

Base pairing and mutagenesis: Observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR

(mismatch/base analogue/replication errors/fidelity replication)

LAWRENCE C. SOWERS*, G. VICTOR FAZAKERLEY†, RAMON ERITJA‡, BRUCE E. KAPLAN‡,
AND MYRON F. GOODMAN*

*Molecular Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089-1481; †Service de Biochimie, Bat. 142, Département de Biologie, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette Cedex, France; and ‡Department of Molecular Genetics, Beckman Research Institute of City of Hope, Duarte, CA 91010

Communicated by Sidney W. Benson, April 11, 1986

ABSTRACT 2-Aminopurine (AP), a potent mutagenic base analogue, most frequently pairs with thymine. In the AP·T base pair, both bases adopt normal tautomeric forms. The mechanism for the mutagenic activity arises from its observed pairing with cytosine, which has been ascribed to an enhanced tendency to adopt the rare imino tautomeric form. NMR studies in H₂O on all the exchangeable protons in an oligonucleotide duplex containing an AP·T base pair show Watson-Crick hydrogen bonding. When the thymine is replaced by cytosine in the duplex, we observe an AP·C base pair. Both amino protons of AP are seen excluding the rare tautomeric form. Although several alternative structures are possible, it is shown that the second hydrogen bond is formed by protonation of the AP·C base pair and that this is the dominant species under physiological conditions.

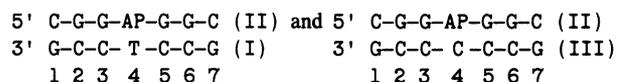
2-Aminopurine (AP), a base analogue of adenine, is a strong mutagen causing A·T → G·C and G·C → A·T transitions (for a review, see ref. 1). During DNA synthesis, AP preferentially forms base pairs with thymine (2, 3). The mutagenicity of AP occurs because AP·C mispairs are found at much higher frequency than A·C mispairs (3). Mispairing of AP with cytosine is the proposed intermediate in the two mutagenic pathways (4, 5). By analogy with mispairing schemes involving rare tautomeric forms of bases originally proposed by Watson and Crick (6, 7), Freese (4) in 1959 proposed that AP-induced mutagenesis resulted from an enhanced tendency of the analogue by comparison with adenine to exist as a disfavored imino tautomer. In spite of a lack of direct experimental evidence, the disfavored tautomer structure is the long-established explanation for the mutagenicity of AP.

Studies using purified polymerases show that AP substitution for adenine in DNA results in a large increase in dCMP misincorporation (3). UV spectral and melting analyses of synthetic DNA polymers containing AP·C base pairs (8) suggested that the mispair is stabilized by two hydrogen bonds, an expected 2-2 H bond between the amino group of AP and the keto oxygen of cytosine, and a shared 1-3 proton, which is not normally present when either base is present as the favored amino tautomer.

Disfavored tautomers involving either AP or cytosine are only two among several possible structures for AP·C stabilized by two H bonds (Fig. 1). Indeed, theoretical calculations (9) indicate that AP is no more likely to form the rare tautomer than adenine or cytosine. Recent thermodynamic calculations (10, 21) suggest that the base pair between AP and cytosine may be stabilized by protonation with both

bases remaining as favored amino tautomers. A further pairing possibility, prompted by recent NMR (11) and crystallographic data (12) on the T·G mismatch, would invoke a wobble base pair between AP and cytosine.

We have synthesized two heteroduplex structures



containing either an AP·T or an AP·C base pair. The proton NMR results demonstrate that whereas AP·T forms a normal base pair, AP·C exists predominantly as a protonated base pair as opposed to an imino tautomer or wobble base pair at physiological pH.

EXPERIMENTAL

The three oligonucleotides were synthesized as described (22).

NMR Spectroscopy: Spectra were recorded at 500 MHz on a Bruker WM-500 at the Southern California Regional NMR facility at the California Institute of Technology (Pasadena, CA). The solvent peak was suppressed by a 1-t-1 hard pulse sequence (13). For most experiments, the carrier was placed at ≈11.5 ppm. For some spectra, it was placed much closer to water (see text).

Oligonucleotides were 4 mM in strand concentration and were dissolved in 90% H₂O/10% ²H₂O/150 mM NaCl/10 mM phosphate, pH 7.2 (unless otherwise stated). Chemical shifts were reported relative to internal tetramethylammonium chloride at 3.18 ppm.

RESULTS

The determination of the mode of base pairing between AP and cytosine depends on the identification and characterization of the exchangeable protons of the AP·C base pair. We first measured the spectra in H₂O of a number of model compounds that relate to possible structures for the AP·C base pair.

At pH 7.2, N³-methyl 2'-deoxycytidine (3 MedC), which is a model compound for the rare tautomer of deoxycytidine, exists predominantly in the protonated form (pK 8.83) unlike deoxycytidine (pK 4.3). Even at 2°C, we do not observe any exchangeable proton resonances for 3 MedC in the NMR spectrum, which indicates that they must exchange very

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AP, 2-aminopurine; 3 MedC, N³-methyl 2'-deoxycytidine; NOE, nuclear Overhauser effect; WC and nWC, protons involved in hydrogen-bonded and non-hydrogen-bonded Watson-Crick base pairs.

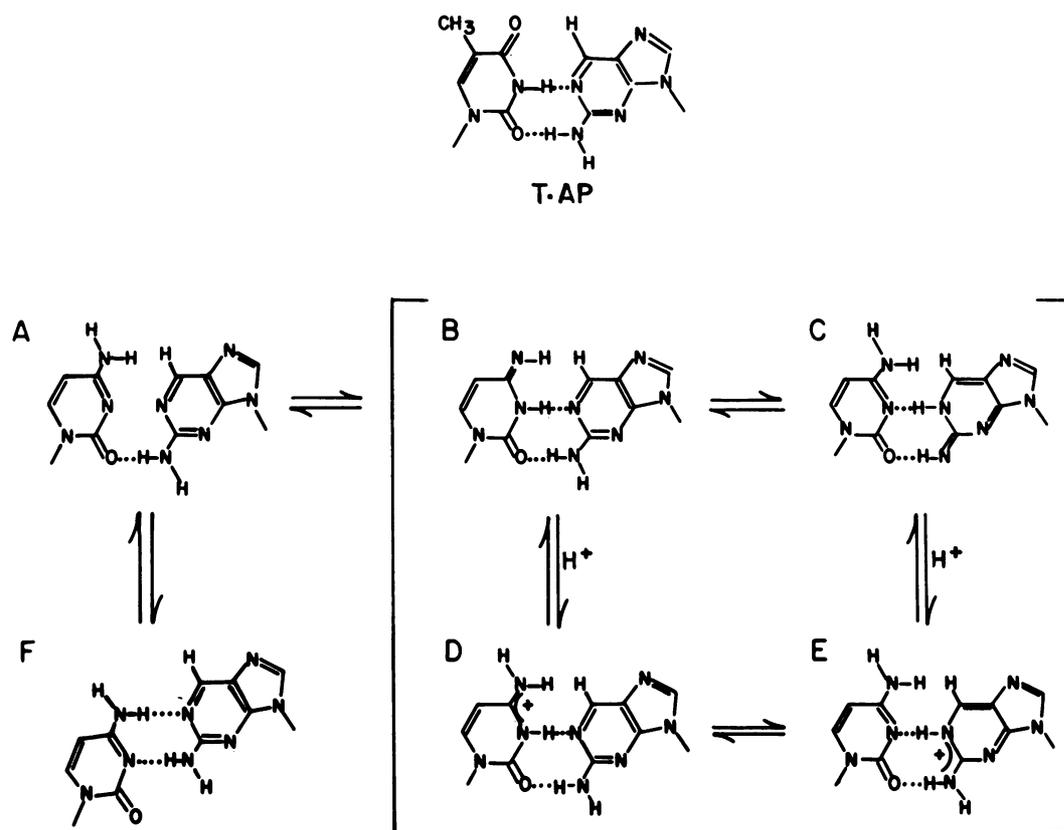


FIG. 1. Hydrogen-bonding scheme between AP and thymine. Possible hydrogen bonding schemes (A-F) between AP and cytosine.

rapidly with water. On lowering the pH to 2.1, we find two sharp resonances corresponding to the two amino protons (Table 1). At pH 11.2, 3 MedC is predominantly in the neutral rare imino tautomer form. Even though the base concentration is high, two sharp resonances of equal intensity, corresponding to the N-4 exocyclic imino protons in the *cis* and *trans* orientations are observed. It should be noted that both the exchangeable and H-6 proton resonances shift considerably between the protonated and rare tautomeric forms. For comparison, the corresponding chemical shifts for deoxycytidine at pH 7.2 are given in Table 1. At the same pH, AP gives one sharp (linewidth, 15 Hz) resonance corresponding to the two amino protons. The amino group of AP, like those of guanine and adenine, rotates rapidly on a proton NMR time scale.

AP-T Base Pairing. We next examined the base pairing between AP and thymine in the helix formed by mixing the oligonucleotides I and II. NMR spectra were recorded in H₂O at 18°C under which conditions, as we have shown (14), the guanine amino proton resonances will be very broad. Between 5°C and 20°C, the hydrogen-bonded Watson-Crick (WC) and the non-hydrogen-bonded non-Watson-Crick (nWC) protons exchange with each other via rotation about the C-2 amino bond and this exchange is in the intermediate

NMR time scale and the rate is independent of the melting temperature of the DNA.

The resolution-enhanced spectrum of the DNA containing the AP-T base pair is shown in Fig. 2. The exchangeable proton resonances are identified by comparison with a spectrum in ²H₂O (not shown). At low field, five resonances are observed, one of which, at 13.6 ppm, is in the normal region for the thymine imino proton of A-T base pairs. The other four resonances are in the expected region for guanine imino protons of the G-C base pairs. At this temperature (18°C), fraying of the ends of the helix (15) results in rapid exchange of the imino protons of the terminal base pairs but not (14) of the cytosine amino protons of these base pairs. The terminal imino protons are observed, although still broad, at 0°C.

Between 7.95 and 8.6 ppm, six resonances arising from exchangeable protons and one from a nonexchangeable proton are observed. This is where we would expect to observe the six WC cytosine amino proton resonances. In the region 6-7 ppm, we find seven exchangeable proton resonances. It is in this region that we would expect to find the six cytosine nWC proton resonances. If rotation of the AP amino group is sufficiently slowed by H-bond formation, two other exchangeable proton resonances could also appear in the spectrum.

Presaturation of a guanine imino proton resonance will give rise to a large nuclear Overhauser effect (NOE) to the WC cytosine amino proton of the base pair and a small one to the nWC proton. Small NOEs may be observed on the exchangeable proton resonances of the adjacent base pairs and on AP H-6 from base pairs 3 and 5.

Presaturation of the resonance at 12.774 ppm gives rise to a large NOE at 8.235 ppm (Fig. 2b) and a smaller one at 6.660 ppm. These must be the two cytosine amino protons. A NOE is observed on a nonexchangeable proton resonance at 7.821 ppm, which can only be the H-6 of AP and shows that the imino proton belongs to base-pair 3 or 5. A small NOE is

Table 1. Chemical shifts observed for model compounds

Compound	pH	Chemical shift, ppm	
		Amino/imino	H-6
Deoxycytidine	7.2	7.22, 6.75	7.82
3 MedC	2.1	9.46, 8.85	8.08
	7.1	—	8.08
	11.2	7.76, 7.29	7.25
AP	7.2	5.91	8.61

Solutions were 10 mM in phosphate and spectra were recorded at 2°C.

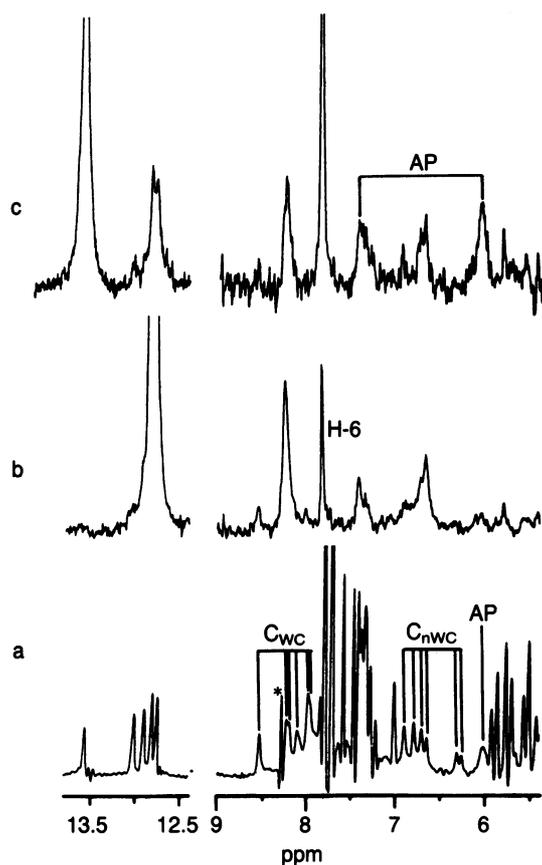


FIG. 2. (a) Resolution-enhanced spectrum (5.4–9.0 ppm and 12.4–13.8 ppm) of the AP-T duplex in 10 mM phosphate/150 mM NaCl/H₂O/²H₂O (9:1), pH 7.2, at 18°C. The peak marked with an asterisk is an artefact. Difference spectra (b) after 0.5-sec presaturation of the guanine imino proton at 12.77 ppm of base-pair 5 and (c) after 0.5-sec presaturation of the thymine imino proton at 13.60 ppm.

observed on a WC amino resonance at 8.538 ppm of an adjacent G-C pair. In addition, a NOE is seen on a resonance at 7.41 ppm, which, from its linewidth, must be an exchangeable proton.

When the lowest field imino resonance is presaturated a large NOE is observed on the H-6 resonance of AP (Fig. 2c), showing that this is the imino resonance of the AP-T base pair. NOEs are observed to the two adjacent G-C imino resonances and to their WC and nWC cytosine amino resonances at 8.2 and 6.6–6.7 ppm, respectively. The NOEs observed at 7.4 and 6.0 ppm must correspond to the amino protons of AP.

All of the imino and WC cytosine amino protons were systematically presaturated and, using known DNA inter-proton distances, the assignment of all the exchangeable proton resonances was made and is given in Table 2. The

relative magnitude of the NOEs is entirely consistent with the expected base pairing shown in Fig. 1.

AP-C Base Pairing. Mixing oligonucleotides II and III gives a DNA with an AP-C base pair. The resolution-enhanced spectrum is shown in Fig. 3a. At 14°C, only four resonances are observed in the normal imino region; an additional two broad resonances are observed at 0°C, as for the AP-T duplex. Seven exchangeable proton resonances are observed between 8.2 and 9.4 ppm and a further seven are in the region 6.4–7.2 ppm.

Systematic presaturation of all the G-C imino resonances at 14°C and 0°C gives the assignment of these imino and the cytosine amino resonances of these six base pairs; these are given in Table 2. The two remaining exchangeable resonances at 9.384 and 6.60 ppm belong to the AP-C base pair. Presaturation at 9.384 (Fig. 3b) gives a large NOE to a nonexchangeable proton resonance at 7.529 ppm, which is assigned to the H-6 of AP. NOEs to resonances of the two adjacent base pairs are observed at 13, 8.4, and 6.9 ppm. The remaining NOE at 6.60 ppm must arise from a proton of the AP-C base pair. It is possible that another small NOE is observed at 5.67 ppm.

When this experiment was repeated with the carrier offset placed at 6.4 ppm, which allows observation closer to the water resonance (Fig. 3c), the NOE to the resonance at 5.67 ppm is more clearly seen. The apparent negative phase of the irradiated resonance at 9.37 ppm is due to the fact that it lies in the second excitation envelope. In addition to the null created at the water position, there is a second one at ≈7.9 ppm, which results in the loss of a number of the NOEs seen in the previous experiment. This experiment was carried out with a 1-sec presaturation of the resonance at 9.38 ppm, under which conditions spin diffusion effects can be important. When repeated with a 0.2-sec presaturation, the NOE at 5.67 ppm was very small compared to that observed at 6.60 ppm, showing a strong spin diffusion contribution via the resonance at 6.60 ppm.

When the resonance at 6.60 ppm is presaturated (Fig. 3d), the reverse NOE at 9.38 ppm is observed but no NOE is observed to the H-6 of AP. In this experiment, the adjacent nWC cytosine amino resonance was partly saturated, resulting in the NOE to its WC proton at 8.23 ppm. The experimental conditions, irradiating close to the water resonance, precluded observation above 6 ppm.

We have only been able to identify three resonances of exchangeable protons, each resonance integrating for one proton, from the AP-C base pair.

To examine the stability of the AP-C base pair we have measured, where possible, the linewidths of the exchangeable proton resonances as a function of temperature, and these are given in Fig. 4.

We have also recorded the spectra at 14°C for a stepwise pH titration in which the pH was varied between 7.2 and 5.5. At pH 6.6, the only observable changes in the spectrum arise from resonances of the AP-C base pair. The resonance at 9.38

Table 2. Chemical shifts (ppm) of all the observed exchangeable protons and of the H-6 proton of AP measured at 18°C for the AP-T duplex and at 14°C for the AP-C duplex

		AP-T						AP-C					
5'	3'	Imino	C _{WC}	AP _{WC}	C _{nWC}	AP _{nWC}	AP(H-6)	Imino	C _{WC}	AP _{WC}	C _{nWC}	AP _{nWC}	AP(H-6)
G	C		8.09		6.79				8.23		7.122		
C	G	12.915	8.002		6.32			13.146	8.316		6.495		
C	G	12.825	8.215		6.710			13.036	8.404		6.987		
N	AP	13.600		7.41		6.03	7.821	9.384		6.60		5.60	7.529
C	G	12.774	8.235		6.660			13.068	8.377		6.933		
C	G	13.030	8.538		6.908			13.255	8.777		7.167		
G	C		7.98		6.30				8.23		6.570		

Solutions were 4 mM in strand concentration, 150 mM NaCl/10 mM phosphate, pH 7.2, in H₂O/²H₂O (9:1).

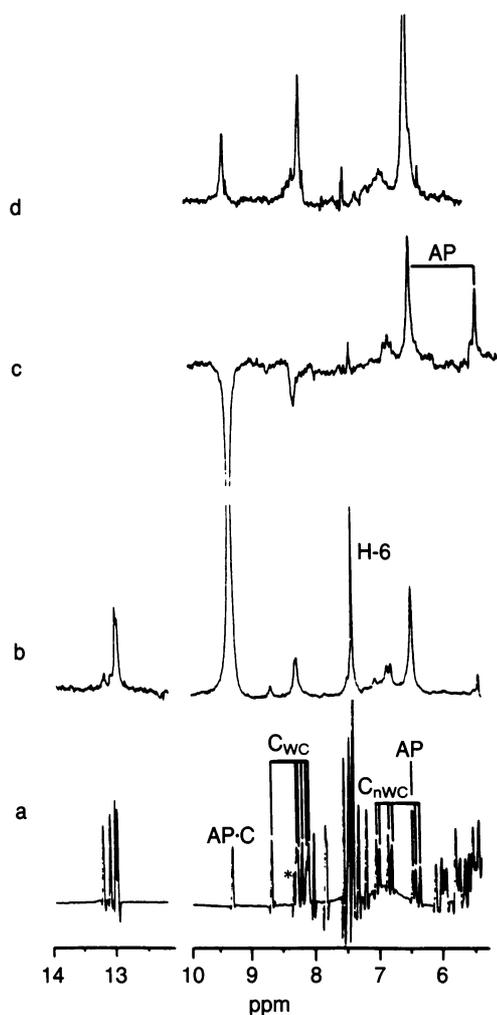


FIG. 3. Resolution-enhanced spectrum (5.5–10.0 ppm and 12.2–14.0 ppm) of the AP-C duplex in 10 mM phosphate/150 mM NaCl/H₂O²/H₂O (9:1), pH 7.2, at 14°C. The peak marked with an asterisk is an artefact. Difference spectra (b) after 0.5-sec presaturation of the AP-C imino proton at 9.38 ppm; (c) same as b except with carrier offset at 6.4 ppm and 1-sec presaturation (see text); (d) after 0.5-sec presaturation of the AP WC amino proton at 6.60 ppm.

ppm broadens to 35 Hz and shifts 0.02 ppm upfield. That at 6.60 ppm broadens similarly and shifts 0.02 ppm downfield. The H-6 of AP shifts also to 0.02 ppm downfield. We cannot observe the resonance at 5.67 ppm except in difference spectra. At pH 6.0, the two exchangeable resonances are broadened to 75 Hz and each has shifted ≈ 0.05 ppm from pH 7.2. The H-6 resonance has shifted 0.03 ppm.

At pH 5.5, the two exchangeable resonances of the AP-C base pair are no longer observable. The H-6 proton of AP has shifted a total of 0.04 ppm. The largest shift observed on a cytosine H-6 resonance, that at 7.43 at pH 7.2, is 0.04 ppm downfield.

DISCUSSION

For the AP-T base pair, we have observed three exchangeable proton resonances. The interproton NOEs demonstrate that the expected base-pairing scheme is the one adopted by this base pair. Hydrogen bonding between one of the amino protons and the thymine carbonyl oxygen slows down the rate of rotation of the AP amino group. However, the amino resonances show linewidths of ≈ 80 Hz and are much broader than those of the cytosine amino resonances, ≈ 20 Hz. The excess linewidth is most probably due to rotation rather than exchange as the thymine imino proton is much narrower. The

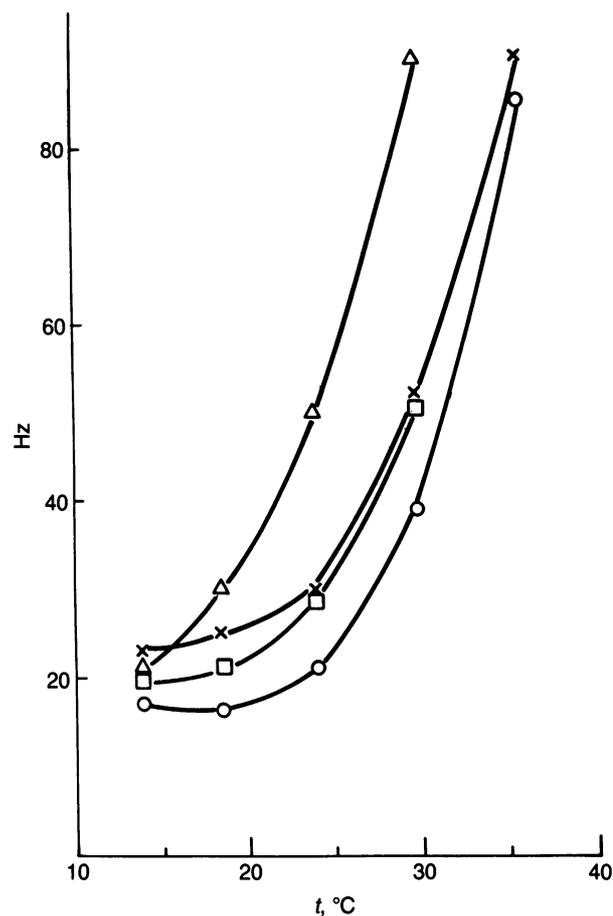


FIG. 4. Variation of linewidth as a function of temperature for the AP-C duplex exchangeable protons. Imino resonances of base pairs 2 and 6 (Δ), base pairs 3 and 5 (\circ), of the AP-C base pair (\times), and the AP nWC proton (\square).

half-life for rotation is thus ≈ 5 msec, much longer than for guanine amino groups at this temperature (14, 16).

Similarly, we have observed three exchangeable proton resonances for the AP-C base pair, but their assignment is less obvious. The proton that gives rise to the resonance at 9.37 ppm must lie close to the helix axis from the relative magnitude of the NOEs observed in Fig. 3, to the neighboring G-C base pairs. It is not a proton of an amino group despite the chemical shift as the NOE to the resonance at 6.60 ppm is 1/4th to 1/5th the NOEs observed between WC and nWC cytosine amino protons under the same conditions. This proton gives a large NOE to the H-6 of the AP and must arise from an imino type proton between cytosine N-3 and AP N-1. It should be noted that the H-6 of AP is the only nonexchangeable proton that can give an NOE from the G-C imino protons of the base pairs above and below in the helix. Presaturation of the resonance at 6.60 ppm gives no NOE to the H-6 of AP and must therefore be one of the AP amino protons. The relative magnitude of the NOEs at 6.60 and 5.67 ppm upon varying the AP-C imino presaturation time shows that the resonance at 6.60 ppm is the WC AP amino resonance, and that at 5.67 ppm is the nWC proton resonance. In none of our experiments could we find a resonance or resonances that could be attributed to the amino protons of cytosine in the AP-C base pair.

Clearly, AP and cytosine are base-paired. The structure A (Fig. 1), in which only one hydrogen bond is formed, can be readily eliminated, as we observe an imino type proton. It would also give four exchangeable proton resonances, and we only observe three. The same observations argue against the wobble base pair, structure F. If the resonance at 9.38

ppm were assigned to the WC cytosine amino proton, because we observe a large NOE to the AP H-6, then the nWC proton must appear in the spectrum as for all the other cytosine residues. No such resonance is observed and also the relative magnitude of the interbase-pair NOEs observed upon presaturation of this resonance are not consistent with it being a WC cytosine amino proton resonance. Also, the rare tautomer of AP, structure C, can be excluded as a major species as we have observed both AP amino proton resonances. In structures D and E, the proton between AP N-1 and cytosine N-3 will lie close to the helix axis like the guanosine imino protons. However, it is not a true imino proton, as in both cases (structures D and E) the hybridization of the nitrogen to which it is attached is between that of an imino and amino proton. This may in part well account for the unusually high chemical shift observed.

Proton exchange is believed to take place only from the open state and yet it must take place from the closed state for the AP-C base pair as, in structures B-E, there are either four or five exchangeable protons. We observed slow exchange with water for the exocyclic imino proton of 3 MedC even at pH 11. This proton would definitely exhibit slow exchange for the neutral rare tautomeric species at pH 7. At pH 11, the *cis* and *trans* forms for 3 MedC were observed to be equally populated. The absence of the cytosine resonances cannot be explained by the rare tautomeric species. Even if, and there is no reason to suppose that it would, it exists entirely in the *trans* isomer relative to N-3, it would be seen from the presaturation experiments on WC cytosine amino proton resonances above and below in the helix.

Not only is it considered that base-pair opening is a prerequisite for proton exchange in normal base pairs but that transient protonation of the base must occur (17) before a suitable base catalyst can accept a proton. That base protonation has a dramatic effect on exchange rates is shown by comparison of deoxycytidine and 3 MedC at pH 7.2. The protonation prerequisite for exchange of the cytosine amino protons is already fulfilled in structure D without the need for base-pair opening. Protonation of cytosine would accelerate the exchange of the amino protons, which are solvent-accessible, but not that of the hydrogen-bonded imino-type proton.

We believe that our results clearly demonstrate the existence of the protonated base pair, but it could be argued that a small percentage of structure D in rapid equilibrium with structure B would also account for the data. We have looked for evidence of such an equilibrium. If it exists, we should push the equilibrium toward the protonated base pair by lowering the pH. We would expect some significant chemical shift changes on going from the rare tautomer form to the protonated base. In the model compound (Table 1), the cytosine H-6 resonance moves 0.83 ppm upon shifting the equilibrium from rare tautomer to protonated base. On lowering the pH from 7.2 to 5.5, the largest upfield shift upon any of the cytosine H-6 resonances was 0.04 ppm. If the rare tautomeric form exists under physiological conditions, its concentration must be extremely small. The observed enhanced exchange of the exchangeable resonances and shift of the H-6 of AP at pH 5.5 can be accounted for by acid denaturing and thus more frequent base pair opening for AP-C.

We have measured at pH 7.2, where possible, the linewidths of the exchangeable proton resonances as a function of temperature (Fig. 4). The increase observed in the linewidths of the AP-C imino proton and WC amino proton of AP follows closely that of the imino type protons of base pairs

3 and 5. There does not appear to be any preferential melting out of the AP-C base pair relative to the adjacent base pairs.

CONCLUSION

The observations reported here show that the dominant form for the AP-C base pair is the protonated structure. From our results, we cannot quantify the equilibrium between protonation on cytosine or AP, although the rate of proton exchange for the cytosine amino protons and the greater basicity of cytosine (pK: deoxycytidine, 4.3; dAP, 3.8) are more consistent with cytosine protonation. The protonated base pair is different from the rare tautomer model commonly accepted. It is important to note that, to date, there is no direct structural evidence for the existence of rare tautomers of normal bases, whereas protonated cytosine residues in DNA duplex molecules at physiological pH have been reported both for the foldback complex of poly(dC) (18) and for the type I silver complex with G-C base pair (19, 20).

We thank Ms. Sarah E. Wright for expert preparation of the manuscript. Work reported here was supported by Grants GM 33863 and GM 21422 from the National Institutes of Health; work performed at the Southern California Regional NMR facility at the California Institute of Technology was supported by Grant CHE 84-40137 from the National Science Foundation; we wish to acknowledge the kind support of Dr. James P. Yesinowski in enabling us to utilize the facility in a timely and efficient manner. L.C.S. was supported in part by National Research Service Award Grant 2T32CA09320-04.

- Roneš, A. (1979) *Mutat. Res.* **75**, 1-47.
- Bessman, M. J., Muzyczka, N., Goodman, M. F. & Schnaar, R. L. (1974) *J. Mol. Biol.* **88**, 409-421.
- Watanabe, S. M. & Goodman, M. F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2864-2868.
- Freese, E. (1959) *J. Mol. Biol.* **1**, 87-105.
- Rudner, R. (1960) *Biochem. Biophys. Res. Commun.* **3**, 275-280.
- Watson, J. D. & Crick, F. H. C. (1953) *Nature (London)* **171**, 964-967.
- Watson, J. D. & Crick, F. H. C. (1953) *Cold Spring Harbor Symp. Quant. Biol.* **18**, 123-131.
- Goodman, M. F. & Ratliff, R. L. (1983) *J. Biol. Chem.* **258**, 12842-12846.
- Danilov, V. I., Kruglyak, Y. A., Kuprievich, V. A. & Shramko, O. V. (1967) *Biofizika* **12**, 726-729.
- Sowers, L. C. (1983) Dissertation (Duke University, Durham, NC).
- Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K. & Breslauer, K. (1982) *Biochemistry* **21**, 437-444.
- Brown, T., Kennard, D., Kneale, G. & Rabinovich, D. (1985) *Nature (London)* **315**, 604-606.
- Bleich, H. & Wilde, J. (1984) *J. Magn. Reson.* **56**, 154-155.
- Fazakerley, G. V., van der Marel, G. A., van Boom, J. H. & Guschlbauer, W. (1984) *Nucleic Acids Res.* **12**, 8269-8279.
- Patel, D. J. & Hilbers, C. W. (1975) *Biochemistry* **14**, 2651-2660.
- Fazakerley, G. V., Téoule, R., Guy, A., Fritzsche, H. & Guschlbauer, W. (1985) *Biochemistry* **24**, 4540-4548.
- McConnell, B. (1978) *Biochemistry* **17**, 3168-3176.
- Courtois, Y., Fromageot, P. & Guschlbauer, W. (1968) *Eur. J. Biochem.* **6**, 493-501.
- Eichhorn, G. L., Butjow, J. J., Clark, P. & Tarien, E. (1967) *Biopolymers* **5**, 283-296.
- Dattagupta, N. & Crothers, D. M. (1981) *Nucleic Acids Res.* **9**, 2971-2985.
- Sowers, L. C., Ramsay-Shaw, B., Veigl, M. & Sedwick, W. D. (1986) *Mutat. Res.*, in press.
- Eritja, R., Kaplan, B. E., Mhaskar, D., Sowers, L. C., Petruska, J. & Goodman, M. F. (1986) *Nucleic Acids Res.*, in press.