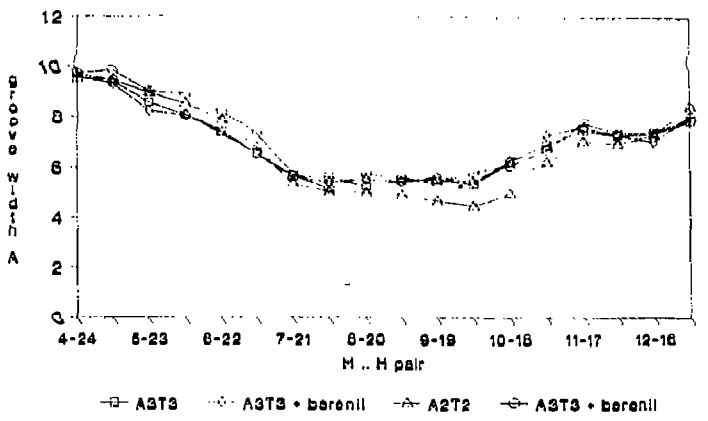

Fig. 2. Plot of H4'...H5' interstrand distances for two oligonucleotides and their berenil complexes.

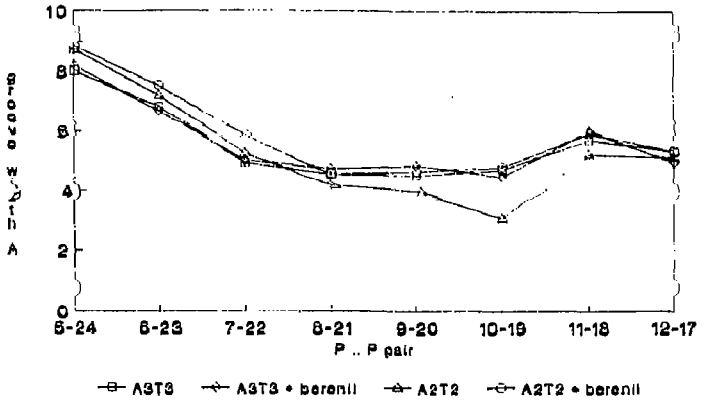

Fig. 3. Plot of P...P interstrand distances.

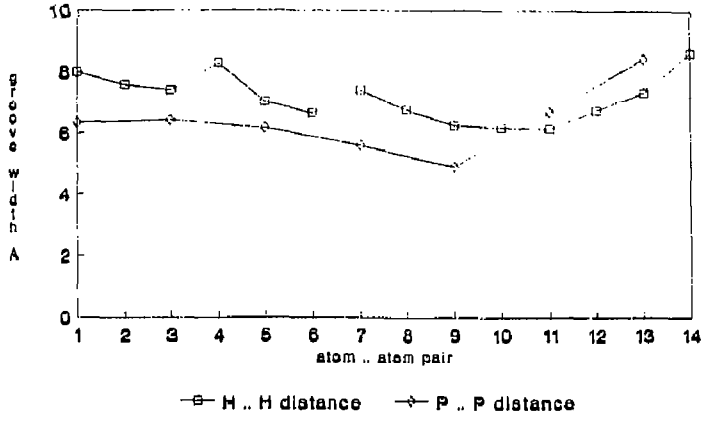

Fig. 4. Plot of H4'...H5' and P...P distances for the AT region in the homeodomain–DNA structure [15].

Å; the two distances per base-pair are necessarily identical.

These hydrogen atoms cannot be directly located in these oligonucleotide–drug and –protein structures, in contrast to phosphorus atoms. However, the puckers of deoxyribose sugars in oligonucleotides, which define the position of C4' and C5' atoms and their bonded atoms, can normally be unambiguously assigned at 2.0–2.5 Å resolution by the process of crystallographic refine-

ment, except in those relatively rare instances when a sugar group is disordered. Such circumstances, or when a pucker has been incorrectly assigned, would be readily revealed in the interstrand H4′...H5′ plots such as Fig. 2 or 4 as marked discontinuities in the otherwise smooth curves. It should also be noted that average temperature factors in the AATT region of the Dickerson–Drew dodecamer are somewhat higher for phosphorus (48 Å^2) than for C5′ or C4′ atoms (42 Å^2).

The P...P minor-groove width plot for these structures (Fig. 3) (which takes into account the 5.8 Å van der Waals radius of phosphate groups), shows a similar trend in width variations. Since there are now fewer sampling points, changes such as at the 3′ end of the narrow groove region are much more abrupt. The apparent exceptionally narrow groove width in d(CGCGAATTCGCG)$_2$ (3.1 Å), is actually an artifact due to the unusual phosphate conformation at this point [26]. Groove widths in this central narrow AATT region in the three other structures average 4.6 Å.

More pronounced differences between the two types of groove width are found for the AT-rich region of the homeodomain-DNA complex (Fig. 4), with the H4′...H5′ width measurements showing several points of local widening and consequent narrowing. These former correspond to the relatively bulky N-terminal amino acid residues Arg-Pro-Arg being accommodated in the minor groove and slightly widening it. By contrast, the P...P plot fails to show these effects.

The interstrand H4′...H5′ distances in the minor groove thus provide a sensitive indication of structural changes produced by ligand binding. This is emphasised by the two sampling intervals per base pair step. Most importantly, these hydrogen atoms form an important part of the minor-groove surface and thus can directly contact ligands binding in the groove. The average H4′...H5′ interstrand separation of 5.5 Å in the AT region of the oligonucleotide crystal structures represents a free space of just ca. 3.5 Å when the van der Waals radii of the hydrogen atoms are taken into account. This is equal to the van der Waals thickness of an aromatic ring system. Thus, the H4′...H5′ distances can also directly indicate the size of ligand or group that can be accommodated in the minor groove without DNA backbone distortion.

Acknowledgements: C.O. Pabo is thanked for providing coordinates. I am grateful to my colleagues David Brown and Karen Edwards for discussions, and to the Cancer Research Campaign for support.

REFERENCES

[1] Zimmer, C. and Wähnert, U. (1986) Prog. Biophys. Mol. Biol. 47, 31–112.
[2] Dervan, P.B. (1986) Science 232, 464–471.
[3] Goodsell, D. and Dickerson, R.E. (1986) J. Med. Chem. 29, 727–733.
[4] Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P.D. and Dickerson, R.E. (1985) J. Mol. Biol. 183, 553–563.
[5] Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) Proc. Natl. Acad. Sci. USA 84, 8385–8389.
[6] Quintana, J.R., Lipanov, A.A. and Dickerson, R.E. (1991) Biochemistry 30, 10294–10306.
[7] Brown, D.G., Sanderson, M.R., Skelly, J.V., Jenkins, T.C., Brown, T., Garman, E., Stuart, D.I. and Neidle, S. (1990) EMBO J. 9, 1329–1334.
[8] Lane, A.N., Jenkins, T.C., Brown, T. and Neidle, S. (1991) Biochemistry 30, 1372–1385.
[9] Hurley, L.H. and Needham-Vandervanter, D.R. (1986) Acc. Chem. Res. 19, 230–237.
[10] Hurley, L.H., Lee, C.-S., McGovren, J.P., Warpehoski, M.A., Mitchell, M.A., Kelly, R.C. and Aristoff, P.A. (1988) Biochemistry 27, 3886–3892.
[11] Suck, D., Lahm, A. and Oefner, C. (1988) Nature 332, 464–468.
[12] Turnell, W.G., Satchwell, S.C. and Travers, A.A. (1988) FEBS Lett. 232, 263–268.
[13] Suzuki, M. (1989) EMBO J. 8, 797–804.
[14] Churchill, M.E.A. and Suzuki, M. (1989) EMBO J. 8, 4189–4195.
[15] Kissinger, C.R., Liu, B., Martin-Blanco, E., Komberg, T.B. and Pabo, C.O. (1990) Cell 63, 579–590.
[16] Qian, T.Q., Billeter, M., Otting, G., Müller, M., Gehring, W.J. and Wüthrich, K. (1989) Cell 59, 573–590.
[17] Sluka, J.P., Horvath, S.J., Glasgow, A.C., Simon, M.I. and Dervan, P.B. (1990) Biochemistry 29, 6551–6561.
[18] Affolter, M., Percival-Smith, A., Müller, M., Billeter, M., Qian, Y.Q., Otting, G., Wüthrich, K. and Gehring, W.J. (1991) Cell 64, 879–880.
[19] Pullman, B. (1983) J. Biomol. Struct. Dynamics 1, 773–794.
[20] Travers, A.A. (1989) Annu. Rev. Biochem. 58, 427–452.
[21] Zasedatelev, A.S. (1991) FEBS Lett. 281, 209–211.
[22] See for example, the NEWHELIX 91 program for calculating helical and conformational parameters.
[23] Drew, H.R. and Dickerson, R.E. (1981) J. Mol. Biol. 151, 535–556.
[24] Arnott, S. and Hukins, D.W.L. (1973) J. Mol. Biol. 81, 93–105.
[25] GEMINI Version 1.3: A Molecular-modelling Program (A.J. Beveridge) CRC Biomolecular Structure Unit, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK.
[26] Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) J. Biol. Chem. 257, 14686–14707.