TAUTOMERIC EXCHANGE IN CYTOSINE

Gregory C. Y. Lee, J. H. Prestegard*, and Sunney I. Chan

Noyes Laboratory of Chemical Physics†, California Institute of Technology, Pasadena, California 91109

Received March 18, 1971

Summary. It has long been noted that the H_5 proton of cytosine and its related derivatives exhibits unusually broad resonances in the pmr spectrum under certain conditions of temperature and solution pD. We have examined this phenomenon as a function of temperature, solution pD as well as the external magnetic field, and have shown that the line broadening arises from chemical exchange between the amino and imino tautomers of the cytosine base. The observation of sharp H_5 resonances in cytosine derivatives existing in only the amino or the imino tautomeric structure supports this interpretation. The imino tautomer was estimated to be present to the extent of $15 \pm 3\%$ at room temperature in neutral aqueous solution.

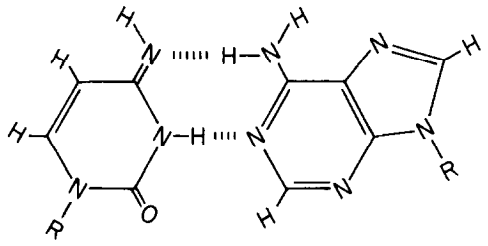
The genetic information carried in the DNA molecule is occasionally misread in duplication or in transcription for protein synthesis, producing either a mutant strain or an abnormal protein. It was suggested some time ago that such errors are the result of the presence of a minor tautomeric form of the normal purine or pyrimidine base in the nucleic acid.[1] The existence of these tautomers has since been demonstrated.[2-9] More recently, however, the wobble hypothesis has been used to explain transcription errors.[10] However, the tautomeric explanation of mutation remains appealing, especially for the replication process, since it allows formation of an abnormal C-A base pair, for example, in a geometry identical to that of a normal C-G pair (Fig. 1).

Recently, experiments in this laboratory have indicated that nmr may be of use in furthering this investigation. It has long been observed that some biological bases of DNA exhibit unusually broad proton resonances under certain

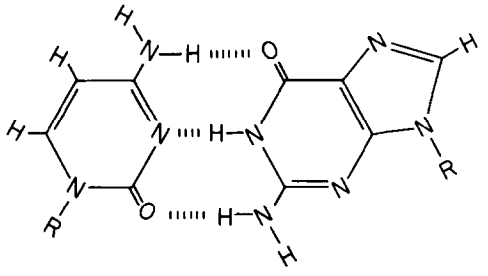
---

* Present address: Department of Chemistry, Yale University, New Haven, Connecticut 06520.

† Contribution No. 4226.



C — A



C — G

Fig. 1. (a) Abnormal C-A base pair; (b) normal C-G base pair.

conditions of pH and temperature.[11] For example, in the case of cytosine, the H_5 resonance is often broadened. The broadening of nuclear resonance lines is usually attributed to chemical exchange, molecular aggregation, complex formation with paramagnetic ions, and/or spin-spin coupling to quadrupolar nuclei.[12] In this communication, we wish to present experimental evidence to indicate that the H_5 resonance broadening observed for cytosine and some of its derivatives is the result of chemical exchange between the amino and imino tautomeric structures of the cytosine base.

Samples of cytidine-5′-monophosphate (5′-CMP), cytidine, 3-methyl cytidine, and 4-N, N-dimethyl cytosine were purified by recrystallization.[13] Cation exchange was used to replace possible ion contaminants with sodium ion. The 100 MHz pmr spectra of these compounds at 0.02 M in D_2O were taken as a function of both pD and temperature. As expected, changes in chemical shifts of the cytosine H_5 and H_6 resonances occur around pD 4.5 as

the cytosine ring becomes protonated. For the present work, however, of more immediate interest are the marked changes in the H_5 linewidths of both 5′-CMP and cytidine which occur with changing pD and temperature. The linewidth of the H_5 resonance as a function of pD at 30°C and as a function of temperature at pD 6.0 is shown in Figs. 2 and 3, respectively. At 30°C the H_5 resonance is narrow at both ends of the pD scale, but between pD 4.0 and pD 7.0, it broadens, reaching a maximum width of 6.0 Hz at pD 5.8. Within the same pD range, the H_6 resonance also broadens somewhat, although to a significantly lesser extent. No detectable broadening was, however, observed for the $H_{1'}$ resonance. When the H_5 linewidth was examined as a function of temperature at pD ~6, maximum broadening was observed at 0°C. At this temperature, the H_5 resonance exhibits a width of ~8 Hz. Above and below 0°C, the H_5 linewidth becomes progressively narrower, reaching 3.0 Hz at 60°C and 5.0 Hz at -15°C. Similar experiments with 3-methyl cytidine and 4-N, N-dimethyl cytosine revealed no broadening of the H_5 resonance in these methylated derivatives.

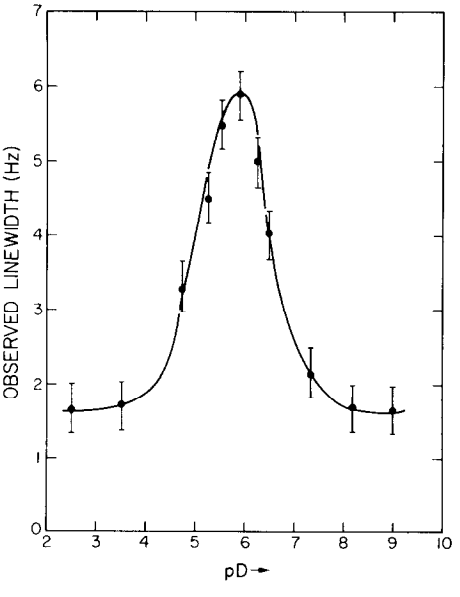


Fig. 2. pD dependence of the linewidth of the H_5 resonance of 5′-CMP (0.02 M) at 30°C.

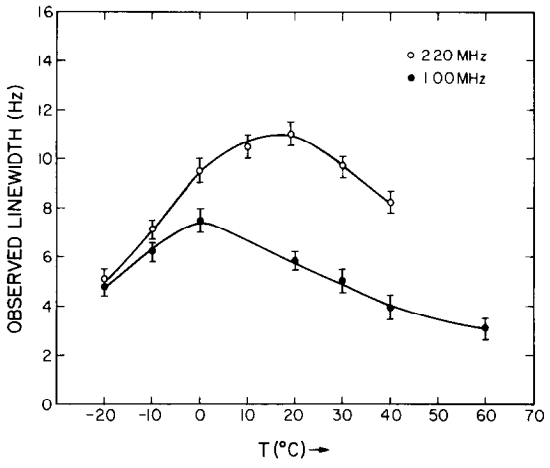


Fig. 3. Frequency and temperature dependence of the linewidth of the H_5 resonance of a 0.02 M 5'-CMP solution containing 5 M NaCl in D_2O at pD 6.0.

On the basis of these and other experiments, many of the possible origins of this previously unexplained spectral broadening can be eliminated. Line broadening by paramagnetic ion complexation can be ruled out not only because we have taken measures to remove such ions from our samples, but also because, with the exception of EDTA, the addition of efficient chelating agents even at concentration levels of 10^{-3} M does not lead to reduction of the H_5 linewidth. Moreover, if the broadening arises from the presence of paramagnetic ion contaminants, the temperature and frequency (vide-infra) dependences which we have observed would suggest that the rate of chemical exchange of the cytosine base between a complexed and free environment is slow compared with the paramagnetic contact shift at low temperatures, and that at the higher temperatures the linewidth is controlled by the rate of relaxation through a change in the precession frequency.[14] At the concentration levels of paramagnetic ions possible ($< 10^{-6}$ M), however, the observed line broadening would imply a nuclear hyperfine coupling constant[15] of $|A/h| \gg 10$ MHz, a value which is at least one or two orders of magnitude too large. Hence, this interpretation of the line broadening is unlikely. Finally, with the excep-

tion of Cu^{2+}, we have observed that the addition of common paramagnetic ions, such as Mn^{2+}, Fe^{3+}, Ni^{2+}, Co^{2+} does not in general lead to specific broadening of the cytosine H_5 resonance, and even then, line broadening is only apparent at ion concentration levels greater than 10^{-5} to 10^{-4} M. In the case[16] of Cu^{2+}, specific H_5 line broadening is indeed observed at ion concentration $\sim 2 \times 10^{-5}$ M, but here the linewidth exhibits a temperature dependence which is quite different from that which we have observed for the undoped solution, as expected.

Of the remaining mechanisms, the absence of a concentration dependence of the H_5 linewidth over the concentration range 0.01 to 0.2 M excludes broadening by molecular aggregation. We can also eliminate effects arising from indirect coupling of the H_5 and H_6 protons to the various quadrupolar ^{14}N nuclei, since the H_5 linewidth is independent of the viscosity of the solution and there is no broadening of the H_5 resonance in 4-N, N-dimethyl cytosine and 3-methyl cytidine, two cytosine derivatives with seemingly rather similar electronic structure to the parent molecule.

Our linewidth data on the H_5 resonance can be qualitatively interpreted in terms of chemical exchange between the amino and imino tautomeric structures of the cytosine base. The absence of H_5 broadening in the two methylated derivatives, 4-N, N-dimethyl cytosine, where the cytosine base can exist only in the amino form, and 3-methyl cytidine, where the base is frozen in the imino structure, supports this interpretation. Within the framework of this interpretation, the decrease of the H_5 linewidth above 0°C would indicate that the chemical exchange is fast in this temperature region. The well-known limiting expression[17]

$$1/\pi\, T_2' = 1/\pi\, T_2 + 4\pi\, p_A^2\, p_B^2\, \Delta^2\, (\tau_A + \tau_B) \qquad (1)$$

should then apply. Here $1/\pi\, T_2'$ is the observed linewidth; $1/\pi\, T_2$ is the width in the absence of chemical exchange; p_A, p_B denote the populations of the major and minor species, respectively; τ_A and τ_B are their corresponding pre-exchange lifetimes; and $\Delta = \delta_A - \delta_B$ (Hz) is the chemical shift difference

of the proton in question between the two environments. Since both τ_A and τ_B are expected to decrease with increasing temperature, the linewidth should decrease with increasing temperature in this temperature region, as observed. Below 0°C, the data suggest that we are approaching the slow exchange limit. We would then be monitoring the major species, which should exhibit a resonance linewidth given by[17]

$$1/\pi\ T_2' = 1/\pi\ T_2 + 1/\pi\ \tau_A \qquad (2)$$

In this limit, one might expect to see distinct resonances for the minor species, B, which, in the case of cytosine, should be the imino tautomer. But, if the percentage of the minor species is low, say 10%, these resonances would be much broader than those of the major species and might escape detection because of low signal intensities. Nevertheless, if the percentage of the minor species is not too low, we expect some reduction in the integrated intensity of the resonances of the major species from that at higher temperatures, to an extent at least commensurate with the population of the minor species. We do observe such an intensity reduction. At 30°C, the ratio of the integrated intensity of the H_5 resonance to that of the $H_{1'}$ resonance was observed to be 0.99 ± .03; at -15°C, this intensity ratio is reduced to 0.89 ± 0.03.

The frequency dependence of the linewidth predicted by equation (1) provides a stringent test of the chemical exchange hypothesis. Since Δ is directly proportional to the magnetic field strength, the exchange contribution to the linewidth at 220 MHz should be nearly five times the contribution at 100 MHz. We observe that the H_5 linewidth increases by a factor of three above 30°C, which, considering the possible limitations of the application of the fast exchange equation (1), supports our interpretation. We have also used equation (1) to predict the relative linewidths of the H_5 and H_6 resonances. Values of $\Delta = \delta_A - \delta_B$ for both the H_5 and H_6 protons can be estimated from the chemical shifts of cytidine and 3-methyl cytidine. Cytidine is believed to be predominantly in the amino form, so that the observed chemical shifts

should approximate δ_A. At pD 8.5, the nonexchangeable methyl group of 3-methyl cytidine freezes the cytosine base in the imino form, and the chemical shifts here may be used to approximate δ_B. At 100 MHz, the H_5 and H_6 resonances of 3-methyl cytidine are 20 Hz and 4 Hz, respectively, downfield from those of cytidine. Since Δ is 20 Hz for H_5 and only 4 Hz for H_6, the effect of exchange broadening on the linewidth of H_5 resonance should be much more pronounced than that of H_6, as observed.

The pD dependence of the H_5 linewidth is also consistent with the exchange broadening mechanism. Below pD 4.5 the cytosine base is protonated at position 3, which would prevent the formation of the imino tautomer. In other words, as the equilibrium is shifted to the protonated form, p_B becomes effectively zero. For pD's $>$ 6, both p_A and p_B are of course non-zero, but if, as might be expected, the exchange is base catalyzed (the presence of a base will aid the removal of protons from either the amino nitrogen or the N-3 ring nitrogen), the rate will be accelerated, correspondingly shortening τ_A and τ_B and reducing the observed linewidth.

We have presented experimental evidence to show that the unusual H_5 broadening observed in pmr spectra of cytosine and its related derivatives is due to the chemical exchange between the amino and imino tautomers of cytosine base. In a separate manuscript to be published elsewhere, we shall present a quantitative analysis of the experimental data presented here, and shall show that this broadening of the H_5 resonance is indeed very sensitive to the amount of the minor tautomer present. On the basis of this analysis, we have ascertained the kinetics of this tautomerism and have concluded that the imino tautomer is present to the extent of 15 ± 3% at room temperature in neutral aqueous solution. It is interesting to note that the percentage of minor tautomer indicated here by the present pmr study is significantly higher than that estimated by other workers using other methods.

Acknowledgment. This work was supported in part by Grant GM 14523-04 and -05 from the National Institute of General Medical Sciences, U. S. Public Health Service, and Grant GP-8540 from the National Science Foundation.

References

1. Watson, J. D. and Crick, F. H., Cold Spring Harbor Symp. Quant. Biol., 18, 128 (1959).
2. Kenner, G. W., Reese, C. B. and Todd, A. R., J. Chem. Soc., 855 (1955).
3. Katritzky, A. R. and Waring, A. J., J. Chem. Soc., 1540 (1962); 3046 (1963).
4. Wolfenden, R. V., J. Mol. Biol., 40, 307 (1969).
5. Brown, D. M. and Hewlins, M. J. E., J. Chem. Soc., C, 2050 (1968).
6. Pullman, A. and Pullman, B., Biochim. Biophys. Acta, 36, 343 (1959).
7. Hoffman, T. A. and Ladik, J., "Advances in Chemical Physics," Vol. 7, Interscience Publishers, New York (1964), Chapter 2.
8. Bodor, N., Dewar, M. J. S. and Harget, A. J., J. Am. Chem. Soc., 92, 2929 (1970).
9. Löwdin, P. O., "Advances in Quantum Chemistry," Vol. 2, Academic Press, New York (1965) p. 213.
10. Crick, F. H., J. Mol. Biol., 19, 548 (1966).
11. Bangerter, B. W. and Chan, S. I., J. Am. Chem. Soc., 91, 3910 (1969).
12. Abragam, A., "The Principles of Nuclear Magnetism," Oxford University Press (1962).
13. Wempen, I., Duschinsky, R., Kaplan, L. and Fox, J. J., J. Am. Chem. Soc., 83, 4755 (1961).
14. Swift, T. J. and Connick, R. E., J. Chem. Physics, 37, 307 (1962).
15. Bloembergen, N., J. Chem. Physics, 27, 595 (1957).
16. Eichhorn, G. L., Clark, P. and Becker, E. D., Biochemistry, 5, 245 (1966).
17. Pople, J. A., Schneider, W. G. and Bernstein, H. J., "High Resolution Nuclear Magnetic Resonance," McGraw-Hill, New York (1959), pp. 218-225.