The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG)

Jie Liu[a], Lucy Malinina[a,b], Tam Huynh-Dinh[c], Juan A. Subirana[a,*]

[a] *Departament d'Enginyeria Química, Universitat Politènica de Catalunya, Av. Diagonal 647, E-08028 Barcelona, Spain*
[b] *Engelhardt Institute of Molecular Biology RAS, Vavilova 32, Moscow 117984, Russia*
[c] *Unité de Chimie Organique, Institut Pasteur, 28 R. du Dr. Roux, 75724 Paris, France*

Received 24 July 1998; received in revised form 30 September 1998

Abstract The title oligonucleotide and many related dodecamers have been extensively studied alone and as DNA-drug complexes. In practically all cases they were found to crystallize in the same space group, stabilized by interactions among the terminal guanine bases. Here we report new packing interactions (R3) in the presence of Ca^{2+}. The oligonucleotides interact by placing their terminal guanines in the narrow groove of a neighbor molecule, an interaction which had never been found in dodecamers.

Key words: B-DNA structure; Oligonucleotide; X-ray crystallography; Calcium ion

1. Introduction

The determination of the structure of the dodecamer d(CGCGAATTCGCG) was a milestone in structural biology [1] since it confirmed unambiguously the double helical model of B-DNA and at the same time showed many details of its structural variability [2]. Since then the structure of this oligonucleotide has been determined under different conditions [3–5] and at a higher resolution [6]. Furthermore this oligonucleotide has been used to establish the structure of about 20 oligonucleotide-drug complexes [7]. Thirty oligonucleotides with a related sequence have also been studied [7]. In all these cases, the dodecanucleotides have been found to crystallize in $P2_12_12_1$ and related space groups, usually with Mg^{2+} as a counterion. The crystalline matrix is stabilized in a very specific manner by guanine-guanine interactions among the two terminal base pairs of each molecule.

In our laboratory it was found [8,9] that several related sequences such as d(CGCTCTAGAGCG), d(CGCAAATTTGCG) and d(CGm·CTTTAAAGCG) could crystallize in the R3 space group, but crystals diffracted up to rather low resolution. Recently we have been able to obtain crystals of the classical d(CGCGAATTCGCG) dodecamer in the same R3 space group and in a similar unit cell, using a rather high concentration of calcium ions. The resolution was limited, but higher than in previous cases [8,9]. In this paper, we present the results obtained.

2. Materials and methods

Crystals of the title compound were obtained in a hanging drop containing dodecamer, 0.2 M $CaCl_2$, protamine, MPD and cacodylic acid as buffer, equilibrated against a reservoir with a higher concentration of MPD at 20°C. Protamine was not essential to obtain suitable crystals. Similar crystals of the various dodecamer sequences mentioned above [8,9] were grown in the presence of either Mg^{2+} or Ca^{2+}, in some cases with small polypeptides or drugs. Interestingly, the dodecamer d(CGCTCTAGAGCG) was crystallized in the R3 unit cell in the presence of Mg^{2+} [9], but in a $P2_1$ unit cell [10] in the presence of Ca^{2+}, just the opposite behavior to that reported here for the title compound.

The X-ray diffraction data of the title compound were collected up to 2.5 Å, giving 1262 $I > 0$ unique reflections with R_{sym} 11.0% in the resolution range 18–3.0 Å (95.7% complete). The space group and cell parameters are R3, $a = b = 41.96$, $c = 101.40$Å, corresponding to nine duplexes in the unit cell or one in the asymmetric unit. Occasionally crystals with a lower symmetry ($P3_1$) were obtained with the same unit cell dimensions. The structure was determined by the molecular replacement method with the AMoRe program [11] using the central decamer of Dickerson structure [1] as searching model. Then the two terminal base pair sets d(CGC) and d(GCG) of the solved structure of d(CGCAATTGCG) [12] were fitted into the solution. Crystallographic refinement of the model with X-PLOR [13] program resulted in R 19.3% and R_{free} 23.5%. The two disordered terminal cytosine residues are excluded from the model. One calcium ion and 15 water molecules were also included in the asymmetric unit. Coordinates have been deposited in the Nucleic Acid Database (Accession Code: HPUB) [7].

3. Results

In each dodecamer molecule only the central decamer has the standard B-form, as shown in Fig. 1. These central decamers pack as infinite columns, although they do not form pseudo-continuous helices. The twist angle between consecutive decamers is about −75°. The terminal base pairs at both ends of the molecule are disrupted in such a way that the guanine is tightly packed inside the minor groove of the neighbor dodecamer. On the other hand the terminal cytosines are highly disordered and have not been detected in the electron density map, even the phosphate group of GUA2 is missing on the map. This mode of interaction had already been found by Spink et al. [12] in a decamer. The latter authors also found that the terminal cytosines were highly disordered. Here we describe for the first time this type of interaction in a dodecamer. All dodecamers with a CG starting sequence previously crystallized are stabilized by guanine-guanine interactions which do not disrupt the terminal Watson-Crick base pairs.

Molecules in the same layer of the unit cell are packed in a pseudo-hexagonal manner, as shown in Fig. 2. A similar packing arrangement had been described long ago for double helical RNA fibers [14]. Each molecule is related by ternary axes with its neighbors. As a result, short phosphate-phosphate distances are found between the reference dodecamer (the central one) and two of its neighbors. Fewer contacts are found with two other neighbors and no contacts with the

*Corresponding author. Fax: (34) (3) 4017150.
E-mail: subirana@eq.upc.es

PII: S0014-5793(98)01295-2

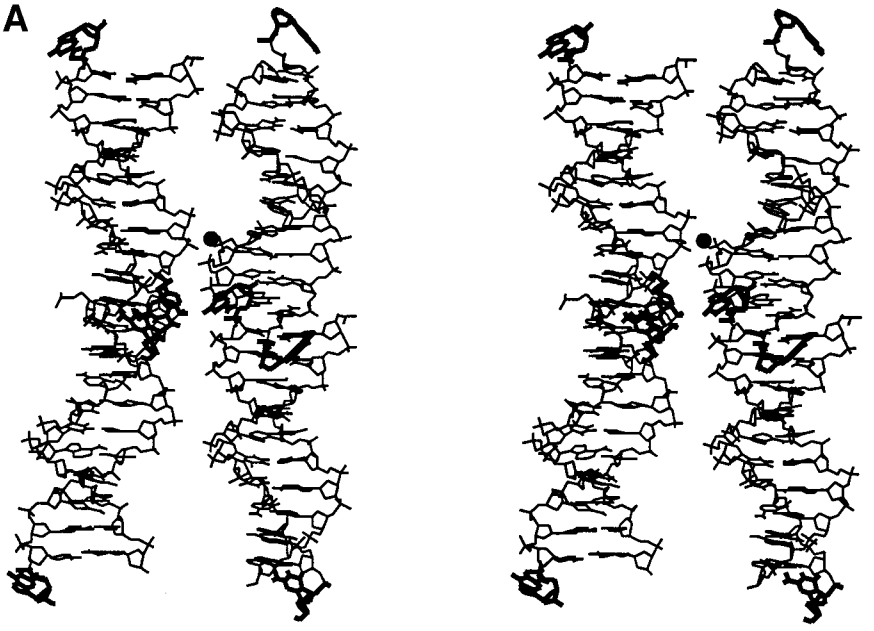

Fig. 1. A: Stereo view showing the packing of duplexes along the *c* direction. The calcium ion is indicated as a black dot. Central decamers form infinite columns and the terminal guanines (shown in heavy trace) pack tightly inside the minor groove of the neighbor duplex. Terminal cytosines are disordered and not shown. Each column is located on a 3_2 axis. Note the close contacts of some phosphate groups. B: Labeling scheme of the title compound.

other two. Short phosphate-phosphate distances indicate the presence of either ordered or disordered cations in their neighborhood and therefore can be considered as a stabilizing factor of the packing arrangement shown in Fig. 2. One calcium ion was found to be present on the ternary axis among those three molecules which have the largest number of contacts and have phosphate-phosphate distances as short as 5.28 Å. The calcium-oxygen distance is 2.31 Å. Besides, among those three molecules with a largest number of contacts, hydrogen bonds are also found between O6 (GUA12) of one molecule and O1P (GUA12) of a neighbor. At the other two symmetry related layers (above and below the layer shown in Fig. 2), there are also the same close phosphate contacts and hydrogen bonds, but arranged around different ternary axes, which thus keep the whole crystal packing stabilized.

4. Discussion

The new packing arrangement found here was due to the use of Ca^{2+} instead of Mg^{2+} in the crystallization solution. There are several reports showing that the change in counterions gives different packing arrangements in oligonucleotide crystals [8–10,15,16]. Calcium ions are usually hydrated and are often found in the narrow groove [15–17]. The comparatively low resolution of our data does not allow us to locate any ions in the narrow groove, but they might also be present. In other cases, the calcium ions play an essential role in the packing interactions among terminal phosphate groups, either by direct interaction with the phosphate oxygens [10] or through their waters of hydration [18]. In our case it appears that one calcium ion also plays a role in stabilizing the packing arrangement, as shown in Fig. 1. When all these data are taken together, it is evident that the use of either Mg^{2+} or Ca^{2+} may result in different packing arrangements of oligonucleotides upon crystallization, but it does not appear possible to predict packing as a function of the crystallization conditions. It is also not clear why in some cases the calcium ions are only found in the narrow groove, whereas in other cases they play a role in the packing interactions among different oligonucleotide molecules.

Comparison with the standard Dickerson dodecamer [1] is shown in Fig. 3. It is apparent that the central decamer has a structure very similar to that found in the original dodecamer (as confirmed by NEWHELIX, results not shown), only the two terminal base pairs have a different arrangement. In Dickerson's case, the two neighbor duplexes in the same column are related by a 2_1 axis, that is essential for the specific guanine-guanine interaction. In our case, three consecutive duplexes in the same column are related by a 3_2 axis and adopt a new guanine-minor groove interaction.

The end-to-end interaction of the dodecamers in the R3 space group is similar to that found in decamers [12] with

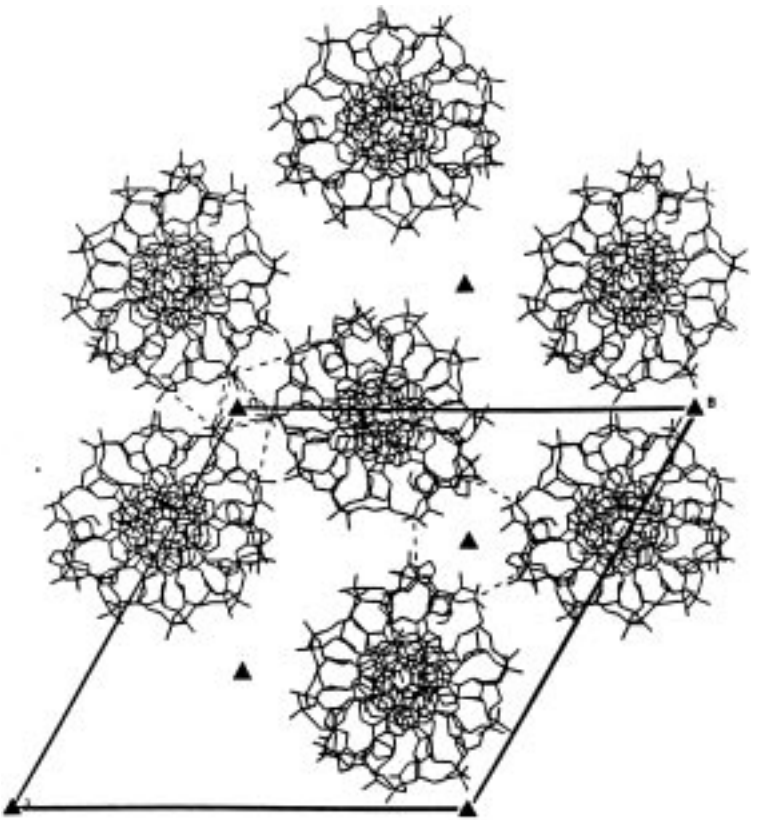

Fig. 2. The pseudo-hexagonal arrangement of molecules at the same layer of columns. Black triangles show the positions of ternary axes which are perpendicular to the projection plane. Three different sets of molecules around each ternary axis can be seen. Dotted lines represent close phosphate contacts with a distance less than 7.0 Å. The three molecules around the origin show the closest contacts (5.28, 6.28 and 6.79 Å). The calcium ion coincides with the ternary axis in this projection.

the same CGC terminal sequence. In retrospect this is not unexpected. In the decamers the upper end is displaced eight base pairs from the lower end and interacts with the next molecule, which is rotated 180° due to the binary screw axis. In the dodecamers, the presence of two additional base pairs in the central decamer region requires that the following dodecamer should be rotated $180°+2\times36°=252°$ in order to show the same interaction. In the R3 space group the rotation is 240°, which is close to this expected value. The difference of 12° can be accommodated by small changes throughout the whole molecule.

In view of this situation it is surprising that the R3 packing

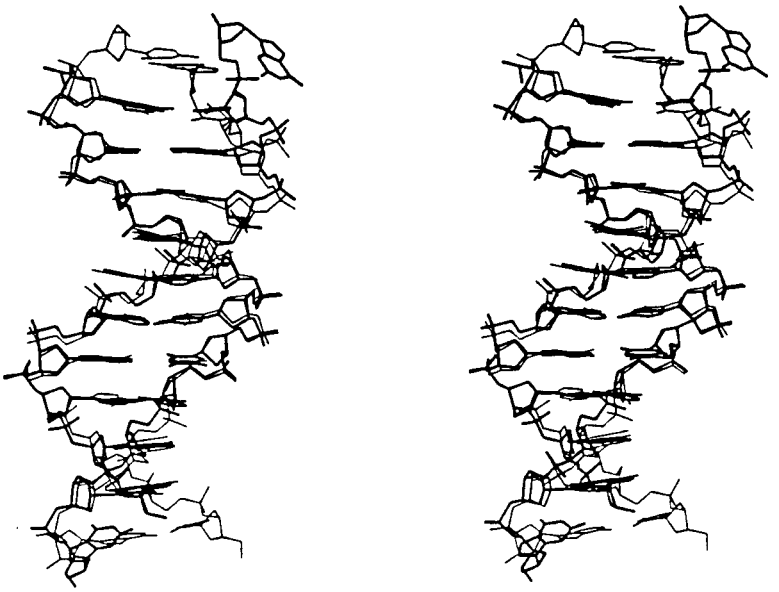

Fig. 3. Comparison of the R3 structure with the standard Dickerson dodecamer. Both central decamers were fitted with the X-PLOR program. Thick lines represent the R3 structure. No significant variation is observed in the central decamers.

in dodecamers has not been described before, in view of the large number of oligonucleotides with this terminal sequence which have been crystallized [7]. A likely explanation is that a low resolution is usually found [8,9] in dodecamers crystallized in this space group and the results which other investigators might have found have not been reported.

In summary our results show that the structure of the central decamer is not strongly influenced by packing forces, which gives confidence to the B-form parameters calculated for the crystalline structure either alone or in the presence of drugs. Furthermore we have demonstrated that the guanine-minor groove interaction described by Spink et al. [12] can also be found in dodecamers.

Acknowledgements: We are thankful to V. Tereshko and G. Minasov for advice and discussion. This work is supported by a grant (PB93-1067) and a fellowship (J.L.) from the Ministerio de Educación y Cultura.

References

[1] Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) Proc. Natl. Acad. Sci. USA 78, 2179–2183.
[2] Fratini, A.V., Kopka, M.l., Drew, H.R. and Dickerson, R.E. (1982) J. Biol. Chem. 257, 14686–14707.
[3] Drew, H.R., Samson, S. and Dickerson, R.E. (1982) Proc. Natl. Acad. Sci. USA 79, 4040–4044.
[4] Holbrook, S.R., Dickerson, R.E. and Kim, S.H. (1985) Acta Crystallogr. B 41, 255–262.
[5] Westhof, E. (1987) J. Biomol. Struct. Dyn. 5, 581–600.
[6] Shui, X., McFail-Isom, L., Hu, G.G. and Williams, L.D. (1998) Biochemistry 37, 8341–8355.
[7] Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A.R. and Schneider, B. (1992) Biophys. J. 63, 751–759.
[8] Tabernero, L. (1991) PhD Thesis, Universitat Politècnica de Catalunya.
[9] Urpí, L. (1992) PhD Thesis, Universitat de Barcelona.
[10] Tereshko, V., Urpí, L., Malinina, L., Huynh-Dinh, T. and Subirana, J.A. (1996) Biochemistry 35, 11589–11595.
[11] Navaza, J. (1994) Acta Crystallogr. A 50, 157–163.
[12] Spink, N., Nunn, C.M., Vojtechovsky, J., Berman, H.M. and Neidle, S. (1995) Proc. Natl. Acad. Sci. USA 92, 10767–10771.
[13] Brünger, A.T., Kuriyan, J. and Karplus, M. (1987) Science 235, 458–460.
[14] Arnott, S., Hutchinson, F., Spencer, M., Wilkins, M.H.F., Fuller, W. and Langridge, R. (1966) Nature 211, 227–232.
[15] Yuan, H., Quintana, J. and Dickerson, R.E. (1992) Biochemistry 31, 8009–8021.
[16] Lipanov, A., Kopka, M.L., Grzeskowiak, M.K., Quintana, J. and Dickerson, R.E. (1993) Biochemistry 32, 1373–1389.
[17] Goodsell, D.S., Grzeskowiak, K. and Dickerson, R.E. (1995) Biochemistry 34, 1022–1029.
[18] Grzeskowiak, K., Goodsell, D.S., Grzeskowiak, M.K., Cascio, D. and Dickerson, R.E. (1993) Biochemistry 32, 8923–8931.