

# The crystal structure analysis of d(CGCGAASSCGCG)<sub>2</sub>, a synthetic DNA dodecamer duplex containing four 4'-thio-2'-deoxythymidine nucleotides

Titus J. Boggon, E. Louise Hancox<sup>1</sup>, Katherine E. McAuley-Hecht<sup>2</sup>, Bernard A. Connolly<sup>3</sup>, William N. Hunter, Tom Brown<sup>2,+</sup>, Richard T. Walker<sup>1</sup> and Gordon A. Leonard\*

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK, <sup>1</sup>School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK, <sup>2</sup>Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ, UK and <sup>3</sup>Department of Biochemistry and Genetics, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne NE2 4HH, UK

Received September 19, 1995; Revised and Accepted December 15, 1995

## ABSTRACT

The crystal structure refinement of the synthetic dodecamer d(CGCGAASSCGCG), where S = 4'-thio-2'-deoxythymidine, has converged at  $R = 0.201$  for 2605 reflections with  $F > 2\sigma(F)$  in the resolution range 8.0–2.4 Å for a model consisting of the dodecamer duplex and 66 water molecules. A comparison of its structure with that of the native dodecamer d(CGCGAATTCGCG) has revealed that the major differences between the two structures is a change in the conformation of the sugar-phosphate backbone in the regions at and adjacent to the positions of the modified nucleosides. Examination of the fine structural parameters for each of the structures reveals that the thiosugars adopt a C<sub>3'</sub>-*exo* conformation in d(CGCGAASSCGCG), rather than the approximate C<sub>1'</sub>-*exo* conformation found for the analogous sugars in the structure of d(CGCGAATTCGCG). The observed differences in structure between the two duplexes may help to explain the enhanced resistance to nuclease digestion of synthetic oligonucleotides containing 4'-thio-2'-deoxynucleotides.

## INTRODUCTION

The biological properties of 4'-thionucleosides (1–5), which are sulphur-containing isoelectronic analogues of natural nucleosides (Fig. 1), are not always apparent from a comparison with those of the corresponding oxysugar-containing deoxynucleoside. Hence, while (*E*)-5-(2-bromovinyl)-2'-deoxyuridine and its sulphur-containing analogue both show potent and selective inhibition of HSV-1 and varicella zoster virus (VZV), the strong antiviral properties of 5-isopropyl- and 5-cyclopropyl-4'-thio-2'-deoxy-

uridine are markedly different to those of the corresponding natural nucleosides, which show no antiviral activity at all. Similarly, 5-vinyl-2'-deoxyuridine is very toxic, whereas the corresponding 4'-thio analogue shows no toxicity, but is active against HSV-1.

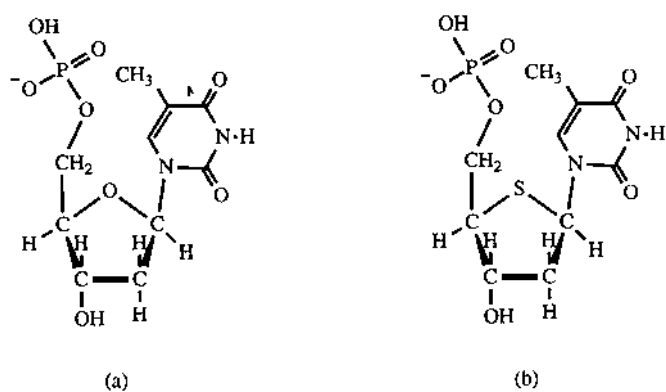
Some of the most interesting properties of pyrimidine 4'-thio-2'-deoxynucleotides are that they are significantly more lipophilic than oxysugar-containing deoxynucleotides and have a much longer serum half-life than normal nucleotides, the latter property arising because they are stable to phosphorolysis. This led to the suggestion that oligonucleotides containing pyrimidine 4'-thio-2'-deoxynucleosides might be useful in the antisense field, especially as structure determination had shown the modified nucleosides to adopt similar conformations to those of the natural nucleosides (6). Additionally, circular dichroism and ultraviolet melting studies (7) of thiosugar nucleoside-containing oligonucleotides had shown them to adopt the same overall conformation as the analogous unmodified oligonucleotides and that the duplexes were only slightly less thermodynamically stable.

Experiments involving the synthesis of oligonucleotides containing the *EcoRV* recognition site revealed that when a 4'-thiothymidine residue was incorporated on the 5'-side of the scissile bond the DNA duplex was resistant to endonuclease attack (7). The reasons for this were not at all clear, as a crystal structure of the complex between *EcoRV* and the modified duplex (Kostrewa, D., Hancox, E.L., Walker, R.T., Connolly, B.A. and Winkler, F.K., unpublished results; 8) showed that the conformation of the DNA had not significantly changed. However, the Mg<sup>2+</sup> ion required for enzyme activity was not present in the structure.

As part of a continuing investigation into the structures and biological properties of oligonucleotides containing 4'-thio-2'-deoxynucleosides we report here the results of an X-ray structural analysis of the self-complementary dodecamer d(CGCGAASSCGCG), where S = 4'-thio-2'-deoxythymidine (Fig. 1). The observed

\* To whom correspondence should be addressed

+Present address: Department of Chemistry, University of Southampton, Highfield, Southampton SO19 5NH, UK



**Figure 1.** Representations of (a) 2'-deoxythymidine-5'-phosphate and (b) its isoelectronic analogue 4'-thio-2'-deoxythymidine-5'-phosphate (S).

structure is compared with that of the native dodecamer d(CGCGATTCGCG) (9) and the structural reasons for the

resistance of 4'-thiosugar-containing oligonucleotides to endonuclease attack are explored.

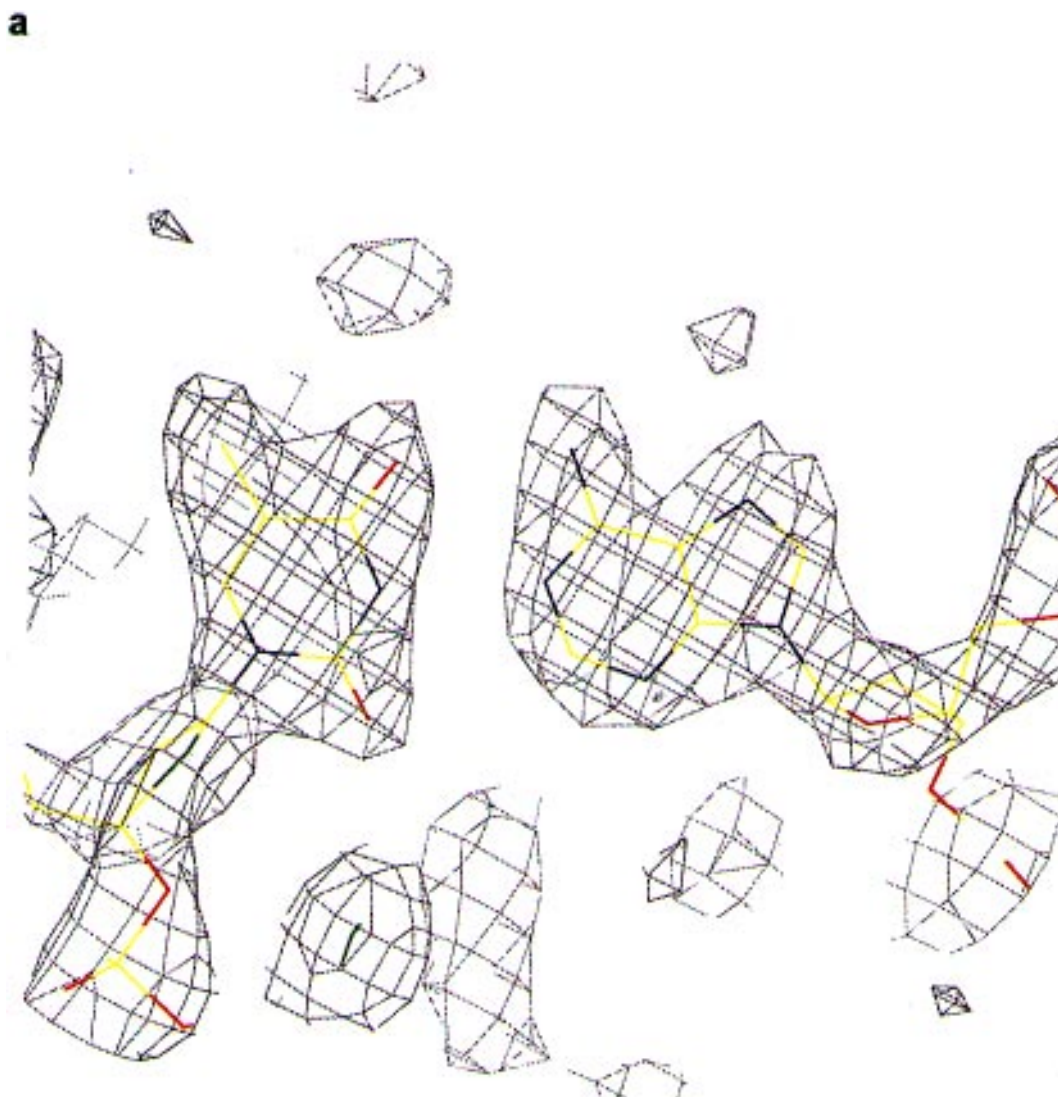
## MATERIALS AND METHODS

### Chemical synthesis

The oligonucleotide was synthesized (10  $\mu$ M scale) using an Applied Biosystems 381A DNA synthesizer. The modified deoxynucleoside phosphoramidite (7) was introduced at the fifth position of the synthesis as previously described (7). The synthesis and purification of the oligodeoxynucleotide was achieved under standard conditions.

### Crystallization, X-ray data collection and structure refinement

Thin, colourless, needle-shaped crystals were grown at ambient temperature in sitting drops which contained the dodecanucleotide (0.5 mM), sodium cacodylate buffer (14 mM, pH 6.0), magnesium chloride (17 mM), spermine tetrahydrochloride (1 mM) and



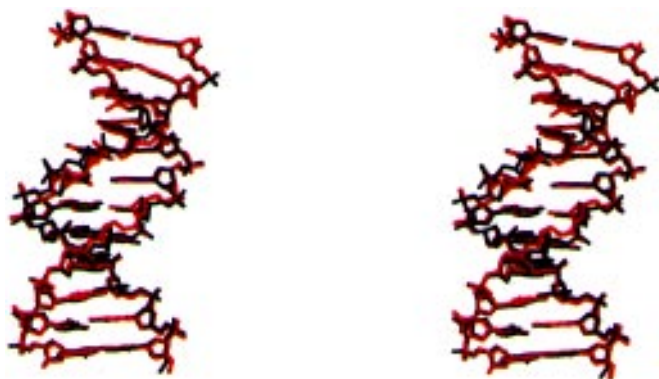


**Figure 2.** (a) The final calculated electron density ( $2F_{\text{obs}} - F_{\text{calc}}$ ,  $\alpha_{\text{calc}}$ ; grey chicken wire) for one of the A.S base pairs in the structure of d(CGCGAASSCGCG). The nucleotides are depicted as sticks with the following atom colouring scheme: carbon, yellow; nitrogen, blue; oxygen, red; sulphur, green; phosphorus, yellow. The electron density is contoured at approximately one standard deviation (SD) from the r.m.s. level found in the unit cell. For clarity solvent molecules have not been shown. (b) A close-up of the final calculated electron density for the sugar-phosphate backbone in the region of one of the 4-thiosugars. Two contour levels are shown: 1 SD from the r.m.s. level (grey) and 2.75 SD from the r.m.s. level (brown). Note the spherical, high level density at the sulphur position of the 4-thio sugar when compared with the density for the unmodified sugar shown.

2-methyl-2,4-pentanediol (MPD) (17% v/v), which were diffused against external reservoirs containing 50% aqueous MPD. Two such crystals were then mounted in thin-walled glass capillaries for X-ray data collection. Data from the first of these, which had dimensions of  $1.0 \times 0.05 \times 0.05$  mm, was collected at  $\lambda = 0.882$  Å on station PX9.6 at the Synchrotron Radiation Source, Daresbury Laboratory, to a resolution limit of 2.35 Å using a 18 cm MAR Research image plate detector and  $17.4^\circ$  oscillation frames, each exposed for 120 s. This data was later supplemented by a set collected from a bigger crystal ( $1.5 \times 0.15 \times 0.15$  mm) to a resolution limit of 2.5 Å on a Rigaku RU200 rotating anode generator ( $\lambda = 1.5418$  Å) operating at 50 kV, 140 mA and equipped with a RAXIS IIc image plate system. This latter data set was collected on 39  $3^\circ$  oscillation frames, each of which were exposed to X-rays for 1 h. All the data was processed using the

program MOSFLM (version 5.20) (10), then scaled, merged and reduced using programs from the CCP4 suite of programs (11) to yield a total of 2820 unique reflections with  $R_{\text{sym}} = 0.065$  to a resolution of 2.4 Å and a multiplicity of 6.3.

Orthorhombic unit cell dimensions of  $a = 24.99$  Å,  $b = 40.36$  Å,  $c = 65.99$  Å and space group  $P2_12_12_1$  suggested that the structure of d(CGCGAASSCGCG) is isomorphous with that of the native dodecamer d(CGCGAATTCGCG) (9). The initial stage of the structure refinement therefore consisted of rigid body minimization, against the observed data, of the structure of the native dodecamer stripped of solvent molecules and with sulphur atoms replacing oxygen atoms where necessary. This procedure, as with the rest of the refinement, was carried out using the program X-PLOR (12,13) and converged with  $R = 0.333$  for the 2605 reflections with  $F > 2\sigma(F)$  in the resolution range 8.0–2.4 Å. The refinement then



**Figure 3.** A least squares superposition of the structure of d(CGCGAASSCGCG) (black sticks) and the 1BNA structure of d(CGCGAATTCGCG) (red sticks). Note that the positions of the bases remains almost constant and that major differences between the two structures are confined to the conformation of the sugar-phosphate backbone in the vicinity of the modified nucleosides.

continued, using the same data, with a round of simulated annealing-based minimization. This consisted of initial energy minimization of the model followed by a slow cooling procedure (initial temperature 1000 K, final temperature 300 K, temperature decrement 25 K, time step 0.5 fs) and one round (120 cycles) of Powell minimization. Throughout this procedure and for subsequent rounds of refinement the dictionary of restraints used was that provided with X-PLOR and modified to include C-S bond length and C-S-C bond angle values for 4'-thio-2'-deoxy-thymidine, which were taken from Dyson *et al.* (3). It should be noted that the values used for dihedral angles and improper angles in the modified nucleoside were those already in place for 2'-deoxythymidine.

At this stage *R* had been reduced to 0.258. Both electron density ( $2F_o - F_c$ ,  $\alpha_{\text{calc}}$ ) and difference density maps ( $F_o - F_c$ ,  $\alpha_{\text{calc}}$ ) were then calculated using the CCP4 suite of programs (11) and displayed and examined on a graphics workstation using the program O (14). Modifications to the model were then made in order to improve its fit to the electron density and a number of potential solvent molecules were added to the model. Solvent molecules were added only if they appeared with approximately spherical density in both electron density and difference density maps and had good hydrogen bonding geometry with other atoms already included in the model. Discrimination of potential peaks on the difference maps was greatly aided by the use of the program WATERSHED (Bond.C.S., personal communication). The refinement then consisted of several rounds, each of which consisted of 40 cycles of positional refinement, 20 cycles of temperature factor (individual, restrained isotropic) minimization, map calculation and examination, model manipulation and the addition of new solvent molecules coupled with the checking of those already incorporated into the model.

The refinement procedure was deemed to have converged when no further solvent molecules could be added to the model and no further manual manipulation of the model was required. At this point the *R* factor was 0.201 for the 2605 reflections with  $F > 2\sigma(F)$  in the resolution range 8.0–2.4 Å for a model consisting of the dodecanucleotide (486 atoms) and 66 solvent molecules, each of which was modelled as an oxygen atom. The data used in the refinement represents 89% of the total theoretical data available between the resolution limits stated, ~50% for the range

2.5–2.4 Å. The final model has root mean square (r.m.s.) deviations from ideality of 0.018 Å for bond lengths and 3.53° for bond angles. The average temperature factors for the phosphate groups, deoxyribose groups, bases and solvent molecules are 44.2 Å<sup>2</sup>, 35.0 Å<sup>2</sup>, 24.4 Å<sup>2</sup> and 61.4 Å<sup>2</sup> respectively. The fit of the model to the final calculated electron density is generally excellent (Fig. 2), although there are occasional discontinuities in the sugar-phosphate backbone. Both observed structure factor data and the final refined coordinates have been deposited with the Brookhaven Protein Databank (PDB) (15) and have been assigned the accession code 233d.

## RESULTS AND DISCUSSION

In the structure of the modified dodecamer the nucleotides are labelled C1–G12 on strand 1 and C13–G24 on strand 2, both in the 5'→3' direction. The 66 solvent molecules are labelled HOH25–HOH91.

As indicated by nucleotide sequence, unit cell parameters and space group, the structure of d(CGCGAASSCGCG) is isomorphous with that of the native dodecamer (9). The parameters that define global double helical conformation do not differ significantly between the two structures. A least squares superposition of the two double helices (Fig. 3) shows the r.m.s. difference in atomic positions between the two structures to be 0.49 Å and also suggests that the global conformation of the two structures is extremely similar. However, a closer inspection of the two superposed structures reveals that incorporation of the four 4'-thiosugar moieties into the sequence has led to some significant differences in the local structure of the two duplexes. This view is supported by an examination of Figure 4. Figure 4a shows the difference in position for every atom in the two duplexes. As can be seen, on each of the two strands the deviation in atomic position remains fairly constant until one reaches a point in the vicinity of the 4'-thio-2'-deoxythymidine residues, when the deviations increase to a maximum before falling away to background level. Although the magnitudes of the positional differences are somewhat greater on strand 1 than strand 2, in both cases the distribution is rather symmetrical, which indicates that these deviations are real and not artefacts of the refinement procedure. Figure 4b shows that for an average deviation per nucleotide the differences in position reach their peak for residues S8 and S20, which are the second of the two modified nucleotides on strands 1 and 2 respectively.

The above analysis refers to the superposition of the modified dodecamer onto the coordinate set 1BNA (9) as deposited in the PDB. There have been several other determinations of the structure of d(CGCGAATTCGCG), either at low temperature (deposition 2BNA in the PDB; 17) or using different refinement techniques (depositions 7BNA, 9BNA in the PDB; 18,19). The refined structure of d(CGCGAASSCGCG) has been superimposed on each of these coordinate sets (Figure 4a), yielding similar results to those found when the structure of d(CGCGAASSCGCG) is superimposed on the original 1BNA coordinate set. This is further evidence that the differences we observe between the structure of d(CGCGAASSCGCG) and that of the native dodecamer are real. Moreover, when the coordinate sets 2BNA, 7BNA and 9BNA are superimposed on those of 1BNA (Fig. 4c) there is no significant increase in the r.m.s. deviations of the sugar atoms in nucleotides 7, 8, 19 and 20 when compared with the rest of the atoms, although, inexplicably, for the 1BNA versus 9BNA superposition anomalously

ly large deviations occur for the C6 and M5 atoms of nucleotides 7, 8, 19 and 20.

Incorporation of the 4'-thiosugars into the sequence does not greatly affect the positions of the bases when both the modified and native structures are compared (Fig. 3), but rather they induce a change in the conformation of the sugar-phosphate backbone in the vicinity of the modified sugars. Indeed, the greatest deviations in the positions of the atoms in the modified and native (1BNA) dodecamers occurs in the phosphodiester linkages on the 5'-side of the thiosugars themselves (Fig. 5)

There are few other differences in structure between d(CGCGAASSCGCG) and its native analogue. However, there is an increase in helical rise at base step 6 (Fig. 6a) in the modified duplex which is coupled with a decrease in helical twist at the same base step (Fig. 6b). As the base step in question is the 5'-ApS-3' step it would appear that this local change in conformation of the duplex is a direct result of the incorporation of sulphur into each strand.

Examination (Table 1) of cross-strand phosphorus-phosphorus distances in d(CGCGAASSCGCG) appears to suggest that compared with d(CGCGAATTCGCG) the width of the minor groove increases in the vicinity of the 4'-thionucleotides. This is sensible, in that in order to accommodate the larger 4'-thio-2'-deoxyribose groups cross-strand phosphate groups will tend to be further apart. However, if we define minor groove width by measuring cross-strand O4'-O4', O4'-S4' or S4'-S4' distances (Table 2) we can see that in d(CGCGAASSCGCG) there appears to be a slight *narrowing* of the minor groove with respect to that in d(CGCGAATTCGCG). In the region of the 4-thiosugars it is tempting to say that this apparent narrowing of the groove is due to the increased van der Waals radius of sulphur compared with oxygen. However, as the narrowing is observed at all points at which the minor groove is measured, it may be that it is a result of the refinement procedure used. Given that the average difference in minor groove width is only 0.5 Å, it is probable that incorporation of 4-thiosugars into this oligonucleotide has not significantly modified the accessibility of the floor of the minor groove and that potential protein-DNA interactions in this area are unlikely to be affected.

Table 3 details the values of the sugar-phosphate backbone torsion angles in the structures of both the modified and native dodecamers. It shows that there are a number of differences in the torsion angle values for the two structures, particularly in the regions containing the 4'-thiosugars (residues S7, S8, S19 and S20). Although there is no real consistency in the differences over both strands, there is a suggestion of an increase in the value of the torsion angle  $\delta$  for the residues which have 4'-thiosugars. This is indicative of a change in the sugar conformations of the modified residues on going from the native to the modified duplex. This indication is supported if one examines the deoxyribose pseudorotation phase angles ( $P$ ) in the structure of d(CGCGAASSCGCG) (Table 1).

For all the modified sugars the average value of  $P$  is 197°, which is close to those values (178° and 180°) found in the crystal structures of two 4'-thio-2'-deoxynucleosides (3), but much higher than the average value of  $P$  (116°) found for the analogous residues in the crystal structure of d(CGCGAATTCGCG) (9). The values of  $P$  suggest (19) that the 4'-thiosugars have a conformation close to C<sub>3'</sub>-*exo*, rather than the approximate C<sub>1'</sub>-*exo* found for the analogous sugars in the native structure.

This observation is supported if we compare the pseudorotation angles in d(CGCGAASSCGCG) with those in the 2BNA (16),

7BNA (17) and 9DNA (18) native structures (Fig. 7a). In each case the pseudorotation angle for the modified sugars is greater than those of the analogous sugars in the native dodecamer, regardless of the refinement technique used or the temperature at which the native structure was determined. Moreover, the values of the pseudorotation angles in the 2BNA, 7BNA and 9BNA native structures closely mimic those of the 1BNA structure (Fig. 7a).

A surprising observation in the structure of d(CGCGAASSCGCG) is that the sugar puckering amplitude ( $\tau$ ) of the 4'-thio-2'-deoxyribose groups appears to be generally lower than those for the analogous sugars in the structures of the native oligonucleotide (Fig. 7b). The cause of our surprise is that one might reasonably expect that in 4'-thiosugars the increase in the length of the C-S bonds with respect to the C-O bonds in unmodified sugars would lead to an increase in the sugar puckering amplitude. Indeed, such an increase is observed if one compares  $\tau$  in the structures of 4'-thiothymidine (47.9°) and (*E*)-5-(2-bromovinyl)-2'-deoxy-4'-thiouridine (48.6°) (3) with that of free thymidine (37.8°) (20).

**Table 1.** A comparison of cross-strand phosphorus-phosphorus distances (Å) which define minor groove width in d(CGCGAASSCGCG) and d(CGCGAATTCGCG) (1BNA)

	d(CGCGAATTCGCG)	d(CGCGAASSCGCG)
P5-P24	8.7	8.9
P6-P23	7.2	7.2
P7-P22	5.3	5.2
P8-P21	4.2	4.9
P9-P20	4.0	5.2
P10-P19	3.1	4.1
P11-P18	5.2	5.2
P12-P17	5.1	4.9

In order to take into account the van der Waals radius of a phosphate group the distance quoted is that measured less 5.8 Å.

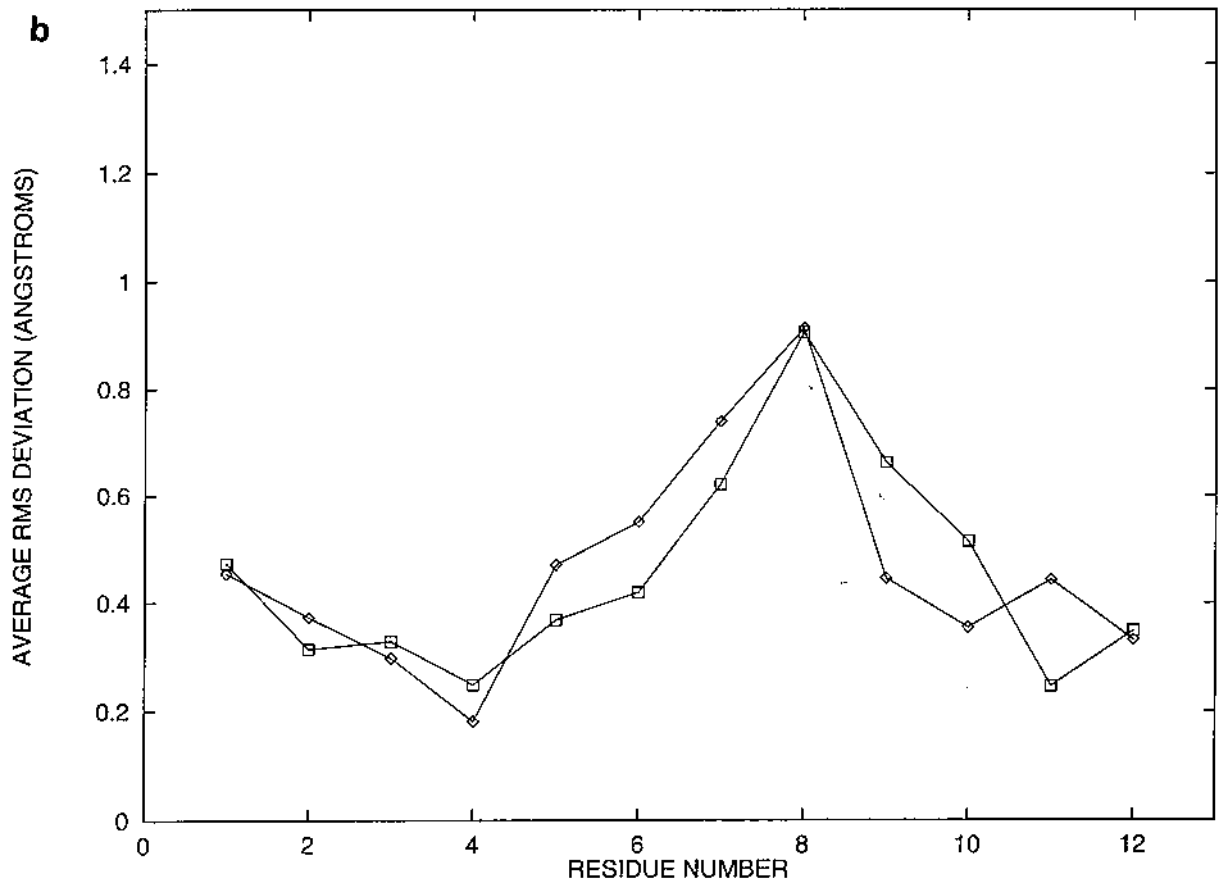
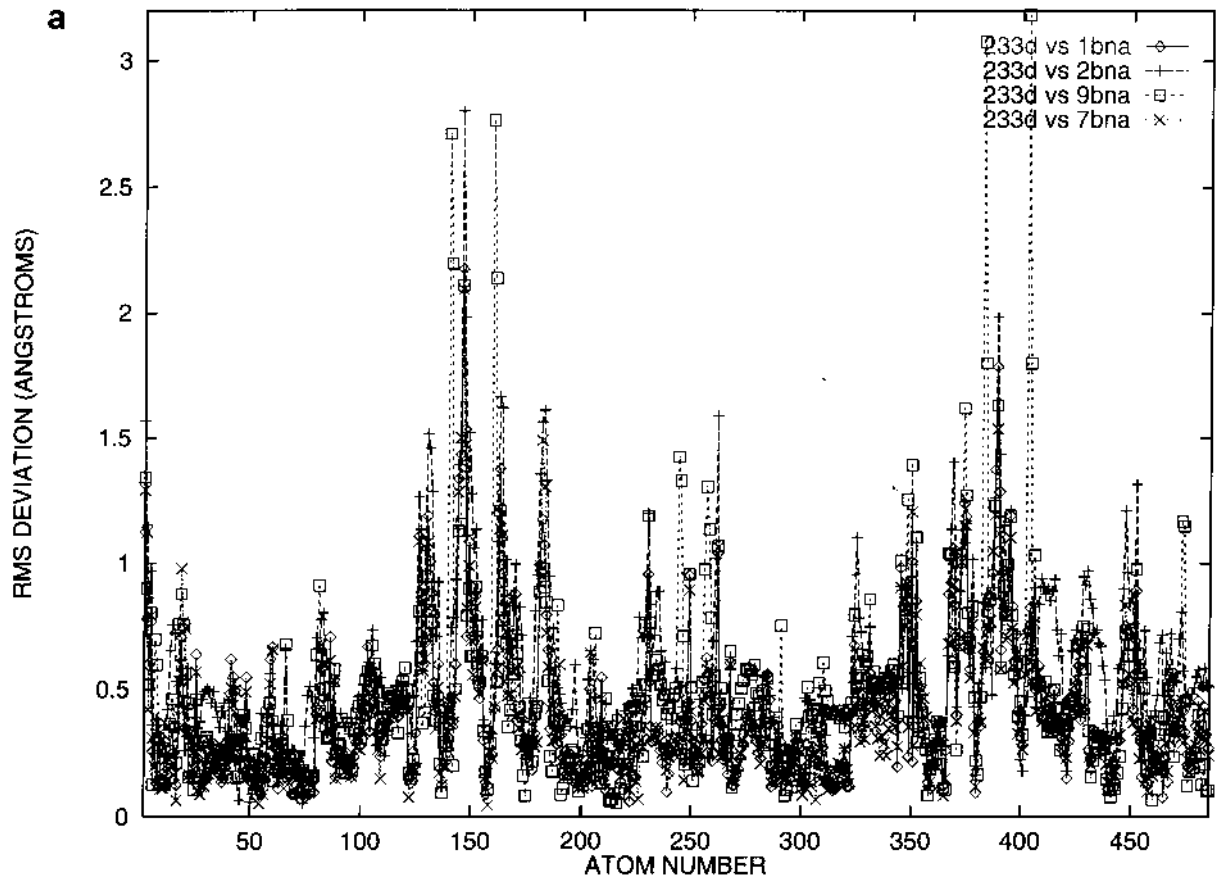
**Table 2.** A comparison of cross-strand O4'-O4', O4'-S4' and S4'-S4' distances (Å) which define minor groove width in d(CGCGAASSCGCG) and the 1BNA structure of d(CGCGAATTCGCG)

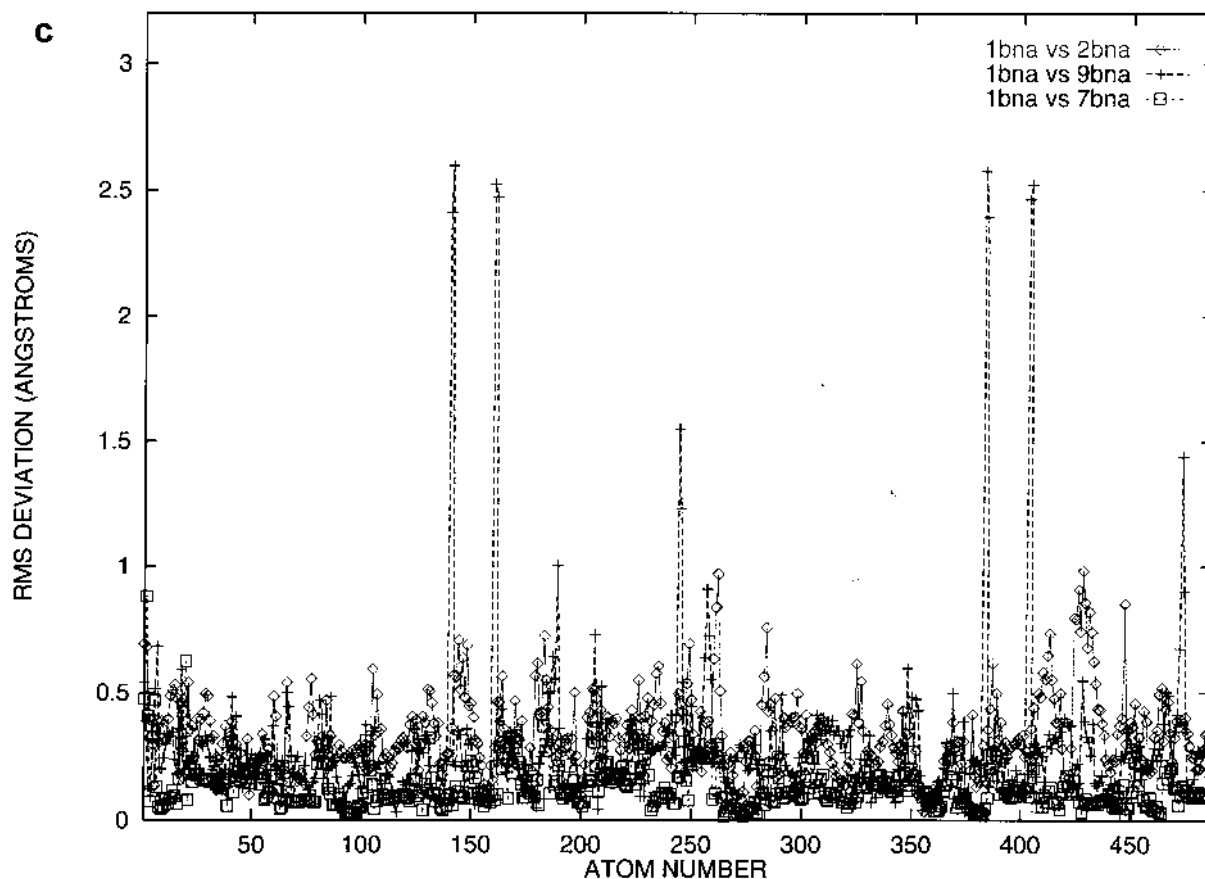
	d(CGCGAATTCGCG)	d(CGCGAASSCGCG)
O4'(4)-O4'(24)	7.0	6.8
O4'(5)-O4'(23)	6.9	6.0
O4'(6)-O4'(22)	5.9	6.1
S4'(7)-O4'(21) <sup>a</sup>	4.6	4.1
S4'(8)-S4'(20) <sup>a</sup>	4.3	3.7
O4'(9)-S4'(19) <sup>a</sup>	3.7	3.2
O4'(10)-O4'(18)	4.7	4.4
O4'(11)-O4'(17)	4.5	3.8
O4'(12)-O4'(16)	5.9	6.3

In order to take into account the van der Waals radii of the atoms the distances quoted are those measured less 2.8, 3.25 and 3.7 Å respectively.

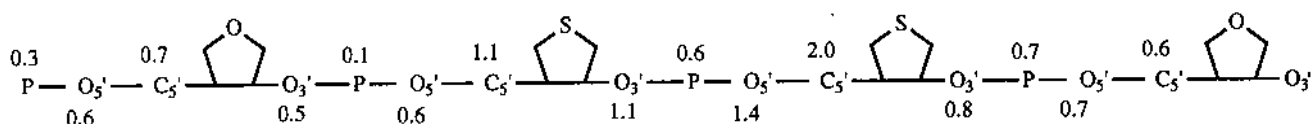
<sup>a</sup>In d(CGCGAATTCGCG) the S4' atoms are, in fact, O4' atoms

Do the observed differences between the native and the modified oligonucleotide structures help to explain the resistance to digestion by restriction endonucleases of oligonucleotides containing





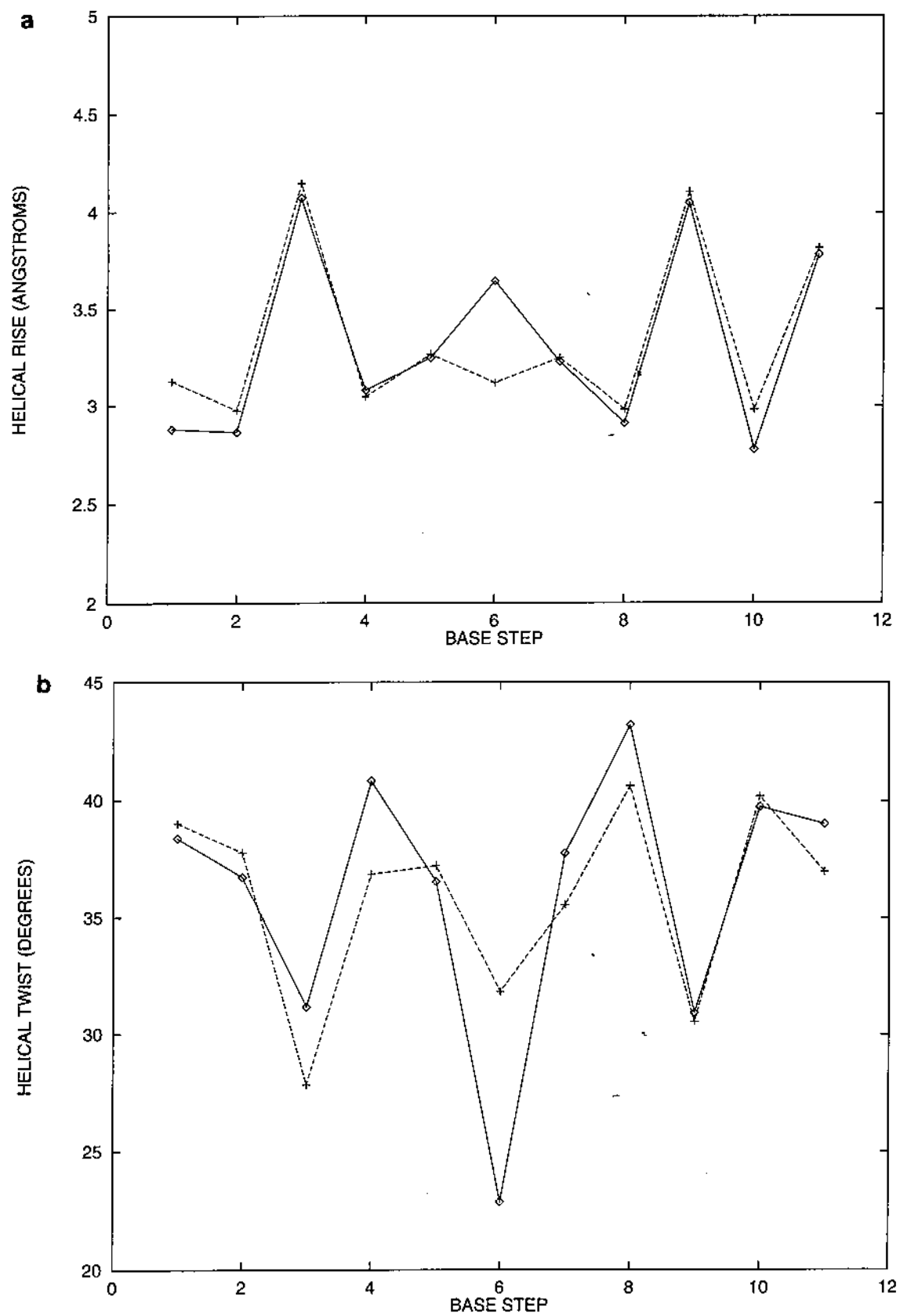
**Figure 4.** (a) A plot of the r.m.s. differences in the positions of the atoms in d(CGCGAASSCGCG) and d(CGCGAATTCGCG) resulting from a superposition of the modified and the 1BNA, 2BNA, 7BNA and 9BNA native structures. The atoms on strand 1 are numbered from 1 to 243, those on strand 2 from 244 to 486. (b) A plot showing the average r.m.s. deviation per residue resulting from the least squares superposition of the modified and the 1BNA native structures. In this figure the residues are numbered from 1 to 12 for both strands in the 5'→3' direction. (c) A plot of the r.m.s. differences in atomic position following the superposition of the 2BNA, 7BNA and 9BNA coordinate sets of d(CGCGAATTCGCG) on the 1BNA coordinate set. In both this figure and Figure 7, 233d refers to the refined coordinates of d(CGCGAASSCGCG) as deposited in the PDB.



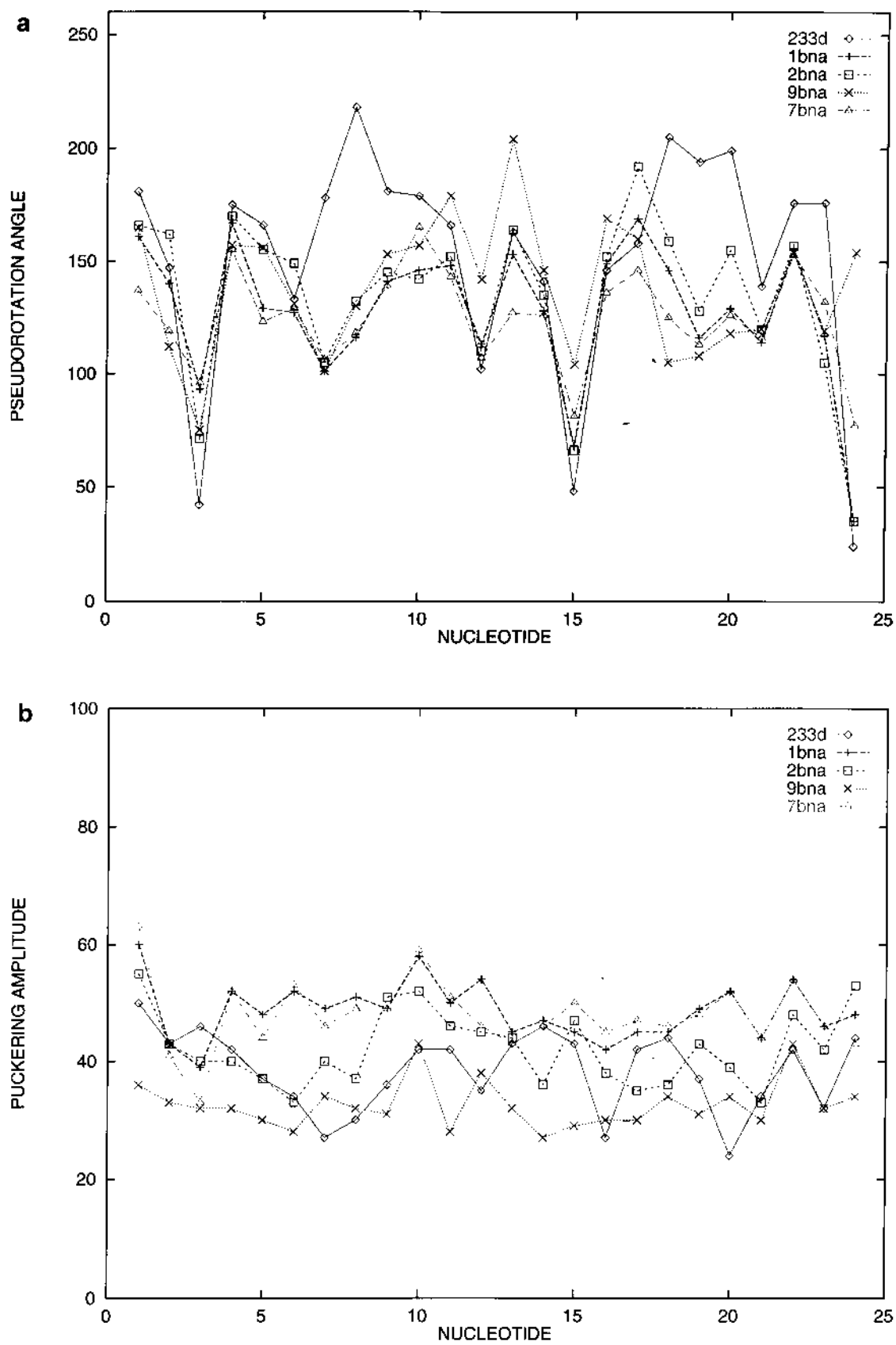
**Figure 5.** A schematic representation of the sugar-phosphate backbone of d(CGCGAASC GCG) in the region of the 4-thio-2'-deoxythymidine moieties showing the average deviation (in Å) between each atom and its counterpart in the 1BNA structure of the native d(CGCGAATTCGCG).

4'-thiosugars? The extremely high specificity of restriction enzymes is due to two major types of interaction between the enzyme and its cognate DNA sequence (19). These are: (i) *direct readout* of the base sequence of the cognate DNA resulting from interactions between the enzyme and the major groove functional groups of the DNA; (ii) *indirect readout* of the base sequence of the cognate DNA which occurs via interactions between the endonuclease and the sugar-phosphate backbone of the DNA. In addition to these modes of inducing sequence specificity there are also *non-sequence-specific interactions* between the protein and the sugar-phosphate backbone of the DNA which usually help to increase the overall binding affinity of the DNA for the protein.

In some cases these interactions can only occur if there is a distortion of the DNA sequence from the canonical B-form and it had been suggested (7) that the presence of thiosugars in oligonucleotides would 'rigidify' them and mean that the sequences could not undergo the distortion required for them to be good substrates for their restriction endonuclease. However, analysis of the structure of *EcoRV* endonuclease complexed with a thiosugar-modified cognate DNA sequence (Kostrewa, D., Hancox, E.L., Walker, R.T., Connolly, B.A. and Winkler, F.K., unpublished results) has shown that the DNA adopts the same kinked structure observed in the complex of the endonuclease with an unmodified cognate DNA sequence (8), but, crucially, that the  $Mg^{2+}$  ion, which is an absolute requirement for endonuclease activity, is not present.



**Figure 6.** Comparison of individual values of (a) helical rise and (b) helical twist in the structures of d(CGCGAASSCGCG) (◊) and the 1BNA structure of d(CGCGAATTCGCG) (+).



**Figure 7.** Comparisons of the values of (a) pseudorotation angle and (b) sugar pucker amplitude ( $\tau$ ) found in the structure of d(CGCGAASSCGCG) and in the 1BNA, 2BNA, 7BNA and 9BNA structures of d(CGCGAATTCGCG).

**Table 3.** Sugar-phosphate backbone and glycosyl torsion angles, pseudorotation phase angles (*P*) and sugar pucker amplitudes ( $\tau$ ) for the structures of d(CGCGAASSCGCG) and d(CGCGAATTCGCG) (1BNA)

Residue	$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	$\zeta$	$\chi$	<i>P</i>	$\tau$
C1			154	172	213	200	272	181	50
			<i>174</i>	<i>157</i>	<i>218</i>	<i>216</i>	<i>255</i>	<i>161</i>	<i>60</i>
G2	312	152	36	144	207	221	248	147	43
	<i>295</i>	<i>170</i>	<i>40</i>	<i>128</i>	<i>174</i>	<i>262</i>	<i>249</i>	<i>140</i>	<i>43</i>
C3	312	134	58	83	199	260	215	42	46
	<i>298</i>	<i>172</i>	<i>59</i>	<i>98</i>	<i>183</i>	<i>272</i>	<i>225</i>	<i>93</i>	<i>39</i>
G4	312	163	49	159	199	215	268	175	42
	<i>297</i>	<i>180</i>	<i>57</i>	<i>155</i>	<i>205</i>	<i>208</i>	<i>268</i>	<i>167</i>	<i>52</i>
A5	304	164	49	147	198	248	250	166	37
	<i>317</i>	<i>143</i>	<i>53</i>	<i>119</i>	<i>180</i>	<i>268</i>	<i>234</i>	<i>129</i>	<i>48</i>
A6	286	171	56	119	186	267	247	133	34
	<i>287</i>	<i>180</i>	<i>66</i>	<i>121</i>	<i>174</i>	<i>272</i>	<i>238</i>	<i>127</i>	<i>52</i>
S7	273	171	66	138	261	175	288	178	27
	<i>303</i>	<i>181</i>	<i>52</i>	<i>99</i>	<i>174</i>	<i>274</i>	<i>232</i>	<i>101</i>	<i>49</i>
S8	69	229	195	151	205	214	262	218	30
	<i>301</i>	<i>174</i>	<i>64</i>	<i>109</i>	<i>170</i>	<i>271</i>	<i>234</i>	<i>116</i>	<i>51</i>
C9	56	132	318	159	198	256	266	181	36
	<i>302</i>	<i>180</i>	<i>61</i>	<i>129</i>	<i>203</i>	<i>266</i>	<i>241</i>	<i>141</i>	<i>49</i>
G10	292	183	40	156	265	145	281	179	42
	<i>293</i>	<i>169</i>	<i>47</i>	<i>143</i>	<i>257</i>	<i>150</i>	<i>270</i>	<i>146</i>	<i>58</i>
C11	286	145	45	147	200	258	251	166	42
	<i>286</i>	<i>139</i>	<i>56</i>	<i>136</i>	<i>198</i>	<i>271</i>	<i>235</i>	<i>148</i>	<i>50</i>
G12	287	177	51	97			262	102	35
	<i>279</i>	<i>176</i>	<i>57</i>	<i>111</i>			<i>248</i>	<i>113</i>	<i>54</i>
C13			57	156	251	193	238	163	43
			<i>56</i>	<i>137</i>	<i>201</i>	<i>234</i>	<i>232</i>	<i>153</i>	<i>45</i>
G14	298	137	43	145	203	218	248	141	46
	<i>309</i>	<i>164</i>	<i>49</i>	<i>122</i>	<i>178</i>	<i>267</i>	<i>244</i>	<i>128</i>	<i>47</i>
C15	316	129	62	82	200	262	226	48	43
	<i>297</i>	<i>169</i>	<i>61</i>	<i>86</i>	<i>175</i>	<i>262</i>	<i>226</i>	<i>68</i>	<i>45</i>
G16	298	157	73	123	151	283	248	146	27
	<i>291</i>	<i>171</i>	<i>73</i>	<i>136</i>	<i>174</i>	<i>262</i>	<i>245</i>	<i>149</i>	<i>42</i>
A17	286	198	58	151	195	213	252	158	42
	<i>303</i>	<i>190</i>	<i>54</i>	<i>146</i>	<i>177</i>	<i>263</i>	<i>254</i>	<i>169</i>	<i>45</i>
A18	56	181	282	193	225	200	254	205	44
	<i>303</i>	<i>186</i>	<i>48</i>	<i>130</i>	<i>174</i>	<i>259</i>	<i>251</i>	<i>146</i>	<i>45</i>
S19	297	132	48	141	238	217	286	194	37
	<i>302</i>	<i>173</i>	<i>60</i>	<i>109</i>	<i>179</i>	<i>272</i>	<i>228</i>	<i>116</i>	<i>49</i>
S20	285	129	68	134	197	265	267	199	24
	<i>301</i>	<i>179</i>	<i>55</i>	<i>122</i>	<i>179</i>	<i>266</i>	<i>240</i>	<i>129</i>	<i>52</i>
C21	278	196	53	126	191	264	258	139	34
	<i>301</i>	<i>185</i>	<i>45</i>	<i>110</i>	<i>183</i>	<i>274</i>	<i>246</i>	<i>114</i>	<i>44</i>
G22	292	180	41	156	240	172	288	176	42
	<i>293</i>	<i>179</i>	<i>50</i>	<i>150</i>	<i>260</i>	<i>171</i>	<i>271</i>	<i>156</i>	<i>54</i>
C23	304	161	28	133	189	245	260	176	32
	<i>289</i>	<i>138</i>	<i>45</i>	<i>113</i>	<i>186</i>	<i>263</i>	<i>235</i>	<i>117</i>	<i>46</i>
G24	310	143	55	73			223	24	44
	<i>295</i>	<i>171</i>	<i>47</i>	<i>79</i>			<i>225</i>	<i>35</i>	<i>48</i>

Values for d(CGCGAASSCGCG) are given on the first line and for d(CGCGAATTCGCG) on the second line in italic. All angles are in degrees. All parameters were calculated using the NEWHEL92 program distributed by R.E. Dickerson through the Brookhaven Protein Data Bank (15).

This observation, coupled with the fact that our analysis of the structure of d(CGCGAASSCGCG) reveals that the major effect of incorporation of 4'-thio-2'-deoxyribose groups into a DNA duplex is to change its backbone conformation, suggests the following reasons for the resistance to nuclease digestion of oligonucleotides containing 4'-thiosugars. (i) Changes in the backbone conformation result in the enzyme being able to bind less well to the modified

DNA than to its unmodified cognate sequence, thus causing a reduction in activity of the endonuclease; (ii) changes in backbone conformation resulting from DNA modification cause a reduction in Mg<sup>2+</sup> ion binding, which again would result in a reduction in the activity of the enzyme. The decision as to which, if either, of these two hypotheses is correct would require further analysis of the structures of 4'-thiosugar-modified DNAs, both in isolation and,

more importantly, complexed to their specific restriction endonucleases.

We conclude with the observation that the inclusion of 4'-deoxy-2'-thiothymidine into DNA has a small, but detectable, effect on the backbone conformation of a DNA duplex. On occasions this may appear to be of little significance, but should it be associated with the conferring of increased nuclease resistance or increased lipophilicity it may make such oligodeoxynucleotides candidates for antisense molecules.

## ACKNOWLEDGEMENTS

This work was supported by the United Kingdom Engineering and Physical Sciences Research Council, the Biotechnology and Biology Science Research Council and the Wellcome Trust. We are grateful to the Wellcome Foundation for financial support (to ELH). We thank Drs Miroslav Papiz and Pierre Rizkallah of the Synchrotron Radiation Source, Daresbury Laboratory, for the allocation of contingency beam time, part of which was used to collect the data described here. TB wishes to thank the Royal Society of Edinburgh for a Caledonian Research Fellowship.

## REFERENCES

- Dyson, M.R., Coe, P.L. and Walker, R.T. (1991) *Carbohydrate Res.*, **216**, 237–248.
- Secrist, J.A., Tiwari, K.N., Riordan, J.M. and Montgomery, J.A. (1991) *J. Med. Chem.*, **34**, 2361–2366.
- Dyson, M.R., Coe, P.L. and Walker, R.T. (1991) *J. Med. Chem.*, **34**, 2782–2786.
- Basnak, I., Hancox, E.L., Connolly, B.A. and Walker, R.T. (1993) *Nucleic Acids Symp. Ser.*, **29**, 101–102.
- Walker, R.T., Whale, R.F., Dyson, M.R., Coe, P.L., Alderton, W., Collins, P., Ertl, P., Lowe, D., Rahim, S.G., Snowden, W and Littler, E. (1994) *Nucleic Acids Symp. Ser.*, **31**, 9–10.
- Koole, L.H., Plavec, J., Liu, H., Vincent, B.R., Dyson, M.R., Coe, P.L., Walker, R.T., Hardy, G.W., Rahim, S.G. and Chattapodhaya, J. (1992) *J. Am. Chem. Soc.*, **114**, 9936–9943.
- Hancox, E.L., Connolly, B.A. and Walker, R.T. (1993) *Nucleic Acids Res.*, **21**, 3485–3491.
- Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.K., Martin, P.D., Petratos, K. and Wilson, K.S. (1993) *EMBO J.*, **12**, 1781–1795.
- Wing, R., Drew, H.R., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) *Nature*, **287**, 755–758.
- Leslie, A.G.W., Brick, P. and Wonacott, A.J. (1986) *CCP4 Newslett.*, **18**, 33–39.
- Collaborative Computational Project Number 4. (1994) *Acta Crystallogr.*, **D50**, 760–763.
- Brünger, A.T. (1990) *X-PLOR (V2.2) Manual*. Howard Hughes Medical Institute, Yale University, New Haven, CT.
- Brünger, A.T., Krukowski, A. and Erickson, J.W. (1990) *Acta Crystallogr.*, **A46**, 585–593.
- Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) *Acta Crystallogr.*, **A47**, 110–119.
- Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. and Weng, J. (1987) In Allen, F.H., Bergerhoff, G. and Seivers, R. (eds), *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*. Data Commission of the International Union of Crystallography, Cambridge, UK.
- Drew, H.R., Samson, S. and Dickerson, R.E. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4040–4044.
- Holbrook, S.R., Dickerson, R.E. and Kim, S.-H. (1985) *Acta Crystallogr.*, **B41**, 255–262.
- Westhof, E. (1987) *J. Biomol. Struct. Dyn.*, **5**, 581–600.
- Saenger, W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York, NY.
- Young, D.W., Tollin, P. and Wilson, H.R. (1969) *Acta Crystallogr.*, **B25**, 1423.