Volume 15 Number 19 1987 Nucleic Acids Research

Inosine.adenine base pairs in a B-DNA duplex

Peter W.R.Corfield[1], William N.Hunter, Tom Brown[2], Paul Robinson[2] and Olga Kennard*

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK, [1]Department of Chemistry, The Kings College, Briarcliff Manor, NY 10510, USA and [2]Department of Chemistry, University of Edinburgh, Edinburgh, UK

Received August 3, 1987; Revised and Accepted September 7, 1987

ABSTRACT

The structure of the synthetic deoxydodecamer d(C-G-C-I-A-A-T-T-A-G-C-G) has been determined by single crystal X-ray diffraction techniques at 2.5Å resolution. The refinement converged with a crystallographic residual, R=0.19 and the location of 64 solvent molecules. The sequence crystallises as a B-DNA helix with 10 Watson-Crick base-pairs (4 A.T and 6 G.C) and 2 inosine.adenine (I.A) pairs. The present work shows that in the purine.purine base-pairs the adenine adopts syn orientation with respect to the furanose moiety while the inosine is in the trans (anti) orientation. Two hydrogen bonds link the I.A base-pair, one between N-1(I) and N-7(A), the other between O-6(I) and N-6(A). This bulky purine.purine base-pair is incorporated in the double helix at two positions with little distortion of either local or global conformation. The pairing observed in this study is presented as a model for I.A base-pairs in RNA codon-anticodon interactions and may help explain the thermodynamic stability of inosine containing base-pairs. Conformational parameters and base stacking interactions are presented and where appropriate compared with those of the native compound, d(C-G-C-G-A-A-T-T-C-G-C-G) and with other studies of oligonucleotides containing purine.purine base-pairs.

INTRODUCTION

Deoxyinosine (I) is effective as a universal base in synthetic hybridisation probes when cloning genes for proteins containing amino acids with degenerate codons (1-3). The use of inosine containing oligonucleotides for this purpose is now widespread. Inosine, despite the absence of a 2-amino group behaves very like guanine. It is able to pair with cytosine, thymine and adenine without destabilising the double helix even when several inosine containing base-pairs occur in a short stretch of DNA (1,2). This implies that inosine forms stable mismatches, a feature that is not surprising when one considers the occurrence of inosine at the 5' hydroxyl end, the wobble position, of some tRNA anticodons where it is known to form hydrogen bonds with A, C or U at the mRNA codon (4). Inosine is also able to occupy the middle position of the anticodon and pair with A(5).

Crick suggested in his wobble hypothesis (4) that in codon-anticodon interactions the first two positions are read according to Watson-Crick complementarity but that a certain latitude is allowed in the interactions of the third position. He proposed an I.A pairing with the bases in anti orientation with respect to the furanose moiety. A purine.purine base pair of this geometry would have a C-1'...C-1' distance of around 12.8Å and is then likely to distort the ribose-phosphate backbone. Any such distortions might be expected to disrupt codon-anticodon interactions as regulated by steric considerations in the ribosome. The influence of a purine-purine base pair and the distortions it can cause when both bases are in the anti orientation is shown in the structure of tRNA phe (6). An A.N-dimethylG base-pair is formed at the junction of the anticodon and D stems. This bulky pairing is thought to be responsible for a 26° kink between these two helical regions of the molecule.

In order to investigate the precise nature of the I.A pairing in a DNA helix we have carried out an X-ray structural analysis of the synthetic deoxydodecamer, d(C-G-C-I-A-A-T-T-A-G-C-G) to 2.5Å resolution. This structure analysis establishes that an I(anti).A(syn) mispair can be formed and is accommodated in the B-DNA duplex without any major distortion of the local or global conformation. The analysis allows for a detailed comparison with the crystal structures of d(C-G-C-G-A-A-T-T-A-G-C-G) (7) which contains two G(anti).A(syn) mispairs and the native structure, d(C-G-C-G-A-A-T-T-C-G-C-G) (8,9) and also with other studies of oligonucleotides containing purine.purine base-pairs.

MATERIALS AND METHODS

Synthesis, Crystallisation and Data Collection

Model studies suggest that deoxyinosine derivatives are unstable to the standard conditions of triester oligonucleotide synthesis. Reactions at the 6-position of the hypoxanthine base can lead to the formation of deoxyadenosine (10). The extent of this side reaction was investigated in the following manner. The sequence d(C-T-T-C-I-T) and the control sequence d(C-T-T-C-A-T) were synthesised and digested with snake venom phosphodiesterase then analysed by HPLC. The control sequence yielded dCMP, dTMP, dAMP and dC as expected. The I containing sequence gave dCMP, dTMP, dIMP and dC in the correct proportions and showed no traces of dAMP or any other modified nucleotides. This experiment indicated that inosine containing oligonucleotides can be prepared using standard solid-phase triester methods (11). These techniques were used to synthesise d(C-G-C-I-A-A-T-T-A-G-C-G). The sequence was purified by

ion-exchange chromatography and by reverse phase HPLC. Crystals were grown from an aqueous solution containing approximately 0.5mM DNA duplex, 7.5mM sodium cacodylate buffer (pH 7.4), 20mM $MgCl_2$, 1.0 mM spermine and 15% (vol/vol) 2-methyl-2,4-pentanediol. The crystals are orthorhombic $P2_12_12_1$ with unit cell dimensions of a=25.85, b=41.99, c=65.10Å as determined from diffractometer measurements. These paramoters indicated isomorphism with the dodecamer d(C-G-C-G-A-A-T-T-C-G-C-G) studied by Dickerson and co-workers (8,9). An assumed crystal density of 1.5 gcm^{-3} indicates the presence of two dodecamer strands (one dodecamer helix) in the asymmetric unit and DNA content of about 50% by weight in the unit cell.

Intensities were measured at 4-5°C with a Syntex $P2_1$ diffractometer equipped with a graphite monochromator (λ = 1.5418Å), a long arm and helium path. A crystal (dimensions 0.50 x 0.20 x 0.12mm) was mounted and sealed in a glass capillary surrounded by mother liquor. Two octants of data were measured independently and averaged (R symm = 0.07) yielding 2207 unique reflections to 2.5Å resolution. The data were corrected for Lorentz polarization and decomposition effects.

Structure Refinement

The starting model for the analysis was the native dodecamer d(C-G-C-G-A-A-T-T-C-G-C-G) (8,9). The nucleotides are numbered C1 to G12 in the 5' to 3' direction on strand 1, C13 to G24 on strand 2. The model was refined against the experimental data as a rigid body using a modified version of SHELX (12). Initially only data within the 10-7Å range was included in the calculations. The resolution was extended in 4 steps of 1Å. The rigid body refinement converged with R=0.40 for 1541 reflections with $I>\sigma(I)$ in the 10-3Å range. Throughout the calculations atoms of nucleotides G4.C21 and C9.G16 were given a low occupancy so that they would not contribute to structure factor calculations. Refinement was continued with the restrained least squares method of Konnert and Hendrickson (13) using NUCLSQ, a program modified for nucleic acid components (14). The resolution was extended to 2.5Å when electron density (2Fo-Fc) and difference (Fo-Fc) maps were calculated and displayed on an Evans and Sutherland PS300 computer graphics system using the program FRODO (15). The required changes were made at the I.A sites and the model coordinates manipulated to fit the electron density. The syn orientation of the adenine was evident at this stage. The atoms of I4.A21 and A9.I16 were given full occupancy and the refinement continued with the location of solvent molecules. Great care was exercised in the identification of solvent molecules from electron density

and difference maps so as to reduce the chances of including artefacts into the refinement process. A total of 64 solvent positions were identified. The final crystallographic residual was 0.19 for 2105 reflections with $I>\sigma(I)$ in the range 8-2.5Å.

RESULTS AND DISCUSSION

The structure of d(C-G-C-I-A-A-T-T-A-G-C-G)

The asymmetric unit consists of two chemically equivalent strands coiled about each other to form an antiparallel double helix with 10 standard Watson-Crick base-pairs and 2 I.A pairs at I4.A21 and A9.I16. The duplex, illustrated in Figure 1, is of the B-DNA type. Symmetry related duplexes interact with hydrogen bonds and van der Waals contacts involving atoms in the minor grooves at the C-G-C ends (8,9).

The I.A base-pair

The geometry of the two I.A(syn) base-pairs is identical within the accuracy of the analysis. Figure 2 presents the I4.A21 superimposed on a fragment Fo-Fc map. The syn orientation of adenine is unambiguous. The purine.purine pair is linked with two hydrogen bonds as depicted for I4.A(syn)21 in Figure 3. The N-1 and O-6 atoms of I form hydrogen bonds to N-7 and N-6 of the syn adenine. The N...N distances are 2.7 and 3.0Å, the O...N separations are 2.9 and 2.7Å in the two I.A(syn) pairs. The tilt of the bases in a pair with respect to each other is represented by propeller twist. At I4.A21 and A9.I16 these values are 15° and 19°. The C-1'...C-1' distances are 11.2 and 10.5Å in these base-pairs. These values agree closely with those observed for G.A(syn) mispairs in a similar structure to that reported here (7) and to the pairing proposed by modeling studies (16). The G.A base pairs were found to be held together with two hydrogen bonds (N-1(G)...N-7(A) and O-6(G)...N-6(A)) with the adenine in a syn orientation. This anti.syn pairing is different from the G.A base-pair identified by crystallographic and nmr techniques in two different duplexes (17-19) and nmr studies of the I.A pair recently reported (20) where both bases occur in the anti orientation (see Figure 4). It should be noted that different types of purine.purine base-pairs are being observed in different sequence environments, a point we shall address further on.

It had previously been suggested (4) that the G(anti).A(anti) base-pair might be destablised by steric hindrance between the 2-amino group of guanine and the 2-hydrogen of adenine. Such an interaction would not occur in the G(anti).A(syn) base-pair so this might be preferred. Since I does not possess the 2-amino group this destabilising factor would not be present in the

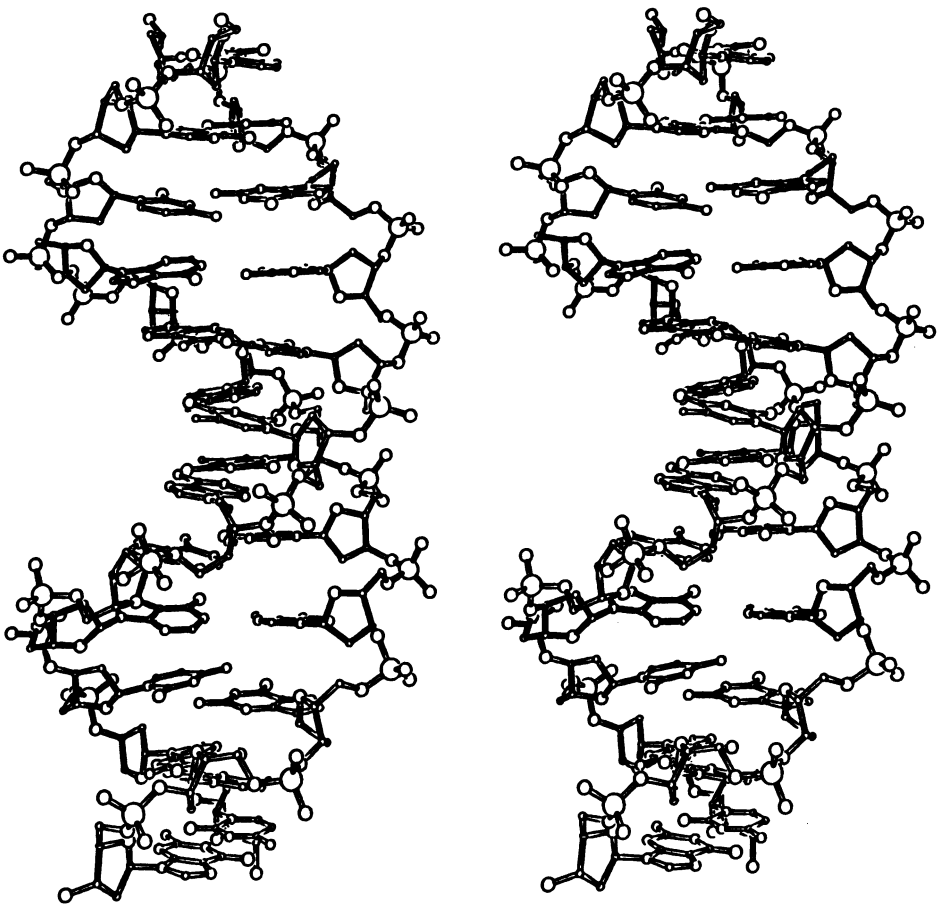

Figure 1
Stereoview of the B-DNA dodecamer. The bonds of nucleotides I4,A9, I16 and A21 are filled in to highlight the I.A(syn) base-pairs. Atoms are shown as spheres decreasing in size in the order P>O>N>C. Base-pair C1.G24 is at the top of the diagram.

I(anti).A(anti) pairing. However, the I(anti).A(syn) base-pair observed in the present structure suggests that other considerations might be important to this pair's stability.

The geometrical features of G.A(syn) and I.A(syn) base-pairs are similar to those of the G.C and A.T Watson-Crick pairs observed crystallographically (7,9,21-23). As mentioned above the C-1'...C-1' distances and propeller twist values do not differ significantly from the standard base-pairs. There is,

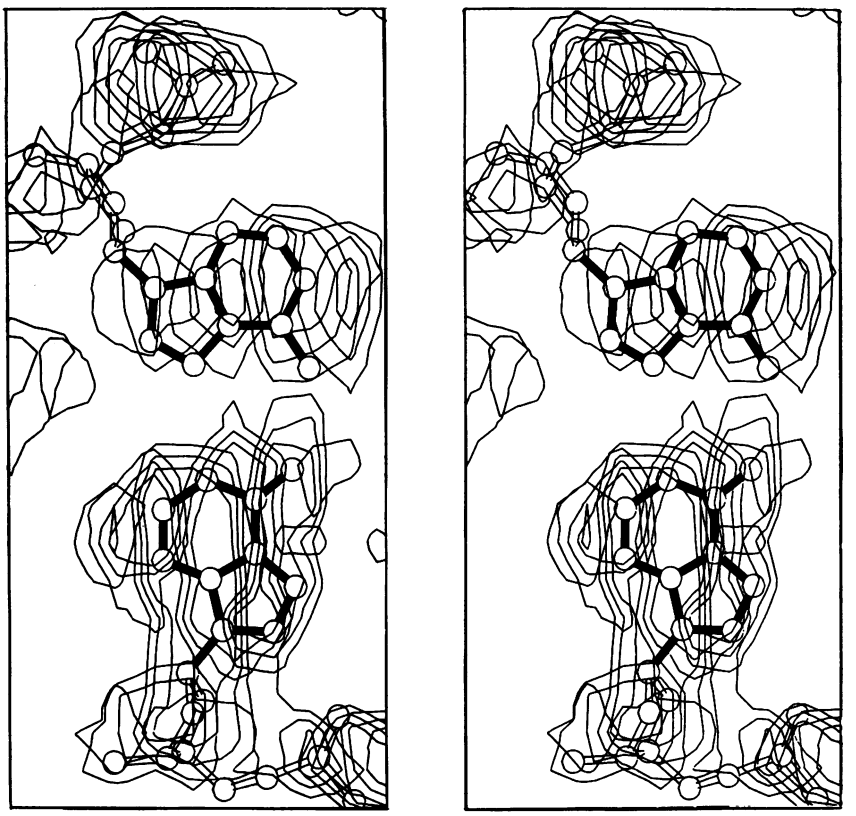

Figure 2
Stereoview of the I4.A(syn)21 base-pair superimposed on a composite section of a "fragment" Fo-Fc map. All atoms shown in this diagram have been omitted from the structure factor calculations (Fc). The Fo-Fc map can then be regarded as an unbiased representation of the electron density. The syn adenine is at the top of the figure.

however, one parameter which does change. This is the angle between the glycosidic bond C-1'-N-1 or N-9 and the C-1'...C-1' vector of a base, designated λ_1 on strand 1. The complementary angle subtended at the base on strand 2 is designated λ_2. In Watson-Crick base-pairs the angle λ falls within a narrow range (52-62°) and the base-pairs have an element of pseudosymmetry (22,23). In the I.A(syn) and G.A(syn) base-pairs the value of λ is normal for the I and G bases, see Table 1 and reference 6. However, the syn adenines have λ values much smaller (41° or less) than those formed by Watson-Crick pairs. It is worth

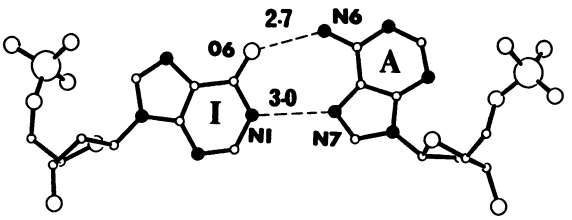

Figure 3
The I4.A(syn)21 base-pair. The hydrogen bonds are marked with dashed lines between the labelled atoms. Atoms are depicted as spheres of decreasing radii such that P>O>N>C; N atoms are dark. Distances are marked in Å.

noting that the degree of asymmetry of base pairs represented by λ may be linked to the efficiency of mismatch repair (24).

Conformation of the Double helix

The insertion of two I.A(syn) base-pairs into the dodecamer produces only small, highly localised changes when compared to the parent compound. The global conformation is not affected: the global twist changes from 37° to 36° and the average rise per residue is 3.3Å in both structures.

The sugar-phosphate backbone

The separations between adjacent phosphorus atoms along each strand range from 6.2 to 7.1Å (average = 6.7Å) in both d(C-G-C-I-A-A-T-T-A-G-C-G) and the native dodecamer. Groove widths may be estimated from the distances of P atoms across the two strands. The width of the minor groove in the present structure ranges from 9.4Å [P(T8)...P(A21)] to 13.8Å [P(A5)...P(G24)]. The width of the major groove ranges from 16.6Å [P(I4)...P(A18)] to 18.6Å [P(T7)...P(C15)]. These values correspond closely to those of the native compound (8,9). The sugar-base glycosidic and individual backbone torsion angles are presented in Table 2 together with their average values and for comparison the average values of the native compound. In the current study a wider range of torsion angles is observed than in the native structure but the average values correspond closely.

The glycosidic bond angle, χ , can be placed in two categories depending on whether a purine or pyrimidine base is involved. For purines in the usual anti conformation the average χ is -104°, for pyrimidines the average is -120°. This feature of χ may represent steric differences between five and six membered rings adjoining the sugar (25). The range of χ observed for bases in the anti conformation is -84 to -144°. This corresponds to -synclinal through to -anticlinal orientation as defined by IUPAC-IUB nomenclature (26). Bases A9 and A21 are syn and have χ values of 71 and 72° respectively.

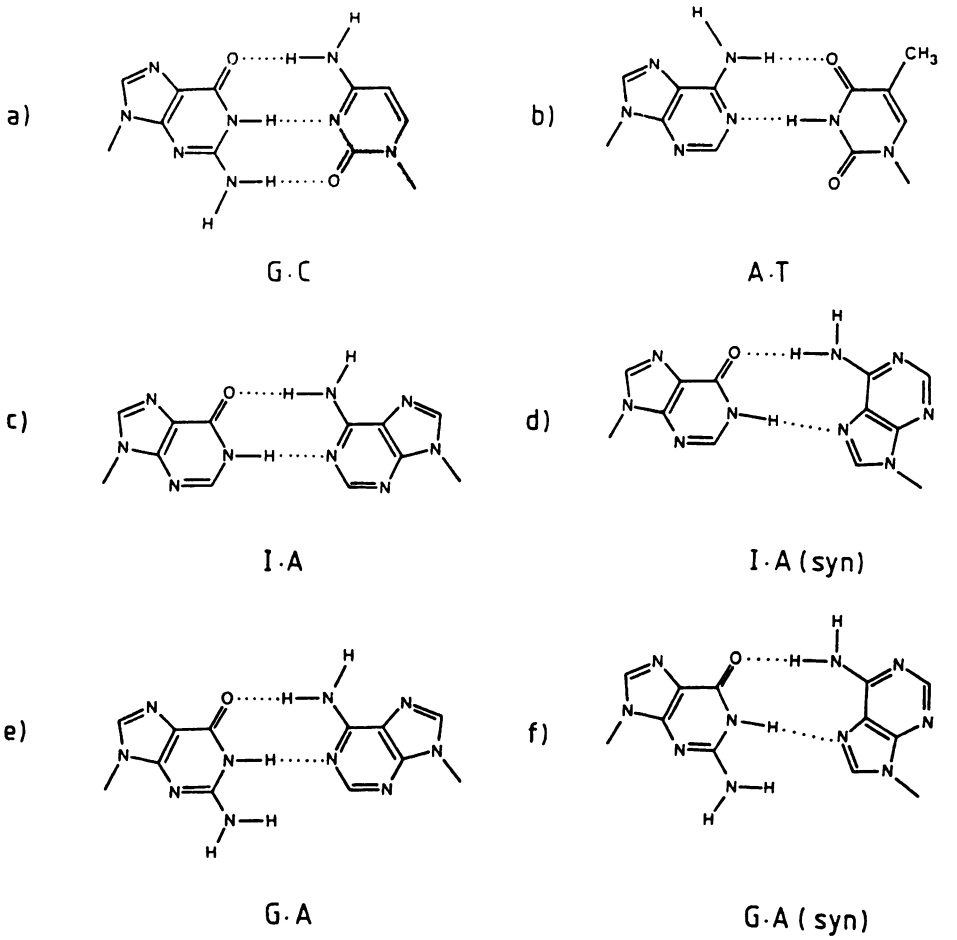

Figure 4
Base-pairing schemes. Watson-Crick G.C and A.T are shown in (a) and (b). Two types of I.A base-pairs, anti.anti (c) and anti.syn (d). Two types of G.A pairs, anti.anti (e) and anti.syn (f). In these schemes only major tautomer forms are used.

Torsion angle δ ranges from 75 to 160° (+synclinal to +antiperiplanar). This torsion angle is related to the conformation of the furanose ring (27). The observed range of δ corresponds to sugar conformations C3'-endo to C3'-exo. A similar variety of sugar conformations was observed in the native structure (9). It may be significant that δ for I4 and I16 is 129° and 102° compared to values of 156° and 136° for G4 and G16 in the native dodecamer.

The conformations of the remaining sugar-phosphate torsion angles may be

Table 1

Geometrical properties of base pair steps and base pairs in d(C-G-C-I-A-A-T-T-A-G-C-G)

Base pair	Step	Roll (°)	Twist (°)	Rise (Å)	Propeller Twist (°)	λ_1 (°)	λ_2 (°)	C1'...C1' (Å)
C1.G24					13	49	58	10.8
	1	2	40	3.2				
G2.C23					9	51	56	10.6
	2	-3	42	2.9				
C3.G22					8	52	59	10.7
	3	5	27	4.0				
I4.A21					15	49	37	11.2
	4	3	36	3.4				
A5.T20					13	57	57	10.4
	5	0	38	3.2				
A6.T19					23	54	58	10.3
	6	-3	32	3.2				
T7.A18					14	61	57	10.2
	7	-1	37	3.1				
T8.A17					18	62	55	10.1
	8	-2	41	3.3				
A9.I16					19	40	59	10.5
	9	2	30	4.0				
G10.C15					9	53	54	10.7
	10	-5	40	2.8				
C11.G14					18	56	52	10.6
	11	-4	34	3.5				
G12.C13					7	55	51	10.8

summarised as follows. Torsion angle α ranges from -synperiplanar to -antiperiplanar, β varies from +anticlinal to -anticlinal while γ is restricted to +synperiplanar/+synclinal when γ C1 and γ C13 are ignored. The torsion angle ε ranges from -anticlinal to +synclinal, ζ from -synclinal to +synclinal.

Base stacking interactions

In d(C-G-C-I-A-A-T-T-A-G-C-G) there are eleven base-pair steps of which seven involve only Watson-Crick base-pairs. These are steps 1 and 11 of type CpG(=CpG), 2 and 10 of type GpC(=GpC), 5 and 7 of type ApA(=TpT) and step 6 of type ApT(=ApT). These base-pairs steps all occur in the parent dodecamer and are very similar in the two structures. There is considerable variation in base-pair overlap from one step to another. Purine-pyrimidine steps (2, 6 and 10) display a greater degree of overlap than the pyrimidine-purine steps (1 and 11). Each ApA(=TpT) step shows good base overlap with the six membered rings of the purine bases interacting strongly. Alterations in base stacking are localised at the I.A(syn) positions. The steps affected are 3 and 9 of type CpI(=ApG), 4 and 8 of type IpA(=TpA). Similar stacking interactions are observed around the two I.A(syn) base-pairs hence only those of I4.A(syn)21 are illustrated (Figure 5) and compared with equivalent steps in the native structure. Step 3, CpG(=CpG) in the parent dodecamer shows moderate overlap of C21 on G22 but only a weak interaction between G4 and C3. At the corresponding

Table 2

Torsion angles (°) and distances between adjacent phosphorous atoms (Å)

Residue	χ	α	β	γ	δ	ξ	ζ	P_i-P_{i+1}
C1	-116	-	-	276	136	-147	-143	-
G2	-96	-17	148	13	151	-167	-127	6.6
C3	-140	-73	150	73	97	-149	-142	6.6
I4	-116	-21	117	63	129	-198	-74	6.9
A5	-92	-69	232	32	153	-200	-89	6.3
A6	-95	-66	201	54	133	-200	-74	6.8
T7	-118	-92	200	56	89	-287	-331	6.5
T8	-124	-165	197	80	82	-289	-317	6.7
A9	71	-120	186	52	116	-175	-87	6.5
G10	-89	-63	180	50	151	-83	-213	6.9
C11	-106	-63	175	16	140	-186	-61	7.0
G12	-116	-112	185	70	92	-	-	-
C13	-93	-	-	172	160	-186	-106	-
G14	-110	-78	181	60	132	-175	-129	6.7
C15	-144	-21	130	55	85	-155	-122	6.2
I16	-127	-32	129	66	102	-194	-69	7.1
A17	-98	-84	203	59	122	-241	-32	6.8
A18	-84	-95	229	37	153	-163	-118	6.3
T19	-134	-58	142	77	75	-183	-66	7.0
T20	-103	-53	202	29	133	-190	-78	6.5
A21	72	-89	199	57	107	-175	-98	6.4
G22	-90	-29	159	37	151	-107	-185	6.9
C23	-121	-64	145	29	115	-170	-102	6.5
G24	-137	-31	164	17	79	-	-	-
Average	-111*	-68	175	49±	120	-183	-126	6.7
Average values for d(C-G-C-G-A-A-T-T-C-G-C-G)								
	-117	-63	171	54	123	-169	-108	6.7

Main chain torsion angles are defined by

$$P\overset{\alpha}{-}O\text{-}5'\overset{\beta}{-}C\text{-}5'\overset{\gamma}{-}C\text{-}4'\overset{\delta}{-}C\text{-}3'\overset{\xi}{-}O\text{-}3'\overset{\zeta}{-}P$$

Glycosyl torsion angle, χ is defined by

O-4'--C-1'--N-1--C-2 for pyrimidines

O-4'--C-1'--N-9--C-4 for purines

* omits χ for A9 and A21 with syn orientation of base

± omits γ for C1 and C13

step in d(C-G-C-I-A-A-T-T-A-G-C-G), CpI(=ApG) the small degree of overlap between bases 3 and 4 (C3 on I4) is conserved. On the opposing strand a large purine has replaced a pyrimidine, A for C. The degree of overlap is maintained. At step 4 of the native structure, GpA(=TpC), there is a good overlap of G4 on A5 and also of C21 on T20. On the 3' side of I4.A(syn)21 the good purine purine overlap (I4 on A5) is conserved in terms of the six membered ring of base 4 lying on the five and six membered rings of A5. The loss of an amino group in the minor groove when replacing guanine with inosine, however, reduces the stacking interaction between bases 4 and 5.

The syn adenine at position 21 displays enhanced overlap of the base with T20 when compared to the corresponding overlap in the native structure. A similar improvement in stacking was observed with the G.A(syn) base-pairs (7,28).

The parameters of Dickerson, Calladine and Drew (9,29,30) allow further analysis of base pair interactions (see Table 1). These parameters have been calculated using the programs HELIX, BROLL and CYLIN kindly supplied by R.E.Dickerson. The twist (rotation per base-pair step) values agree closely with those observed in the native structure. The largest difference in twist at corresponding steps is 3^{o} at steps 2 and 11. The angle by which adjacent base-pairs open up to the minor groove is the roll. This parameter is positive if the angle between base-pair planes opens to the minor groove, negative if towards the major groove. As previously observed in the native structure, at Watson-Crick base-pair steps, purine-pyrimidine steps open to the major groove and pyrimidine-purine steps to the minor groove. The steps involving I.A(syn) pairs are 3, 4, 8 and 9. On the 5' side of the inosine the roll values are the same as found in the native structure; 5^{o} and 2^{o} at steps 3 and 9 respectively. Step 4, the 3' side of I4 has a positive roll angle but this is reduced from 6^{o} in the native structure to 3^{o}. At step 8, 3' side I16, a reduction of 5^{o} in roll is observed from 3^{o} in the native structure to -2^{o}. The rise per base-pair between the two structures is very similar at corresponding steps. The largest difference is at step 11 where an increase of 0.7Å is observed.

Thermal parameters

Isotropic thermal parameters for individual atoms were refined against the diffraction data. These parameters ($B=8\pi^{2}U^{-2}\text{Å}^{2}$) reflect molecular motion (31) but may also contain contributions from errors in the data, the model and static disorder. The final thermal parameters in Å^{2} ranged from 13 to 35 (average = 21) for the bases, 23 to 47 (average = 32) for the sugars and 33 to 55 (average = 42) for the phosphates. These values agree well with those found in the

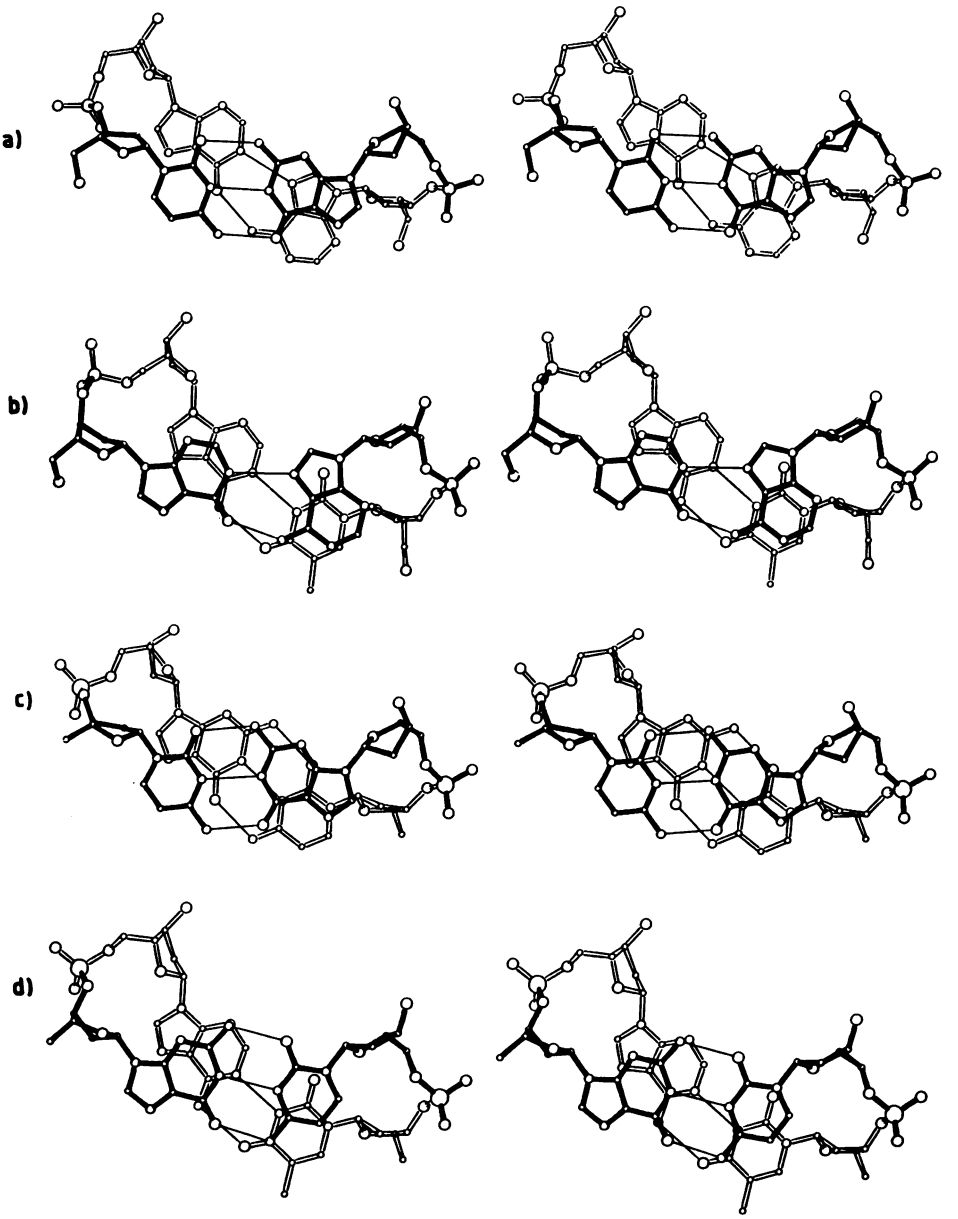

Figure 5
Base stacking interactions presented as stereoviews of steps involving the I4.A(syn)21 mispair, (a) and (b), and the corresponding steps in the parent compound. Hydrogen bonds are represented as thin lines. The bonds of the upper residue are filled in. Atoms are depicted as spheres of decreasing radii in the order P>O>N>C.

native structure where the ranges were 20 to 44 (average = 28) for bases, 32 to 51 (average = 42) for sugars and 38 to 61 (average = 40) for the phosphates. Base A21 has the highest B value of all the bases, 35 $Å^2$. However, the fit to electron density, see Figure 2, is good with no evidence of the anti conformer. The other bases involved in the I.A(syn) pairing have normal B values well within the ranges in Watson-Crick base-pairs. The solvent molecules have B values ranging from 17-72 $Å^2$ (average is 49 $Å^2$).

CONCLUSIONS

The detailed analysis of the structure of d(C-G-C-I-A-A-T-T-A-G-C-G) confirms our previous observations about anti-syn guanine.adenine base-pairs. The presence of a bulky pair such as the I.A or G.A does not appear to affect the global conformation of the double helix. Even at a local level we do not observe any large perturbations in either the torsion angles or other local parameters defining conformation. This is a direct consequence of the adenine adopting a syn orientation with respect to the sugar.

Although inosine is able to form stable base-pairs with the other nucleotides these base-pairs have been reported as being less stable thermodynamically than the equivalent guanine containing base-pairs (3). The notable exception is the I.A pair which, in a number of sequences, is more stable than the G.A base-pair (3). There are several factors which determine the stability of mispairs: the number and type of hydrogen bonds which are formed and the base stacking interactions. The I.A(syn) and G.A(syn) base-pairs have the same inter-base hydrogen bonds and display very similar base-stacking interactions in the dodecamer framework. A difference is that in the G.A(syn) base-pair the 2-amino group of the guanine protrudes into the minor groove. One of the hydrogens on this group is directed away from the base into bulk solvent and the hydrogen donating capacity could be satisfied by interactions with solvent. The other hydrogen is shielded from bulk solvent, tucked in towards the syn adenine and the hydrogen bonding capacity is unsatisfied. This is likely to destabilise the G.A(syn) mispair relative to the I.A(syn) pair. Similar arguments could apply to the I(anti).A(anti) and G(anti).A(anti) base-pairs. However, the crystal structure of d(C-C-A-A-G-A-T-T-G-G) which contains two G(anti).A(anti) mispairs, indicates an alternative way in which the hydrogen bonding capacity of the 2-amino of the mispaired guanine can be satisfied (19). The mispairs have a large propeller twist and this allows the formation of an inter-base-pair hydrogen bond between the 2-amino group of guanine and the carbonyl group of a thymine stacked on the 3' side of the

mismatched adenine. Thus in this case the sequence contributes to fulfilling the hydrogen bonding capacity of the 2-amino group of the mismatched guanine. The stabilising effect of this hydrogen bond may be counteracted by distortions to the double helix at the mismatch positions due to the anti.anti conformation of the bulky purine bases.

In summary, two types of G.A and I.A base-pairs have been characterised to date. These pairs differ in the orientation of the adenines, syn or anti. Theoretical studies indicate that the two types of base-pair are near equivalent (32). We conclude that the neighbouring sequence makes a large contribution to determining which conformation is present and agree with further theoretical work (33) that points to the importance of base stacking interactions in this respect. More structural and thermodynamic information of purine.purine base-pairs in different sequence environments is required before we can hope to fully understand the mechanisms which stabilise different conformations of purine.purine base-pairs.

ACKNOWLEDGEMENT

We are grateful to the M.R.C. Laboratory of Molecular Biology, Hills Road Cambridge for use of their computer graphics system, to Mrs H.Belchamber for preparation of this manuscript and to G.G.Prive and R.E.Dickerson for communicating results in advance of publication. This work was carried out under an M.R.C. program grant to O.K and grants from the S.E.R.C and Nuffield Foundation to T.B.

*To whom reprint requests should be addressed

REFERENCES

1 Takahashi, Y., Kato, K., Hayashizaki, Y., Wakabayashi, T., Ohtsuka, E., Matsuki, S., Ikehara, I. and Matsubara, K. (1985) Proc. Natl. Acad. Sci. USA, 82,1931-1935

2 Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. and Matsubara, K. (1985) J. Biol. Chem. 260, 2605-2608

3 Martin, F.H., Castro, M.M., Aboul-ela, F. and Tinoco, I. (1985) Nucleic Acids Res. 13, 8927-8938

4 Crick, F.H.C. (1966) J. Mol. Biol. 19, 548-555

5 Davis, B.D., Anderson, P. and Sparking, P.F. (1973) J. Mol. Biol. 76, 223-232

6 Ladner, J.E., Jack, A., Robertus, J.D., Brown, R.S., Rhodes, D., Clark, B.F.C. and Klug, A. (1975) Proc. Natl. Acad. Sci. USA 72, 4414-4418

7 Brown, T., Hunter, W.N., Kneale, G. and Kennard, O. (1986) Proc. Natl. Acad. Sci. 83, 2402-2406

8 Wing, R., Drew, H.R., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) Nature (London) 287, 755-758

9 Dickerson, R.E. and Drew, H.R. (1981) J. Mol. Biol. 149, 761-786
10 Adamiak, R. and Biata, E. (1985) Nucleic Acids Res. 13, 2989-3003
11 Gait, M.J., Matthes, H.W.D., Singh, M., Sproat, B.S and Titmus, R.C. (1982) Nucleic Acids Res. 10, 6243-6248
12 Sheldrick, G.M. (1976) SHELX 76 System of Computing Programs (Univ. of Cambridge, Cambridge, England)
13 Hendrickson, W.A. and Konnert, J.H. (1981) in Biomolecular Structure, Conformation, Function and Evolution, ed. Srinivasan, R. (Pergammon, Oxford), Vol. 1, 43-57
14 Westhof, E., Dumas, P. and Moras, D. (1985) J. Mol. Biol. 184, 119-145
15 Jones, T.A (1978) J. Appl. Cryst. 11, 268-272
16 Traub, W. and Sussman, J.L. (1982) Nucleic Acids Res. 19, 2701-2705
17 Kan, L.S., Chandrasegaran, S., Pulford, S.M. and Miller, P.S. (1983) Proc. Natl. Acad. Sci. USA 80, 4263-4265
18 Patel, D.J., Kozlowski, S.A., Ikuta, S. and Itakura, K. (1984) Fed. Proc. Fed. Am. Soc. Exp. Biol. 43, 2663-2670
19 Prive, G.G., Heinemann, U., Kan, L-S., Chandresegaran, S. and Dickerson, R.E. (1987) Fifth Conversation in Biomolecular Stereodynamics, 57-58 (ED Sarma, R.H. Adenine press)
20 Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y. and Ohtsuka, E. (1987) J. Biol. Chem. 262, 6965-6968
21 Hunter, W.N., Brown, T., Kneale, G., Anand, N.N., Rabinovich, D. and Kennard, O. (1987) J. Biol. Chem. (1987) 262, 9962-9970
22 Seeman, N.C., Rosenberg, J.M., Suddath, F.L., Kim, J.J.P. and Rich, A. (1976) J. Mol. Biol. 104, 109-144
23 Rosenberg, J.M., Seeman, N.C., Day, R.O. and Rich, A. (1976) J. Mol. Biol. 104, 145-167
24 Hunter, W.N., Brown, T., Anand, N.N. and Kennard, O. (1986) Nature 320, 552-555
25 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
26 IUPAC-IUB Joint Commission on Biochemical Nomenclature. (1983) Eur. J. Biochem. 131, 9-15
27 Levitt, M. and Warshel, A. (1978) J. Amer. Chem. Soc. 100, 2607-2613
28 Hunter, W.N., Brown, T. and Kennard, O. (1986) J. Biomolec. Stereodynamics 4, 173-191
29 Dickerson, R.E. (1983) J. Mol. Biol. 174, 663-695
30 Calladine, C.R. and Drew,H.R. (1984) J. Mol. Biol. 179, 773-782
31 Ringe, D. and Petsko, G.A. (1985) Prog. Biophys. Molec. Biol. 45, 197-235
32 Churpina, V.P. and Poltev, V.I. (1983) Nucleic Acids Res. 11, 5205-5222
33 Keepers, J.W., Schmidt, P., James, T.L. and Kollman, P.A. (1984) Biopolymers. 23, 2901-2929