Nucleic Acids Research, Vol. 20, No. 18 4753–4759

The conformational variability of an adenosine·inosine base-pair in a synthetic DNA dodecamer

Gordon A.Leonard[1,2]*, Ewan D.Booth[1], William N.Hunter[2] and Tom Brown[1]
[1]Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ and [2]Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received July 2, 1992; Revised and Accepted August 21, 1992

ABSTRACT

A crystal structure analysis of the synthetic deoxydodecamer d(CGCAAATTIGCG) which contains two adenosine·inosine (A·I) mispairs has revealed that, in this sequence, the A·I base-pairs adopt a A(*anti*)·I(*syn*) configuration. The refinement converged at R = 0.158 for 2004 reflections with F ≥ 2σ(F) in the range 7.0 – 2.5Å for a model consisting of the DNA duplex and 71 water molecules. A notable feature of the structure is the presence of an almost complete spine of hydration spanning the minor groove of the whole of the $(AAATTI)_2$ core region of the duplex. pH-dependent ultraviolet melting studies have suggested that the base-pair observed in the crystal structure is, in fact, a protonated AH⁺(*anti*)·I(*syn*) species and that the A·I base-pairs in the sequence studied display the same conformational variability as A·G mispairs in the sequence d(CGCAAATTGGCG). The AH⁺(*anti*)·I(*syn*) base-pair predominates below pH 6.5 and an A(*anti*)·I(*anti*) mispair is the major species present between pH 6.5 and 8.0. The protonated base-pairs are held together by two hydrogen bonds one between N6(A) and O6(I) and the other between N1(A) and N7(I). This second hydrogen bond is a direct result of the protonation of the N1 of adenosine. The ultraviolet melting studies indicate that the A(*anti*)·I(*anti*) base-pair is more stable than the A(*anti*)·G(*anti*) base-pair but that the AH⁺(*anti*)·I(*syn*) base pair is less stable than its AH⁺(*anti*)·G(*syn*) analogue. Possible reasons for this observation are discussed.

INTRODUCTION

The conformation of the guanosine·adenosine (G·A) base-pair has been extensively studied. Proofreading enzymes remove this base pair relatively inefficiently (1, 2) and it has been shown that this base-pair exhibits significant conformational variability as a function of both pH and base stacking environment. For example, in the sequence d(CCAAGATTGG) (3) and the non self-complementary d(GCCACAAGCTC)·d(GAGCTGGTGGC) (4) the base-pairs adopt a G(*anti*)·A(*anti*) configuration. In the sequence d(CGCGAATTAGCG) (5) a G(*anti*)·A(*syn*) conformation has been found while in the sequences d(CGAGAATTCGCG) and d(CGCAAATTGGCG) it has been shown that the base-pair adopts a G(*anti*)·A(*anti*) configuration at high pH and a protonated G(*syn*)·AH⁺(*anti*) conformation at more acidic pH (6–9).

Deoxyinosine (I) has chemical properties very like guanosine although the lack of a N2 amino group means that it can pair with cytosine (C), thymine (T) or adenosine (A) without destabilising a duplex even when several inosine-containing base-pairs are present in a short stretch of DNA (10, 11). This is utilised by some tRNA molecules which have I at the 5′-hydroxyl ends of their anticodons enabling pairing with A, C or U and also in the middle position of an anticodon to pair with A (12).

There have been few crystal structure analyses of DNA duplexes containing inosine: only an I·C base-pair in a Z-DNA octamer (13), an I·T wobble base-pair in the A-DNA duplex $d(GGIGCTCC)_2$ (14) and an I(*anti*)·A(*syn*) base-pair the B-DNA duplex d(CGCIAATTAGCG) (15) have been reported. As part of our continuing study into the nature of purine·purine base-pairs and in order to further investigate the nature of inosine-containing base-pairs we have carried out an analysis of the synthetic deoxydodecamer d(CGCAAATTIGCG) combining single-crystal X-ray techniques and ultraviolet melting methods.

EXPERIMENTAL METHODS

Chemical synthesis

The self-complementary dodecanucleotide d(CGCAAATTIGCG) was synthesised on an ABI 380B DNA synthesiser using the phosphoramidite method (6×1μmole preps). The monomer 5′-dimethoxytrityldeoxyinosine-3′-cyanoethylphoshoramidite was purchased from Cruachem. The oligomer was purified by ion-exchange high-pressure liquid chromatography (HPLC) followed by reversed-phase HPLC and Sephadex gel filtration. This protocol yielded approximately 10mg of pure product which eluted as a single peak when injected on analytical reversed-phase HPLC. (Ion-exchange gradient: 0.04M to 0.67M potassium

* To whom correspondence should be addressed at: Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK

phosphate buffer pH 6.4 in 20% acetonitrile, 30 minutes. Reversed-phase gradient: 0.1M ammonium acetate buffer, 0% to 20% acetonitrile, 30 minutes).

Ultraviolet melting studies

Melting curves were measured at 264nm on a Perkin–Elmer Lambda 15 ultraviolet spectrometer equipped with a Peltier block and controlled by an IBM PS/2 microcomputer. In all cases a heating rate of 1°/minute was used and the curves were measured in triplicate, the data being stored and processed using the PECSS2 software package. The oligonucleotides were dissolved in a buffer consisting of aqueous potassium dihydrogen orthophosphate (0.1M) and EDTA (1mM) which had been adjusted to the appropriate pH by the addition of sodium hydroxide.

Crystallisation, X-ray data collection and structure refinement

Crystals were grown in sitting drops at 4°C from a solution containing d(CGCAAATTIGCG) (0.5mM), sodium cacodylate buffer (15mM, pH 6.5), magnesium chloride (40mM), spermine tetrahydrochloride (0.3mM) and 2-methyl-2,4-pentanediol (MPD) (14% vol/vol) equilibrated against 50% vol/vol aqueous MPD.

Two crystals were mounted in thin-walled glass capillaries and used for X-ray data collection on a Siemens AED2 4-circle diffractometer using CuK_α radiation ($\lambda=1.5418$Å), a graphite monochromator, a long arm and a helium path. For both crystals data were measured at 4°C using ω-scans with a scan width of 1° and scan speeds ranging from 2°/minute to 0.17°/minute depending on the results of a prescan. For the first crystal of size 1.0×0.2×0.2mm a total of 4866 reflections were measured. These reduced to 2307 unique reflections after first omitting those reflections with $F<\sigma(F)$ and merging ($R_{merge}=0.09$). The second crystal which was of a similar size was then used to measure 5426 reflections of which 1745 were unique after omission of those data with $F<\sigma(F)$ and merging ($R_{merge}=0.09$). The two data sets were then merged to give 2462 unique reflections ($R_{merge}=0.13$) to a resolution of 2.25Å. However, since only 10% of the reflections in the resolution range 2.5–2.25Å had $F\geq 2\sigma(F)$, it was decided to restrict our analysis to a resolution limit of 2.5Å.

Orthorhombic unit cell parameters of a=25.16Å, b=41.06Å, c=65.45Å and space group $P2_12_12_1$ indicated that the structure was quasi-isomorphous with the structure of the native dodecamer d(CGCGAATTCGCG) (16,17) and the starting model used in the refinement was the native structure which had been subjected to several cycles of idealisation of geometry using NUCLSQ (18). The starting model was refined as a rigid-body using a modified version of SHELX (19). Starting with data in the region 10.0–7.0Å the resolution limit was increased in steps of 1Å until all data with $F > 0$ in the range 10.0–3.0Å were included. This part of the refinement converged at R=0.41 for 1409 reflections. The nucleotides G4, C9, G16 and C21 were then removed from the structure factor calculations and the refinement was continued, extending the resolution to 2.5Å, using Hendrickson-Konnert techniques (20) with the program NUCLSQ (18). Six cycles of positional refinement reduced R from 0.43 to 0.38 for 2090 reflections with $F\geq 2\sigma(F)$.

Electron density ($2F_o-F_c$) and difference (F_o-F_c) maps were then calculated and examined on an Evans and Sutherland ESV graphics workstation using the program FRODO (21). These maps indicated that the conformation of both mispairs was A(*anti*)·I(*syn*). The relevant changes were made to the model and the refinement continued with the *anti·syn* base pairs included in the structure factor calculations, the inclusion of isotropic thermal parameters and the progressive addition of solvent molecules. The solvent molecules were included into the model on the criteria of good spherical density on the F_o-F_c maps coupled with good $2F_o-F_c$ electron density and acceptable hydrogen bonding distances and geometry. Occasionally a peak was accepted as a solvent molecule if it was too far from any other atoms in the model to form hydrogen bonds if the density on both types of map warranted it. At regular intervals the A·I mispairs were removed from the structure factor calculations and the resulting fragment difference maps were examined to ensure that the conformations of the base-pairs were correct.

At the end of the refinement procedure a total of 71 solvent molecules had been included into the model and the final crystallographic residual was R=0.158 for 2004 reflections with $F\geq 2\sigma(F)$ in the resolution range 7.0–2.5Å.

The completeness of the data at the $2\sigma(F)$ level was as follows: 7.0–4.25Å, 93%; 4.25–3.05Å, 87%; 3.05–2.65Å, 77%; 2.65–2.5Å, 50%. As has already been mentioned, the intensity fell off sharply at about 2.5Å resolution. This behaviour is typical of DNA sequences based on the native d(CGCGAATTCGCG) (22).

The geometry of the final model is excellent with average deviations from ideality of 0.011Å for both sugar-base distances and phosphate distances. For angle distances the average deviations are 0.028Å and 0.025Å respectively. A Luzatti plot (not shown) (23) indicates that the coordinate error in our structure ranges from 0.2Å to 0.3Å with the largest errors likely

Table 1. Sugar/phosphate backbone and glycosyl torsion angles (°) for d(CGCAAATTIGCG).

Residue	χ	α	β	γ	δ	ξ	ζ
C1	−106	–	–	301	146	−164	−135
G2	−74	−29	166	22	164	−145	−172
C3	−139	−38	129	43	84	−179	−87
A4	−88	−69	182	72	158	−189	−95
A5	−93	−39	198	12	177	−133	−156
A6	−108	−63	148	53	127	−166	−122
T7	−107	−6	148	12	125	−160	−114
T8	−128	−86	153	79	121	−146	−165
I9	86	−10	129	17	136	−151	−129
G10	−103	−45	145	41	151	−113	−197
C11	−107	−53	139	46	147	−160	−91
G12	−67	−58	201	1	146	–	–
C13	−109	–	–	310	162	−157	−136
G14	−115	−43	142	48	137	−137	−132
C15	−145	−59	131	67	71	−163	−90
A16	−105	−55	168	60	133	−186	−112
A17	−96	35	178	−44	170	−178	−114
A18	−84	−12	178	−2	156	−184	−110
T19	−131	−54	160	65	105	−174	−120
T20	−94	22	153	−20	148	−164	−130
I21	90	−28	160	16	142	−130	−149
G22	−104	−50	147	38	123	−151	−141
C23	−125	−41	152	23	103	−163	−84
G24	−110	13	223	−52	156	–	–

Main chain torsion angles are defined by

α β γ δ ξ ζ

O3′--P--O5′--C5′--C4′--C3′--O3′--P--O5′

The glocosyl torsion angle χ is defined by

O4′-C1′-N1-C2 for pyrimidines

O4′-C1′-N9-C4 for purines

to be on the flexible sugar/phosphate backbone, the smallest on the bases. The refined coordinates have been deposited with the Brookhaven Database.

RESULTS AND DISCUSSION

The Structure of d(CGCAAATTIGCG)

The nucleotides are labelled C1 to G12 in the 5′ to 3′ direction on strand 1 and C13 to G24 (5′ to 3′) on strand 2. The 71 solvent molecules are labelled W25 to W95.

The B-DNA duplex contains 10 standard Watson–Crick base-pairs and two A(*anti*)·I(*syn*) mispairs at positions A4·I21 and A16·I9. As with other purine·purine *anti·syn* base-pairs for which crystal structures are known (5,8,15) the A(*anti*)·I(*syn*) mispairs are incorporated into the double helix without causing any major perturbations. The average helical rise of 3.4Å compares to a value of 3.3Å for the native dodecamer while the global twist changes from 37° to 36°. The separations between adjacent phosphorous atoms along each strand range from 7.2Å to 6.0Å with average distance of 6.6Å almost the same as found

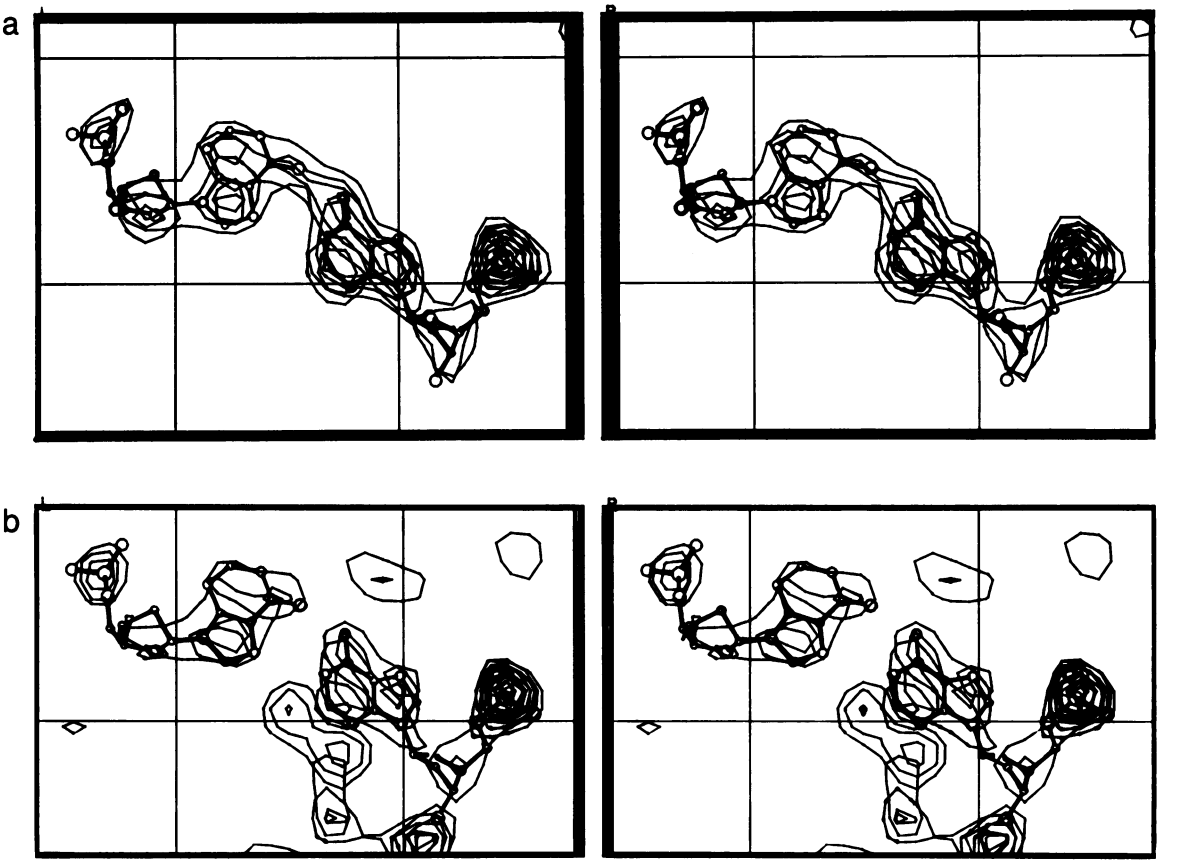


Figure 1. (a) A stereoview of the A(4)·I(21) base-pair superimposed on the relevant sections of a 'fragment' F_o-F_c map. All of the atoms shown in this figure were omitted from the structure factor calculations and the map thus represents an unbiased representation of the electron density. The minimum contour level is at approximately twice the root mean square (r.m.s) deviation fromthe average density on the map. (b) The same base-pair superimposed on the relevant sections of a $2F_o-F_c$ map in which all atoms were included in the structure factor calculations. The minimum contour level is at approximately one r.m.s. deviation from the average density on the map.

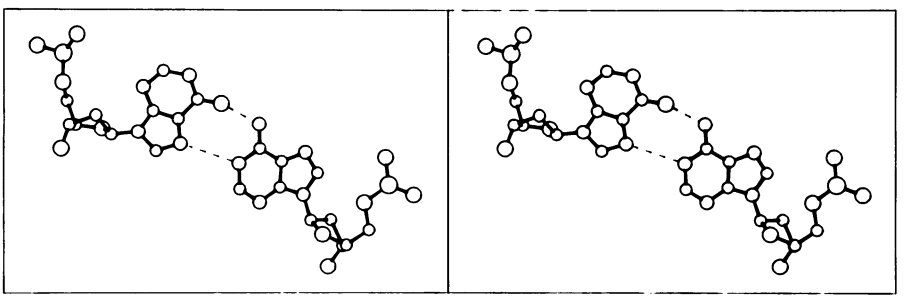


Figure 2. A representation of the A(*anti*)·I(*syn*) base-pairs showing the two interbase distances indicative of the formation of hydrogen bonds. The distances for both the A(4)·I(21) and A(16)·I(9) base-pairs are given in the text.

in the native structure (6.7Å). The sugar/phosphate backbone torsion angles (Table 1) are generally within the range expected for B-DNA (16,17,24) with the exception of the nucleotide A18. However, the abnormal value of γ for this nucleotide may well be the result of disorder in the backbone and is probably not meaningful. All helical parameters, base-pair geometries and torsion angles quoted in this paper were calculated using the program NEWHELIX distributed by R.E.Dickerson with nomenclature according to Dickerson et al., (25).

In a similar fashion to other DNA sequences containing runs of adenines, notably d(CGCAAATTTGCG) (26), d(CGCAAAAATGCG)·d(CGCTTTTTAGCG) (27) and d(CGCAAAAAAGCG)·d(CGCTTTTTTGCG) (28) there seems to be a network of inter base-pair bifurcated hydrogen bonds running through the major groove of the central adenosine-rich core of the duplex. The N6(A) to O4(T) distances range from 2.9Å to 3.4Å and are well within the range expected for the formation of this type of hydrogen bond.

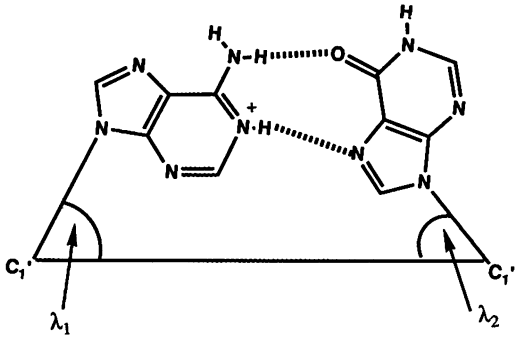


Figure 3. The definition of λ_1 and λ_2 for the A(*anti*)·I(*syn*) base-pairs.

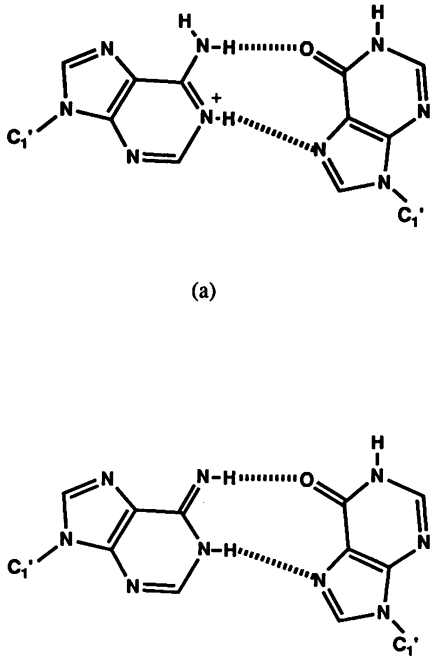


Figure 4. The two possible A(*anti*)·I(*syn*) base-pairing schemes consistent with the results of our crystal structure analysis and with the formation of two interbase hydrogen bonds. (a) The protonated AH^+(*anti*)·I(*syn*) base-pair in which both bases adopt their major tautomeric forms. (b) The A(imino,*anti*)·I(*syn*) base-pair. Here the adenosine is in its rare imino tautomeric form.

The A(*anti*)·I(*syn*) base-pairs

Given the accuracy of our analysis the geometry of the two mispairs is consistent. Figure 1(a) presents the A4·I21 base-pair superimposed on a fragment F_o-F_c map. Figure 1(b) shows the same base-pair superimposed on a $2F_o-F_c$ electron density map. The orientations of the bases are unambiguous. There are two interbase distances indicative of the formation of hydrogen bonds and these are shown in Figure 2. The two distances are N6(A)−O6(I) (2.3Å for the 4·21 pair and 2.8Å for the 16·9 base-pair) and N1(A)−N7(I) (3.2Å and 3.1Å respectively). The propellor twists of the base-pairs are 10° (4·21) and 15° (16·9) while C1′−C1′ distances of 11.5Å and 11.0Å respectively are close to those values found in both the native dodecamer (16,17) and other sequences containing purine·purine *anti*·*syn* base-pairs (5,8,15).

As for all purine·purine *anti*·*syn* base-pairs and purine·pyrimidine wobble mispairs there is a significant asymmetry in the angles between the glycosidic bonds and the C1′−C1′ vector. These are designated λ_1 and λ_2 (Figure 3). In Watson−Crick base-pairs both λ_1 and λ_2 have similar values which fall within a narrow range (52°−62°) and thus the base-pairs have an element of psuedosymmetry. For the A(*anti*)·I(*syn*)

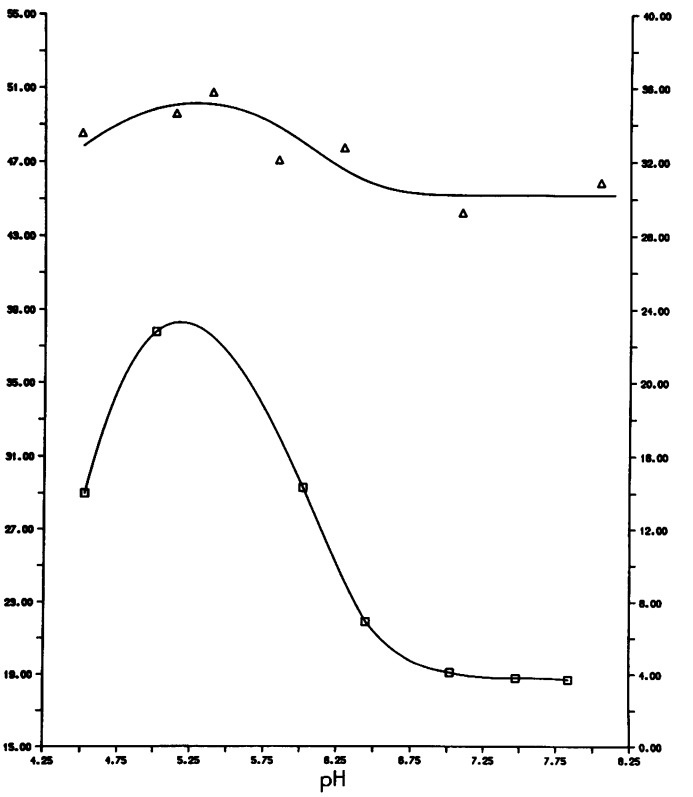


Figure 5. pH-dependence of the duplex melting temperature in O.1M phosphate buffer. Δ: d(CGCAAATTIGCG), temperature scale (°C) on the right-hand side. □: d(CGCAAATTGGCG) temperature scale (°C) on the left hand side.

pairs observed here this psuedosymmetry is lost with values of λ_1 and λ_2 of 49° and 31° (4·21) and 58° and 35° (16·9).

The resolution of this X-ray structure determination means that we cannot locate the positions of hydrogen atoms so at least two tautomeric forms for the adenosine are consistent with the A(*anti*)·I(*syn*) base-pairing that we observe. If the adenosine adopts its major tautomeric form (Figure 4(a)) then protonation of N1(A) is required to allow any hydrogen bonding between N1(A) and N7(I). No protonation is required if the adenosine is in its rare imino tautomeric form (Figure 4(b)). Ultraviolet melting studies in the pH range 8.0 to 4.5 (Figure 5) indicate that the adenosine is in its major tautomeric form and that N1(A) is indeed protonated. The shape of the stability profile of the sequence d(CGCAAATTIGCG) containing the A·I base-pairs is the same as that of the sequence d(CGCAAATTGGCG) which contains two A·G mispairs. Below pH 6.5 the A·G mispair in the sequence studied has been shown to be the protonated AH$^+$(*anti*)·G(*syn*) base-pair (6,7,9) while from pH 6.5 to 8.0 it adopts an A(*anti*)·G(*anti*) configuration. Hence, by direct analogy with previous X-ray, NMR and ultraviolet melting studies we can say that the A·I base-pair that we have observed in our crystal structure analysis of d(CGCAAATTIGCG) is an AH$^+$(*anti*)·I(*syn*) base-pair and that the A·I base-pairs in this sequence display similar conformational variability to the A·G mispair in the sequence d(CGCAAATTGGCG).

The pH-dependent stability profiles for both d(CGCAAATTIGCG) and d(CGCAAATTGGCG) (Figure 5) indicate that at pH 6.5—the pH at which the crystallisation of d(CGCAAATTIGCG) was carried out—there is almost a 1:1 ratio of the protonated *anti·syn* and non-protonated *anti·anti* base-pairs in solution. The results of our structure refinement show that at this pH there is preferential crystallisation of the *anti·syn* base-pairs and we see no evidence of the *anti·anti* pairs. This may be an artefact of the crystal packing requirements of the duplex in which the base-pairs are contained. We are however crystallising both the A·G and A·I sequences at pH 8.0 in the hope that this will result in crystals containing *anti·anti* base-pairs.

Hydration

The most notable feature of the solvent environment of the double helix d(GCGAAATTIGCG)$_2$ is the presence of an almost complete spine of hydration in the minor groove of the duplex similar to that found in the structure of the native dodecamer (29). This is shown in Figure 6(a) and, more schematically, in Figure 6(b). We had thought that the two A(*anti*)·I(*syn*) base pairs in the duplex might curtail the extent of the solvent spine to the (AATT)$_2$ region of the double helix as both the *syn* inosine bases have C8 hydrogen atoms protruding into the minor groove rather than the hydrogen bond acceptor groups that would

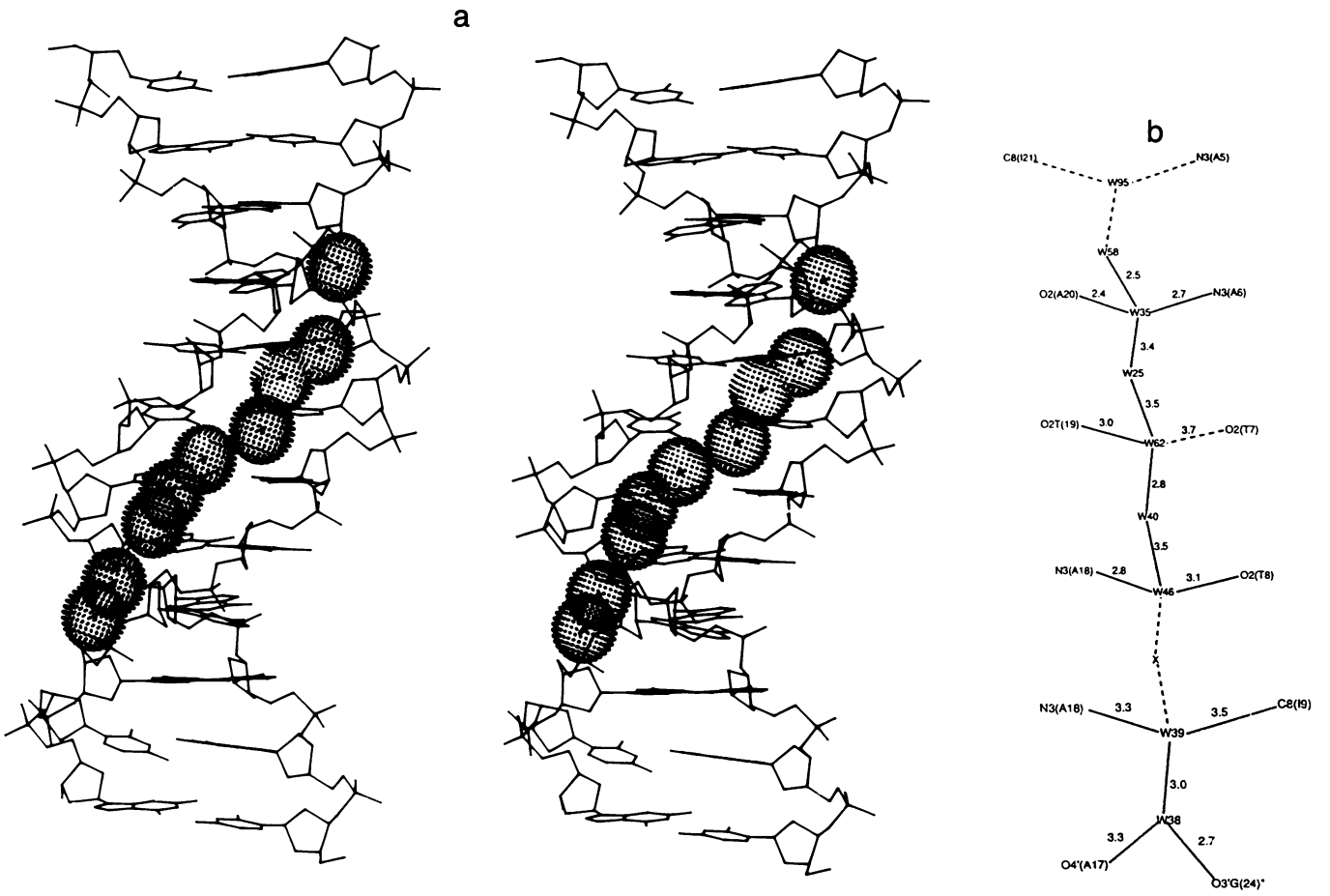
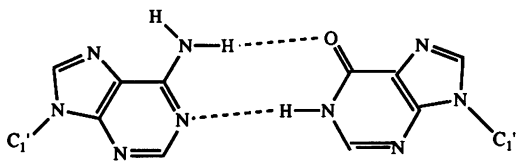
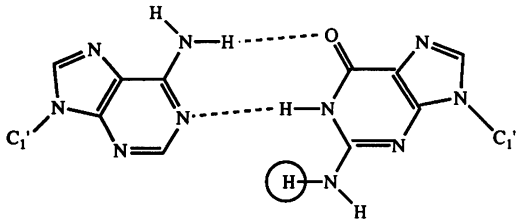


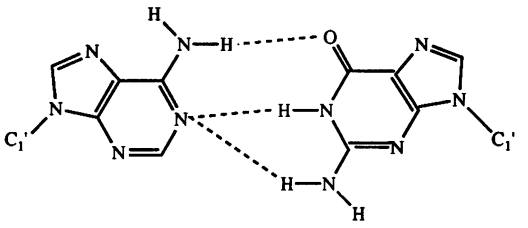
Figure 6. (a) A stereoview of the spine of hydration present in the minor groove of d(CGCAAATTIGCG). The water molecules are shown as stippled spheres. (b) A schematic representation of the spine of hydration shown in (a) showing hydrogen bonds to the base edges and other solvents in the spine. X denotes a water molecule missing from the pattern while dashed lines represent distances longer than would be expected for hydrogen bond formation. We have taken the maximum distance for hydrogen bond formation to be 3.5Å. A * indicates that an atom belongs to a symmetry-related duplex.



(a)



(b)



(c)

Figure 7. (a) A representation of an A(*anti*)·I(*anti*) base-pair. (b) A representation of the A(*anti*)·G(*anti*) base-pair. The N2-amino group hydrogen atom that results in destabilisation of the base-pair is circled. (c) The reverse three-centre inter-base hydrogen bond that could occur in the A(*anti*)·G(*anti*) base-pair.

be associated with Watson–Crick base-pairs. However, this is not the case and the spine of hydration covers the whole of the core $(AAATTI)_2$ region of the duplex. N3(A17) and C8(I9) are bridged by the solvent molecule W39 (Figure 6(b)). This solvent is reasonably well-ordered with a B-value of 32Å^2 compared to an average value of 44Å^2 for all the solvent atoms in the spine of hydration. At the other end of the core region there is a solvent molecule (W95) situated between N3(A5) and C8(I21). This atom does not form any hydrogen bonds to the minor groove atoms. However, it is not very well-ordered and the associated electron density is rather diffuse suggesting the presence of more than one solvent molecule in this region although we can adequately model the density using only one water molecule with a rather high temperature factor (B=68Å^2). The presence of this atom is further evidence that the hydration spine includes the whole of the core region of the duplex and is not perturbed by the presence of *syn* purine bases.

The relative stability of the A·I and A·G base-pairs

From an examination of Figure 5 it can be seen that while the A(*anti*)·I(*syn*) base-pair is more stable (T_m = 34°C) than its A(*anti*)·G(*anti*) counterpart (T_m = 19°C), the AH$^+$(*anti*)·I(*syn*) base-pair is less stable (T_m = 34°C) than the AH$^+$(*anti*)·G(*syn*) analogue (T_m = 39°C).

Figures 7(a) and 7(b) show a comparison of the two *anti·anti* base-pairs and suggests a reason why the A(*anti*)·I(*anti*) base-pair is more stable than the G(*anti*)·A(*anti*) pair. If the inter-base hydrogen bonds are exactly as shown, one of the N2-amino group hydrogens in the A(*anti*)·G(*anti*) base-pair cannot form any sort of hydrogen bond due to steric hinderance. This presence of an unfulfilled hydrogen bond donor is bound to have a large destabilizing influence on the base-pair. It is possible that the hydrogen atom in question could be involved in a reverse three-centre hydrogen bond (Figure 7(c)) of the type put forward in the G·O8A base-pairs in d(CGCGAATT(O8A)GCG) (O8A = 8-oxoadenosine) and the G·A base-pairs in d(CGCGAATTAG-CG) (30) although even this situation would not completely alleviate the destabilization of the base-pair. Inosine lacks an N2-amino group and thus the A(*anti*)·I(*anti*) base-pair is likely to be more stable than the A·G analogue. A further possible consequence of the presence of an N2-amino group on the minor groove side of the A(*anti*)·G(*anti*) base-pair is that it may interfere with the formation of the spine of hydration confining it to the $(AATT)_2$ region of the d(CGCAAATTGGCG)$_2$ duplex at pH 7. This situation would not arise in a d(CGCAAATTIGCG)$_2$ duplex containing A(*anti*)·I(*anti*) base-pairs and the spine of hydration could then cover the whole of the $(AAATTI)_2$ region of the duplex. This could also contribute to the stabilization of the A(*anti*)·I(*anti*) base-pair relative to the A(*anti*)·G(*anti*) base-pair.

A possible reason for the greater stability of the AH$^+$(*anti*)·G(*syn*) base-pair compared to its A·I counterpart also involves the N2 amino group on the guanosine. In the *anti·syn* conformation it is no longer on the minor groove side of the base-pair and, in fact, is linked, via a network of water-mediated hydrogen bonds (see Figure 6 of reference 8), to a neighbouring phosphate oxygen atom. This will undoubtedly result in a stabilisation of the AH$^+$(*anti*)·G(*syn*) with respect to the AH$^+$(*anti*)·I(*syn*) pair which, of course, lacks this N2 amino group. This may explain why the AH$^+$(*anti*)·G(*syn*) base-pair is more stable than the AH$^+$(*anti*)·I(*syn*) mispair.

CONCLUSIONS

The crystal structure analysis of the synthetic dodecanucleotide d(CGCAAATTIGCG) at pH 6.5 has shown that the A·I mispairs adopt an A(*anti*)·I(*syn*) geometry. pH-dependent ultraviolet melting studies have suggested that this base-pair is in fact a protonated AH$^+$(*anti*)·I(*syn*) species rather than A(imino, *anti*)·I(syn) and that, like the A·G mispair in d(CGCAAATT-GGCG) (7,8), the A·I base-pair is conformationally variable with the protonated base-pair dominating below pH 6.5. Between pH 6.5 and 8.0 it is likely that the dominant configuration of the base-pair is A(*anti*)·I(*anti*). The AH$^+$(*anti*)·I(*syn*) mispair is stabilized by two hydrogen bonds one of which is a result of the protonation of the N1 of adenosine.

Our analysis of the A·I base-pairs in d(CGCAAATTIGCG) provides further evidence that inosine behaves very like guanosine when it is incorporated into oligonucleotides. It is also further confirmation of the conformationally variable nature of

purine·purine mispairs a factor which may explain why these base-pairs are inefficiently repaired by DNA proofreading mechanisms.

ACKNOWLEDGEMENTS

This work was supported by the United Kingdom Science and Engineering Council through the Molecular Recognition Initiative, the Wellcome Trust and Shell U.K. Ltd.

REFERENCES

1. Fersht, A.R., Knill-Jones, J.W. and Tsui, W.C. (1982). *J. Mol. Biol.* **212**, 36–51.
2. Kramer, B., Kramer, W. and Fritz, H-J. (1984). *Cell* 878–887.
3. Prive, G.G., Heinemann, U., Kan, L-S., Chandrasegaran, S. and Dickerson, R.E. (1987). *Science* **238**, 498–504.
4. Carbonnaux, C., van der Marel, G.A., van Boom, J.H., Guschlbauer, W. and Fazackerley, G.V. (1991). *Biochemistry* **30**, 5449–5458.
5. Brown, T., Hunter, W.N., Kneale, G.G. and Kennard, O. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 2402–2406.
6. Gao, X. and Patel, D.J. (1988). *J. Amer. Chem. Soc.* **110**, 5178–5182.
7. Brown, T., Leonard, G.A., Booth, E.D. and Chambers, J. (1989). *J. Mol. Biol.* **207**, 455–457.
8. Leonard, G.A., Booth, E.D. and Brown, T. (1990) *Nucleic Acids Res.* **18**, 5617–5623.
9. Lane, A.N., Jenkins, T.C., Brown, D.J.S. and Brown, T. (1991). *Biochem. J.* **279**, 269–281.
10. Takashaki, Y., Kato, K., Hiyashizaki, Y., Wakabayashi, T., Ohstuka, E., Matsuki, S., Ikehara, I. and Matsubara, K. (1985). *Proc. Natl. Acad. Sci. USA.* **82**, 1931–1935.
11. Ohtsuka, E., Matsuki, S., Ikehara, M., Takashaki, Y. and Matsubara, K., (1985). *J. Biol. Chem.* **260**, 2605–2608.
12. Davis, B.B., Anderson, P. and Sparking, P.F. (1973). *J. Mol. Biol.* **76**, 223–232.
13. Kumar, V.D., Harrison, R.W., Andrews, C.A. and Weber, I.T. (1992). *Biochemistry* **31**, 1541–1550
14. Cruse, W.B.T., Aymani, J., Kennard, O., Brown, T., Jack, A.G.C. and Leonard, G.A. (1989). *Nucleic Acids Res.* **17**, 55–72.
15. Corfield, P.W.R., Hunter, W.N, Brown, T., Robinson, P. and Kennard, O. (1987). *Nucleic Acids Res.* **15**, 7935–7949.
16. Wing, R., Drew, H.R., Takano, T., Broka, C., Takana, S., Itakura, K and Dickerson, R.E. (1980) *Nature*, **287**, 755–758.
17. Dickerson, R.E. and Drew, H.R. (1981). *J. Mol. Biol.* **149**, 761–786.
18. Westhof, E., Dumas, P. and Moras, D. (1985). *J. Mol. Biol.* **184**, 119–145.
19. Sheldrick, G.M. (1976). SHELX76 System of Computer Programs (Univ. of Cambridge, Cambridge, U.K.)
20. Hendrickdon, W.A. and Konnert, J.H. (1981) in Srinivasan, R. (ed.), Biomolecular Structure, Conformation, Function and Evolution. Pergammon Press, Oxford, Vol 1, pp43–57.
21. Jones, T.A. (1978). *J. Appl. Cryst.* **11**, 268–272.
22. Larsen, T.A., Kopka, M.L. and Dickerson, R.E. (1991). *Biochemistry* **30**, 1403–1412.
23. Luzatti, P.V. (1952). *Acta.Cryst.* **5**, 802–810.
24. Frederick, C.A., Quigley, G.J., van der Marel, G.A., van Boom, J.H., Wang, A.H-J. and Rich, A. (1988). *J.Biol.Chem*, **263**, 17872–17879.
25. Dickerson, R.E., Bansal, M., Calladine, C.R., Diekmann, S., Hunter, W.N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H.C.M., Olson, W.K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D.M., Tang, C-S., Wang, A,H-J. and Khurkin, V.B. (1989). *Nucleic Acids Res.* **17**, 1797–1803.
26. Coll, M., Frederick, C.A., Wang, A.H-J. and Rich, A. (1987). *Proc. Natl. Acad. Sci. USA.* **84**, 8385–8389.
27. DiGabriele, A.D., Sanderson, M.R. and Steitz, T.A. (1989). *Proc Natl. Acad. Sci. USA.* **86**, 1816–1820.
28. Nelson, H.C.M., Finch, T.J., Luisi, B.F. and Klug, A. (1987). *Nature.* **330**, 221–226.
29. Drew, H.R and Dickerson, R.E. (1981). *J. Mol. Biol.* **151**, 535–556.
30. Leonard, G.A., Guy, A., Brown, T., Téoule, R. and Hunter, W.N. (1992). *Biochemistry* in press.