Proc. Nat. Acad. Sci. USA
Vol. 69, No. 9, pp. 2488–2491, September 1972

Tautomerism of Uracil and Thymine in Aqueous Solution: Spectroscopic Evidence

(lactam/lactim/enol/keto/triplet states/fluorescence)

MALCOLM DANIELS

Istituto di Chimica Fisica, Università di Perugia, Italy

Communicated by J. D. Watson, May 24, 1972

ABSTRACT **Excitation spectra for triplet state formation and fluorescence emission from uracil and thymine in neutral aqueous solution at room temperature are anomalous when compared to the absorption spectra of uracil and thymine.**

The experimental data are critically examined with respect to three molecular models; *Model 1*, which is characterized by increased intersystem crossing from higher vibrational levels of S_1; *Model 2*, in which fluorescence is attributed to a (π,π^*) state, whereas intersystem crossing occurs from an ($n\pi^*$) state; and *Model 3*, in which fluorescence is attributed to a tautomer I, while triplet yields originate from tautomer II. Reasons are presented for discarding *Models 1* and *2*, and it is demonstrated that the observed excitation spectra complement each other with respect to the absorption spectrum, such that the quantum yields of triplet formation and fluorescence emission become independent of exciting wavelength. It is suggested that the fluorescing tautomer I has the N_3–C_4 enol structure, while the triplet-forming tautomer II is most likely the predominant diketo form.

Since the earliest days of the Watson–Crick formulation of the secondary structure of DNA, it has been recognized that mismatching of the base pairs could lead to errors in replication and could thus be a molecular mechanism for mutagenesis. Tautomerism of the bases in those groups of the molecule involved in inter-base hydrogen bonding (giving amino–imino forms of adenine, guanine and cytosine, and lactam–lactim forms of thymine and uracil) was an obvious mechanism for "natural" mismatching; this suggestion has spurred much effort in the search for evidence for tautomeric forms, particularly in aqueous solution. The net effect of investigations by UV absorption spectroscopy, infrared in D_2O, nuclear magnetic resonance, and Raman spectroscopy has been negative (1). However, these latter three methods are not sensitive to minor components, and UV spectroscopy is beset by the problem of unresolved overlapping transitions, so that it can only be said that the amino–lactam structures are the predominant forms of bases. As only minor amounts of tautomers would be necessary to make viable such a mechanism for the manifestation of mutagenesis through tautomeric mismatching, the problem really remains unresolved.

In recent years, two types of experiments have been performed on the pyrimidine bases to give excitation spectra that must be accounted for in terms of the absorption spectrum. Considered individually, each of these results are anomalous in that the excitation spectrum for the process is markedly different from the absorption spectrum. The purpose of the present communication is to demonstrate that in the light of the present knowledge, these anomalies can be quantitatively resolved if, and only if, the solution contains two distinct molecular species, tautomeric forms; reasons are presented for the suggestion that they are the diketo (lactam) form (major component) and the N_3–C_4 lactim form (minor component).

By studying the photochemical kinetics for dimer formation in aqueous solutions of uracil, Brown and Johns (2) were able to obtain from their reaction mechanism values for ϕ_T*. When determined at different wavelengths of excitation, the quantum yield did not have the expected independence of wavelength, but increased 13-fold from 280 to 230 nm (Fig. 1). This behavior is also exhibited by thymine and orotic acid (3), and has been confirmed recently by Lamola and Eisinger (4) using an entirely different technique in which triplet yields are determined by energy transfer to Eu^{3+} and monitored by subsequent emission from the excited state of Eu^{3+}. The same frequency dependence was found (4) for TMP as for thymine, and it was stated that similar behavior was found for 1,3-dimethylorotic acid, 1,3-dimethylthymine, and 3-methyl-TMP.

By use of a signal accumulation technique, the existence of fluorescences from all the purine and pyrimidine bases of the nucleic acids has recently been demonstrated (5), and preliminary confirmation of these results has appeared (6). The fluorescence excitation spectrum for uracil did not coincide with the absorption spectrum (Fig. 1 of ref. 5); consequently, the quantum yield of fluorescence also varied with wavelength of excitation, in the opposite sense, however to ϕ_T, decreasing by 3-fold from 282 to 256 nm (Fig. 1).

To account for this unusual behavior of the singlet and triplet states, three models have been proposed; they are now considered in turn.

Model 1 refers to a single molecular system with connected singlet and triplet states, and is illustrated in Fig. 2*a*. The increasing yield of triplets with increase in exciting frequency has been rationalized (2) by the proposition that intersystem

* ϕ_T: quantum yield for triplet formation. It was later shown (3) that the experimental value of ϕ_T determined by this method is related to the true quantum yield by ϕ_T (exp) = ϕ_T ($k_1/k_1 + k_1'$), where k_1 is the rate constant for reaction of the triplet leading to dimer and k_1' is the rate constant for triplet self-quenching. For uracil, the fraction ($k_1/k_1 + k_1'$) is about 0.5–1.0 and, as the rate constants are independent of wave-length, this point has no effect on the present considerations that concern the wavelength dependence of ϕ_T.

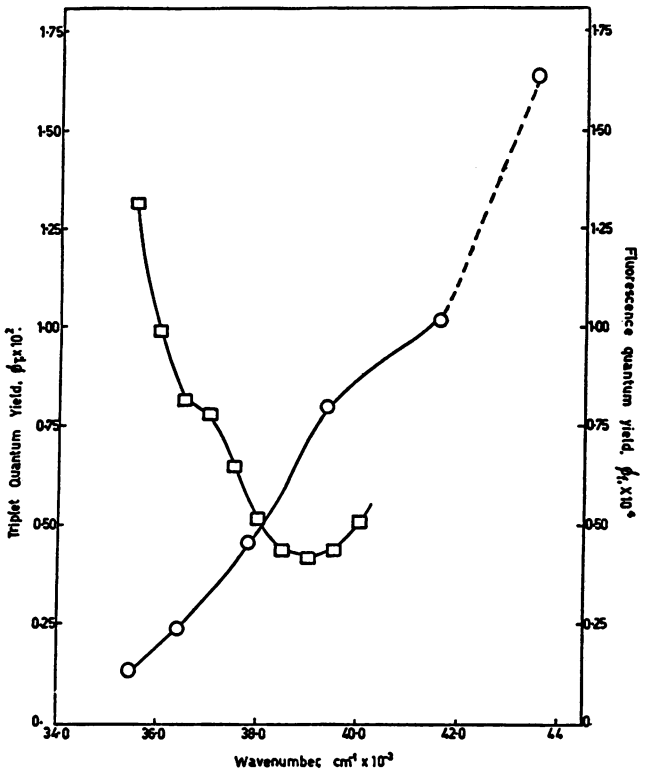


FIG. 1. Quantum yields of triplet formation, ϕ_T and of fluorescence emission, ϕ_f, from uracil as functions of exciting frequency. O——O, ϕ_T (taken from ref. 2); □——□, ϕ_f (calculated from data in ref. 5a).

crossing can occur from the higher vibrational levels of S_1, as well as by the more normal process from the lowest level of S_1. An attractive feature of this model is that it immediately provides an explanation for the form of variation of the fluorescence quantum yield. If there is a significant probability of intersystem crossing at the higher vibrational levels, the probability of populating the lowest level after excitation of the higher level must correspondingly decrease and, hence, so must ϕ_f. However, examined more closely, this model is not satisfactory. To compete effectively with vibrational deactivation, which is considered (7) to occur with lifetimes of about 10^{-12} sec, k'_{isc} will have to have values of about 10^{12} sec^{-1}. Recent values (8) for k^0_{isc} are 10^9 sec^{-1}, so that k'_{isc} must increase above k^0_{isc} by 10^3 to make the mechanism possible. More importantly, the model is quantitatively inconsistent with the experimental values. According to the model (Fig 2a), the triplet yield when excitation is to higher levels can be written (4)

$$\phi'_T = \left(\frac{k'_{isc}}{k_{vr} + k'_{isc}}\right) + \left(\frac{k_{vr}}{k_{vr} + k'_{isc}}\right)\phi_T{}^0$$
$$= f'_{isc} + (1 - f'_{isc})\phi_T{}^0,$$

where f'_{isc} is the probability of intersystem crossing from the higher level. Similarly (5b),

$$\phi_f' = (k_{vr}/k_{vr} + k'_{isc})\phi_f{}^0 = (1 - f'_{isc})\phi_f{}^0,$$

and f'_{isc} can thus be calculated from both the triplet and fluorescence yields. To validate the model, concordant results must be obtained. Over the range 35,500 to 39,000 cm^{-1}, the fluorescence yields give $f'_{isc} \simeq 0.7$, whereas the triplet yields give $f'_{isc} \simeq 0.6 \times 10^{-2}$. The discrepancy is so great that it must be concluded that the singlet and triplet levels that are observed experimentally cannot be related by this model and, thus, are not molecularly connected.

A more rewarding way of considering the experimental results is in terms of spectral distributions (Fig. 3). The cross-section for triplet formation σ_T (related to ϕ_T through the cross-section for absorption, σ_A i.e. $\phi_T = \sigma_T/\sigma_A$) clearly does not parallel the total absorption curve, and neither does the relative cross-section for fluorescence. An obvious approach is to regard these distributions as components of the overall experimentally determined absorption curve, and to attempt to identify their origin. This is the approach of *Models 2* and *3*.

Model 2. A two-component model for a single molecular system may be constructed in terms of $n\pi^*$ and $\pi\pi^*$ transitions (5). Although the low fluorescence quantum yield might suggest that fluorescence originates from an $n\pi^*$ state, studies of solvent effects on the fluorescence of thymine (W. Hauswirth and M. Daniels, unpublished results) show it to have the characteristics of $\pi\pi^*$ emission, and there is no reason to expect the emission from uracil to be different. Evidence available from studies at 77°K (9) indicates the triplet also to be $\pi\pi^*$ and the model can be completed if intersystem crossing occurs from an $n\pi^*$ state i.e., $^1(n\pi^*) \rightarrow {}^3(\pi\pi^*)$. The net results of these considerations are embodied in the energy-level scheme of Fig. 2b (note that although the $\bar{\nu}_{max}$ of the distributions of Fig. 3 are quite distinct, the thresholds are almost degenerate). The complete uncoupling

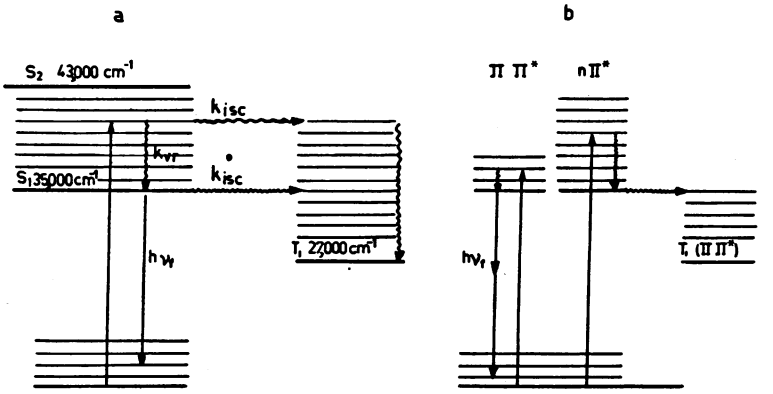


FIG. 2. Energy level diagrams illustrating (a) *Model 1* and (b) *Model 2*. In a the singlet and triplet levels are connected by intersystem crossing, with rate constant k^0_{isc} for the lowest level and k'_{isc} for a higher vibrational level. In b the $n\pi^*$ and $\pi\pi^*$ states are not connected (see text), and intersystem crossing $^1(n\pi^*) \rightarrow {}^3(\pi\pi^*)$ predominates over the process $^1(\pi\pi^*) \rightarrow {}^3(\pi\pi^*)$.

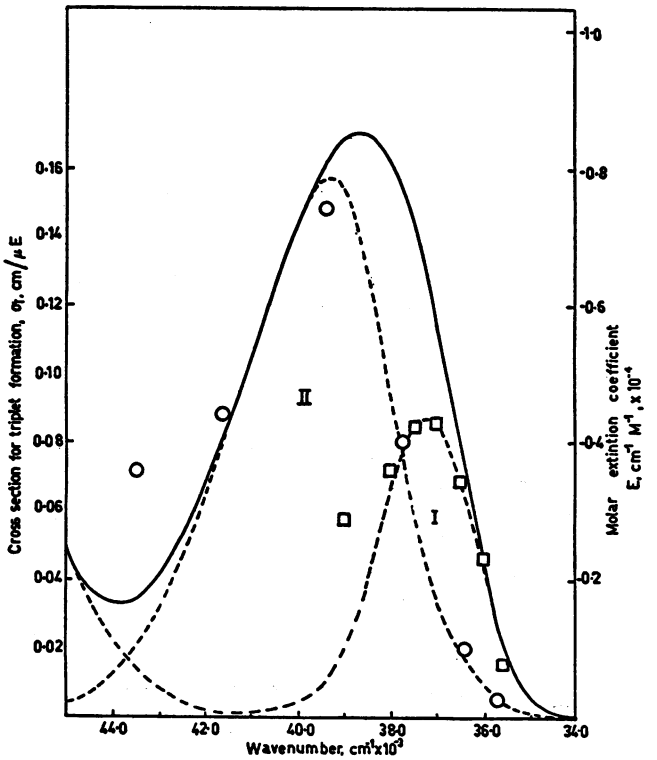


FIG. 3. Frequency distribution of triplet cross-section and the relative probability of fluorescence emission (relative excitation spectrum) fitted to the UV-absorption spectrum of uracil. (——), absorption spectrum, measured on Cary model 1501; O, σ_T (taken from ref. 2); □, relative fluorescence excitation spectrum taken from ref. 5*a*, points omitted for clarity); (– – –), resolved absorption spectra for proposed tautomers I and II (see text).

of the singlet $n\pi^*$ and $\pi\pi^*$ levels can be understood in terms of a symmetry restriction on the transition $^1(n\pi^*) \leftrightarrow {}^1(\pi\pi^*)$, which has to compete with vibrational deactivation rates to S_1 about 10^{12} sec^{-1} and singlet deactivation rates to S_0 about 10^{12} sec^{-1} (5). However, the major drawback to this model is the evidence from polarized reflectance spectroscopy† that a 1-methyluracil crystal has only one electronic transition (of $\pi\pi^*$ nature) in the region of the fluorescence and triplet cross-section results. A weak $n\pi^*$ transition is seen beginning at 42,000 cm^{-1}, but this will probably be blue-shifted in aqueous solution. In this situation, *Model 2* is not tenable. Thus, the possibilities of a single molecular system seem to be exhausted by *Models 1* and *2*, and it is necessary to turn to consideration of a two-molecule system.

Model 3 is based on tautomerism of a uracil molecule in aqueous solution. Quite simply, spectrum I of Fig. 3 is associated with one tautomer, which fluoresces weakly and has a triplet yield presently unmeasurable, whereas tautomer II,

† Unpublished results of Dr. Leigh B. Clark, University of California, San Diego; these results, which confirm and extend the polarized absorption work of Stewart, R. F. & Davidson, N. (1963) *J. Chem. Phys.* **39**, 255, cast considerable uncertainty on the interpretation by Eaton, W. A. & Lewis, T. P. (1970) *J. Chem. Phys.* **53**, 2164, of a perpendicularly polarized absorption component as being due to an $n\pi^*$ state underlying the first absorption band.

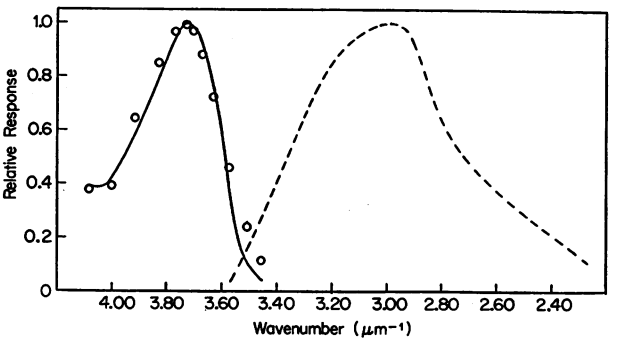


FIG. 4. Fluorescence characteristics of 2,4-diethyoxythymine (50 μM) aqueous solution at pH 6.8 determined on a Turner model 210 corrected spectrofluorimeter by a signal accumulation technique at 300°K. (– – –), relative fluorescence emission spectrum, excited at $\bar{\nu}$ Abs. (max.); O, relative fluorescence excitation spectrum, monitored at $\bar{\nu}_f$ (max.); (——), relative absorption spectrum.

associated with spectrum II, is responsible for the measured triplet yield. On this basis, it is possible to show that the reported (2) triplet cross-section and the relative fluorescence excitation spectrum (5) can be fitted to the absorption spectrum within experimental uncertainty. Spectra I and II were constructed as follows: The relative fluorescence excitation spectrum has a Gaussian form up to 38,000 cm^{-1}, and this form has been fitted to the total absorption curve subject to the constraint that it does not exceed the numerical value of the total molar absorption coefficient at any frequency. This is then considered the absorption due to tautomer I, and the absorption attributable to tautomer II is obtained as the difference between the overall absorption and that due to tautomer I. Thus, spectrum II is entirely determined by two experimental "shape" quantities and a proportionality factor. It is apparent from Fig. 3 that the absorption spectrum II obtained in this way, from the fluorescence results, closely parallels the distribution of σ_T values. The validity of this approach is strengthened by the results of Table 1, in which quantum yields of triplet formation have been recalculated with σ_A values derived from spectrum II. It is striking that the

TABLE 1. *Quantum yield of triplet formation for tautomer II*

Wavelength, nm	Effective* molar extinction coefficient ϵ(II) (× 10^{-3})	Absorption cross-section† σ_A(II)	Triplet formation cross-section‡ σ_T	ϕ_T§ (× 10^2)	
230	1.1	2.53	0.071	2.7	
240	4.1	9.45	0.088	0.93	
254	7.8	18.0	0.148	0.82	
265	4.0	9.2	0.080	0.89	0.89
275	0.9	2.07	0.020	0.96	
280	0.25	0.58	0.0049	0.85	

* "Effective" because of the unknown magnitude of the tautomeric concentration, which cannot be evaluated to give the molar extinction coefficient of spectrum II

† $\sigma_A(II) = 2.303 \times 10^{-3} \epsilon(II)$ (cm^2/mol).

‡ From ref. 2.

§ $\phi_T = \sigma_T/\sigma_A(II)$.

quantum yields are essentially constant and independent of wavelength, except for the value for 230 nm, which can be attributed to more effective intersystem crossing from the higher singlet state S_2. A corollary to the present interpretation is that the fluorescence quantum yield is (by definition) independent of exciting wavelength, with a value of 2.5×10^{-4}, and thus two anomalies are removed.

Clearly, this procedure of curve-fitting maximizes the contribution of I to the total absorption. However, lesser values would change the shape of the resultant curve II on the low-energy side, so that ϕ_T would no longer be independent of exciting wavelength.

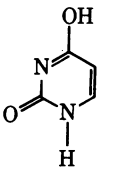
Supporting evidence that the fluorescence originates from a tautomer, and that this is the origin of the difference between the absorption and the excitation spectra, comes from the fluorescence results for 2,4-diethoxythymine, shown in Fig. 4. No tautomerism is possible for this molecule, and it is the only pyrimidine for which we have found the excitation spectrum and absorption spectrum to coincide‡.

The structure of the tautomers I and II may be inferred from fluorescence and UV-absorption data. Earlier studies (10) of various methyl- and ethoxy-substituted derivatives of uracil and thymine showed that only the ethoxy derivatives had measurable fluorescence at room temperature in neutral solution. Our own studies (Hauswirth W. & Daniels, M., in preparation) of thymine fluorescence as a function of pH show that the quantum yield increases with each stage of ionization i.e., as the molecule progressively enolizes.

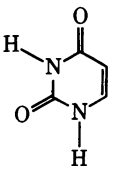
$$\phi_T(T) = 1 \times 10^{-4}; \phi_f(T^-) = 1.7 \times 10^{-3}; \phi_f(T^{2-}) = 3.5 \times 10^{-3},$$

while for 2,4-diethoxythymine, the unchanged analog of the dienolic form of thymine, $\phi_f = 1.6 \times 10^{-3}$. These results suggest that the fluorescing tautomer I contains an enol group, and the agreement between the absorption λ_{max} of tautomer I (269 nm) and λ_{max} of 4-ethoxyuracil (11) (269 nm) favors the 4-hydroxy structure. Tautomer II, more weakly fluorescing, is probably the expected predominant diketo form.

Tautomer I



Tautomer II



In conclusion, it is considered that the present analysis§ of the excitation spectra for triplet state formation and for fluorescence emission provides strong evidence for the existence of uracil and thymine in tautomeric forms in aqueous solution.

The experimental work reported here was supported in part by the United States Atomic Energy Commission (Div. of Biology and Medicine) under contract AT-(45-1)-2014, and aided considerably by the facilities of the Pacific Northwest Water laboratory. I thank Dr. Leigh B. Clark for making his results available ahead of publication. The preparation of this paper would not have been possible without the support of the Consiglio Nazionale delle Ricerche of Italy, which is gratefully acknowledged. Permanent address: Chemistry Dept. and Radiation Center, Oregon State University, Corvallis, Ore. 97331

1. Tinoco, I. & Holcomb, D. N. (1968) *Annu. Rev. Phys. Chem.* **15**, 371–394.
2. Brown, I. H. & Johns, H. E. (1968) *Photochem. Photobiol.* **8**, 273–286.
3. Fisher, C. J. & Johns, H. E. (1970) *Photochem. Photobiol.* **11**, 429–444.
4. Lamola, A. A. & Eisinger, J. (1971) *Biochim. Biophys. Acta*, **240**, 313–325.
5. (a) Daniels, M. & Hauswirth, W. (1971) *Science* **171**, 675–677; (b) Hauswirth, W. & Daniels, M. (1971) *Photochem. Photobiol.* **13**, 157–163.
6. Vigny, P. (1971) *Compt. Rend. Acad. Sci. Paris*, **272**, 2247–2249. Ser. D.
7. McGlynn, S. P., Azumi, T. & Kinoshita, M. (1969) in *Molecular Spectroscopy of the Triplet State* (Prentice Hall, Englewood Cliffs, N.J.), chap. 1.
8. Hauswirth, W. & Daniels, M. (1971) *Chem. Phys. Lett.* **10**, 140–142.
9. Eisinger, J. & Shulman, R. G. (1968) *Science* **161**, 1311–1319.
10. Longworth, J. W., Rahn, R. O. & Shulman, R. G. (1966) *J. Chem. Phys.* **45**, 2930–2939.
11. Shugar, D. & Fox, J. J. (1952) *Biochim. Biophys. Acta* **9**, 199–218.

‡ The report of Lamola and Eisinger (4) that triplet yields from 1,3-dimethylorotic acid, 1,3-dimethylthymine, and 3-methyl TMP (which cannot tautomerize) are dependent on frequency is not necessarily incompatible with the present model. It is entirely feasible according to the discussion of these authors that the triplet yield of any one tautomer could still vary with frequency. The essential point of this communication is that for uracil, triplet formation and fluorescence do not originate from the same state.

§ Optimization techniques have not been applied because of the small number of values of ϕ_T available. This paper is largely worked out in terms of the uracil data. Thymine data are very similar, ϕ_f decreasing by a factor of about 2.5 from 285 to 250 nm (5) and ϕ_T increasing by a factor of about 4 from 280 to 240 nm (3), but only three values of ϕ_T are available and curve resolution has not been attempted.