*Carbohydrate Research,* 134 (1984) 23–38
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

# F.T.-I.R. AND LASER-RAMAN SPECTRA OF THYMINE AND THYMIDINE*

MOHAMMED MATHLOUTHI, ANNE-MARIE SEUVRE[†],
*Departement "Biologie Appliquée", Institut Universitaire de Technologie, Université de Dijon, B.P. 510 – Dijon cedex (France)*

AND JACK L. KOENIG
*Department of Macromolecular Science, Case Western Reserve University, Cleveland, Ohio 44106 (U.S.A.)*

(Received March 26th, 1984; accepted for publication, May 17th, 1984)

## ABSTRACT

F.t.-i.r. and laser-Raman spectra of thymine and thymidine in the solid state were recorded. Assignments were proposed for the frequencies observed. The influence of the deoxy sugar on the vibrations of the nucleoside are discussed as a function of its particular puckering. The aim of this work is to elucidate the differences between the molecules constituting the nucleic acids, in order the better to comprehend their biological functions.

## INTRODUCTION

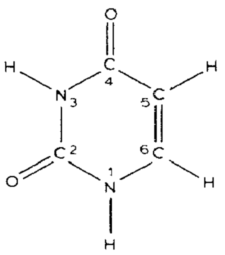
It is well known[2] that certain sequences of DNA's are richer than others in G–C or A–T pairs of bases. These characteristics confer on the sequences special functions that are undoubtedly dependent on the structure of the constitutive units (bases, nucleosides, and nucleotides). Vibrational spectroscopy is one of the physical techniques that permits elucidation of the geometry and of the modification (by the environment) of the vibrational energy of the groups of atoms constituting the molecules. In order to proceed logically in our investigation of the structure of nucleic acids, this work started[3] with interpretation of the vibrational spectra of D-ribose and 2-deoxy-D-*erythro*-pentose, and then of those of the purine base adenine and its D-ribonucleoside adenosine[1]. The pyrimidine bases present in nucleic acids are cytosine, uracil (**1**), and thymine (**2**). The last two bases differ only in the substituent on carbon atom 5.

Uracil and thymine had been studied by i.r. and Raman spectroscopy[4–7]. In these investigations, different techniques for assignment of vibrational frequencies

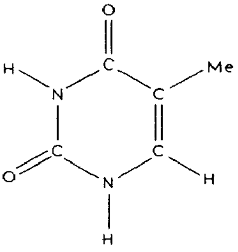
---

*F.t.-i.r. and Laser-Raman Spectra of Constituents of Nucleic Acids, Part III. For Part II, see ref. 1.
[†] Present address: Laboratoire de Chimie des Oligomères, Faculté des Sciences et Techniques de Rouen, 76130 Mont-Saint-Aignan, France.

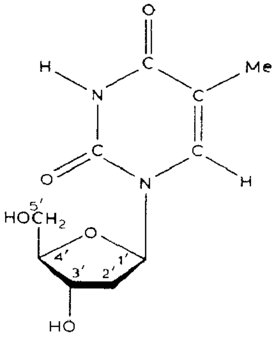
0008-6215/84/$ 03.00



**1**



**2**



**3**

were adopted. Susi and Ard[4,5] used normal-coordinate analysis and isotopic substitution in their study of uracil, thymine, and thymidine (**3**). Tsuboi *et al.*[6,7] based their interpretation of the Raman and infrared spectra of the uracil residue and of polycrystalline uracil on different methods of isotopic substitution: $D_2O$ solutions and $^{18}O$ and $^{15}N$ isotopes. Comparison of the spectra and calculation of the normal coordinates led them to propose assignments of the most prominent vibrations.

As in our earlier[1,2] investigations, discussion of the infrared and Raman results will be based on the known physical and structural properties of the molecules studied. In this context, such information as the stability and structure of metal complexes[8], ionization sites[9] and their influence on the bond lengths[10] and conformation[11] of thymine and thymidine, and optical rotation[12] of the latter will be taken into account.

## EXPERIMENTAL

The F.t.-i.r. and Raman spectra of thymine (**2**) and thymidine (**3**) were obtained by use of the techniques described previously[2]. The ranges of frequencies explored were 1700–200 and 3600–2700 $cm^{-1}$. Thymine (5-methyluracil) and thymidine [1-(2-deoxy-$\beta$-D-*erythro*-pentofuranosyl)-5-methyluracil][13] were Sigma Grade products, used without purification.

## RESULTS AND DISCUSSION

*A. Thymine* (**2**). — F.t.-i.r. and laser-Raman spectra of crystalline thymine (**2**) are respectively shown in Figs. 1 and 2. The laser-Raman spectrum of a solution of **2** in a sodium hydroxide is also given in Fig. 2. Important differences between the observed spectra are explicable by applying selection rules; this is particularly evident for the carbonyl region. The spectra will be divided into groups of frequen-

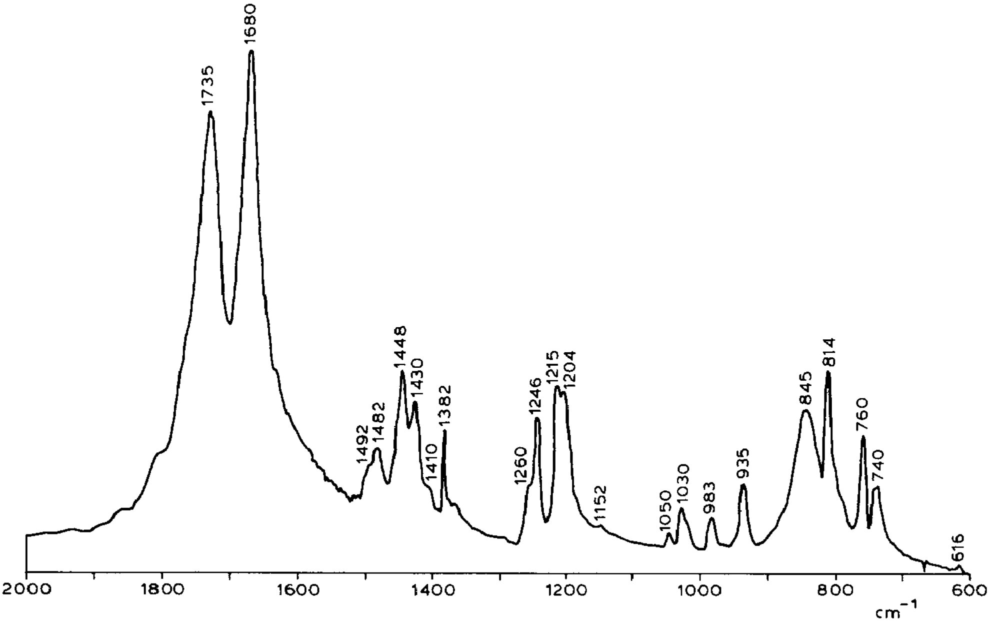


Fig. 1. F.t.-i.r. spectrum of solid thymine (**2**).

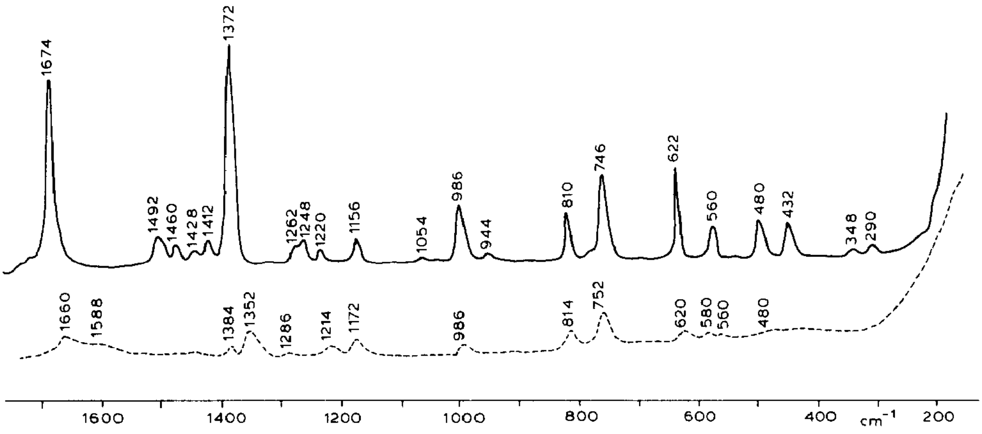


Fig. 2. Laser-Raman spectra of solid thymine (——) and in sodium hydroxide solution (----).

cies of the same kind: double bond (or carbonyl) in the 1800–1500-cm$^{-1}$ region, symmetrical vibration region of 1500–1300 cm$^{-1}$, out-of-ring vibration at 1300–900 cm$^{-1}$, and skeletal vibrations, below 900 cm$^{-1}$.

*Analysis of the bands observed. — a. The 1800–1500-cm$^{-1}$ region.* The most-intense i.r. absorptions are observed at 1735 and 1680 cm$^{-1}$. These frequencies are assignable to C=O and C=C stretching modes. However, information from the previous work[7,13a] using isotopic substitution and coordinate calculations is con-

TABLE I

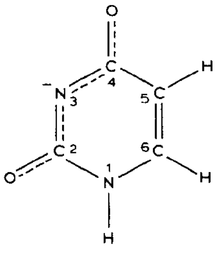
BANDS OBSERVED$^a$ IN LASER-RAMAN AND F T -I R SPECTRA OF THYMINE (**2**)

| *Solid samples* | | | | *Aqueous solutions* | |
|---|---|---|---|---|---|
| *I.r.* | | *Raman* | | | |
| *ν (cm$^{-1}$)* | *I* | *ν (cm$^{-1}$)* | *I* | *ν (cm$^{-1}$)* | *Assignments (modes)* |
| | | 290 | 5.2 | | $\delta$(OH...O) |
| | | 348 | 3.1 | | |
| | | 432 | 16.5 | | $\delta$(C=O) |
| | | 480 | 17 5 | 480 | $\delta$(N–C=C) |
| | | 560 | 15.5 | 560 | $\delta$(C–C=C) |
| | | | | 580 | |
| 616 | 2.0 | 622 | 42.3 | 620 | $\delta$(N–C–C) |
| 740 | 17.2 | | | | $\delta$(C–H) |
| 760 | 27.2 | 746 | 38.1 | 752 | $\nu$(C-5–CH$_3$) |
| 814 | 39.7 | 810 | 21.6 | 814 | $\nu$(C-4–C-5) |
| 845 | 31.9 | | | | $\delta$(N–H) out-of-plane |
| 935 | 17.4 | 944 | 3.6 | | $\delta$(C-5–C–H) |
| 983 | 11.3 | 986 | 24.7 | 986 | $\delta$(C–N–C) |
| 1030 | 13.2 | | | | $\delta$(C–N–H) |
| 1050 | 8.1 | 1054 | 2.1 | | $\delta$(N–C–H) |
| 1152 | 9.8 | 1156 | 11.3 | | r(CH$_3$) |
| | | | | 1172 | |
| 1204 | 35.5 | | | | $\nu$(C–N) |
| 1215 | 36.3 | 1220 | 6.2 | 1214 | |
| 1246 | 30.4 | 1248 | 11.3 | | $\delta$(C-5=C-6–H) |
| 1260 | 17.2 | 1262 | 6.2 | | $\nu$(C–N) |
| | | | | 1286 | |
| | | | | 1352 | |
| 1382 | 27.5 | 1372 | 100$^b$ | 1384 | $\delta$(CH$_3$) symmetrical |
| 1410 | 17.1 | 1412 | 13.4 | | |
| 1430 | 33.3 | 1428 | 6.7 | | $\delta$(CH$_3$) antisymmetrical |
| 1448 | 39.0 | 1460 | 9.3 | | |
| 1482 | 24.5 | | | | $\delta$(N-3–H) |
| 1492 | 21 1 | 1492 | 13.9 | | $\delta$(N-1–H) |
| | | | | 1588 | |
| | | | | 1660 | |
| 1680 | 100$^b$ | 1674 | 86.6 | | $\nu$(C-5=C-6) |
| 1735 | 88.2 | | | | $\nu$(C-4=O), $\nu$(C-2=O) |

$^a$Key: I = relative intensity; $\delta$ = bending mode; $\nu$ = stretching mode; r = rocking mode. $^b$Taken as reference.

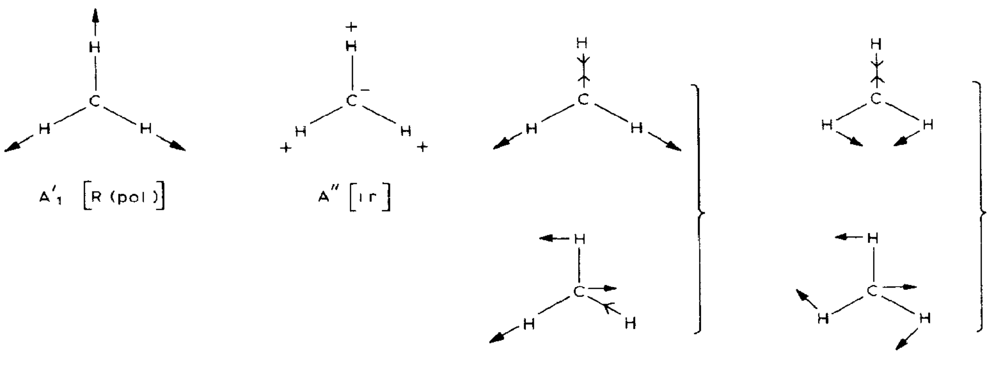
troversial as regards contribution of C-2=O to the 1735-$cm^{-1}$ vibration. An observed band at 1680 $cm^{-1}$ was assigned by different authors[4–6] to C-4=O and C-5=C-6 stretchings. Using $^{18}O$ substitution, Miles[13a] localized the C-2=O and C-4=O vibrations at 1692 and 1657 $cm^{-1}$, respectively, with a combination of C=C stretching to the latter band. Observed bands are listed in Table I. It appears that Raman lines are localized at lower frequencies than i.r. absorptions for C=O and C=C vibrations. This result could be due to intermolecular hydrogen-bonding in the crystal[4,14]. Moreover it has been noted[14] that, in alkaline solutions of uracil (and, probably, of thymine), deprotonation takes place at N-3. This deprotonation results in a delocalization of the $\pi$ electrons, and helps explain differences between i.r. and Raman intensities. Comparison of the vibrations observed for the solid form in the double-bond region in Fig. 1 (F.t.-i.r.) and Fig. 2 (laser-Raman) shows that the 1735-$cm^{-1}$ band is almost absent from the Raman spectrum.

Fig. 2 shows the Raman spectrum of a solution of **2** in aqueous 0.1M sodium hydroxide. The 1800–1500-$cm^{-1}$ region exhibits only a broad vibration at 1660 $cm^{-1}$, probably due to the coupling of $H_2O$ bending and C-5=C-6 stretching modes. According to the previous results[13a,14] and to the fact that i.r. absorptions are sensitive to dipole interactions, it is to be expected that the $\nu$(C=O) band should be more intense than $\nu$(C=C). However, the hydrogen bonding of the C=O group in the crystal[13a], as well as the delocalization of double bonds leading to a structure like that (**4**) suggested by Lord and Thomas[14] for the uracil anion may be responsible for C=O i.r. absorption of lower intensity than that of C=C and for the absence of C=O stretching in the Raman spectra. Therefore, we propose assignment of the 1735-$cm^{-1}$ i.r. band to a combination of C-2=O and C-4=O vibrations, and that the 1680-$cm^{-1}$ i.r. band and the corresponding 1674-$cm^{-1}$ Raman line have as their origin the C-5=C-6 stretching mode. When the environment of the C=O bonds is changed, as in thymidine (**3**), it may be observed (see Fig. 1) that the two carbonyl vibrations result in a double peak more intense than the C=C absorption band. The C=O vibration shown in Fig. 1 at 1735 $cm^{-1}$ is a broad, nonsymmetrical band originating from more than a single, stretching C=O mode, which supports our assignment.



**4**

*b. The 1500–1300-$cm^{-1}$ region.* This region is generally called the local-symmetry region. The most prominent vibrations observed are due to the $CH_3$



Scheme 1 Vibrations of planar $CH_3$

group, which possesses a $D_3h$ point-group symmetry[15]. The expected vibrations and their i.r. and Raman activities are shown in Scheme 1. The vibration of the methyl group consists of symmetrical, $CH_3$ bending; this is relatively strong in the Raman spectrum and is observed at 1372 $cm^{-1}$ (Fig. 2), and exhibits a medium-intensity i.r. absorption at 1382 $cm^{-1}$ (see Fig. 1). Three antisymmetrical vibrations are to be expected from group-theory predictions; these vibrations give very weak Raman lines and relatively strong i.r. absorptions, so that i.r. absorptions at 1410, 1430, and 1448 $cm^{-1}$, which correspond to the Raman lines at 1412, 1428, and 1460 $cm^{-1}$, could be assigned to the antisymmetrical $CH_3$ deformations.

Deuteration techniques and calculations allowed Susi and Ard[5] to assign the 1492-$cm^{-1}$ vibration to $\delta$(N–H), and the 1484-$cm^{-1}$ vibration to a ring-stretching mode. Comparison of the observed and calculated frequencies of di-*C*-deuterouracil and perdeuterouracil shows that the 1498-$cm^{-1}$ vibration is absent for perdeuterouracil, and remains at the same value for di-*C*-deuterouracil. This shift demonstrates the contribution of NH deformation to this frequency, and this result is in agreement with Kazakova's results[16] on uracil and its deuterated derivatives. As expected, the Raman intensity of this out-of-plane vibration of (C–N–H) is weak. The environment of N-1–H is different from that of N-3–H, which explains the presence of two coupled absorption bands, at 1482 and 1492 $cm^{-1}$ in the F.t.-i.r. spectrum (see Fig. 1) and a broad, weak band in the Raman spectrum (see Fig. 2), at 1492 $cm^{-1}$. Accordingly, we propose assignment of the 1492-$cm^{-1}$ vibration to $\delta$(N-1–H), and that at 1482 $cm^{-1}$ to $\delta$(N-3–H).

*c. The 1300–900-$cm^{-1}$ region.* The deformation of angles, including those for atoms in and out of the ring is expected in this region. This is particularly true for sugars, where this range of frequencies is called the anomeric or "fingerprint" region, in reference to C-1–OH and C-1–H deformations. The antisymmetrical character of these deformations emerges from the difference in intensities of i.r. and Raman bands. Individual stretchings from the ring are also expected in the 1300–900-$cm^{-1}$ range. Two vibrations (a band at 1246 and a shoulder at 1260 $cm^{-1}$) having medium intensities are observed in Fig. 1 and, in Fig. 2, at 1262–1248 $cm^{-1}$

with weak Raman intensities. The frequency at 1242 $cm^{-1}$ was assigned by Hartman *et al.*[17] to thymine residues of DNA, and that at 1243 $cm^{-1}$, to uracil residues in RNA. The same conclusions were drawn[18] from a Raman investigation of yeast RNA and calf-thymus DNA. An experimental and theoretical study[4,5] of uracil and thymine and their derivatives showed frequencies at 1248, 1245, 1238 and 1236 $cm^{-1}$ which were assigned to $\delta$(C–H) and ring stretching.

Deformation of the C–H and N–H bonds was localized[16] in the 1400–1000-$cm^{-1}$ region. Analysis of i.r. spectra of uracil-5-*d* and -6-*d* allowed Kazakova[16] to assign the 1240–1220-$cm^{-1}$ vibrations to CH deformations. Such results led us to attribute 1248-$cm^{-1}$ (Raman) and 1246-$cm^{-1}$ (i.r.) vibrations to $\delta$(C-5=C-6–H). The shoulder at 1260 $cm^{-1}$ (i.r.) and 1262 $cm^{-1}$ (Raman) could be a contribution of an endocyclic C–N stretching. The ring stretchings are generally found[4,5] in the 1300–900-$cm^{-1}$ region. C–N contributions to ring stretchings certainly occur at higher frequency than $\nu$(C–C). Moreover, the environment of N-3, with the proximity of $\pi$ electrons, enhances the dipole character of N-3–C bonds, which explains the fact that i.r. absorptions corresponding to $\nu$(C–N) are stronger than the Raman peak observed at 1220 $cm^{-1}$. Analysis of the assignments proposed in the literature[4,5] does not permit ascertaining whether there is any correspondence between observed frequencies and a well defined vibration in this region. However, we may base our assignments, listed in Table I, of 1215 and 1204 $cm^{-1}$, to the different C–N stretchings, on the preceding arguments and on observations of Tsuboi *et al.*[6,7] of C–N stretching at ~1200 $cm^{-1}$.

The vibration observed at 1152 $cm^{-1}$ (i.r.) and 1156 $cm^{-1}$ (Raman) was found[19] at 1154 $cm^{-1}$ in the Raman spectrum of 1-methylthymine and at 1156 $cm^{-1}$ by Nishimura *et al.*[6] in that of a methyluracil. This frequency was assigned to a carbonyl[19] or a ring stretching[5,6]. The fact that this vibration is present only in spectra of thymine or other methyluracils is used as an argument for assignment of 1156–1152 $cm^{-1}$ to a $CH_3$ rocking (see Table I). Four bands are observed in the F.t.-i.r. spectrum (see Fig. 1), at 1050, 1030, 983, and 935 $cm^{-1}$. Only three Raman lines are shown in the same region of frequencies, with a very weak intensity for those at 1054 and 944 $cm^{-1}$ and a medium intensity for that at 986 $cm^{-1}$. Comparison of Raman and i.r. intensities led to the conclusion that these vibrations originate from endocyclic and exocyclic deformations.

An infrared absorption was observed[5] at 1028 $cm^{-1}$ that is absent from the Raman spectrum of thymine; this vibration was assigned to a combination of ring stretchings and bendings. The i.r. band at 1030 $cm^{-1}$ (see Fig. 1) certainly originates from a combination of the deformations of angles of the same nature but differing in environment, which is the case for $\delta$(N-1–H) and $\delta$(N-3–H). The antisymmetrical character of such a vibration is confirmed by its vanishing in the Raman spectrum. It must be noted that, besides the vibration listed at 1030 $cm^{-1}$ (see Table I), a shoulder is observed (see Fig. 1) which could be a differentiation of the $\delta$(NH) vibrations. The weak band at 1050 $cm^{-1}$ (see Fig. 1) was not observed in previous work[4–7] on thymine and uracil. Its origin, as indicated earlier, is probably a peripheral deformation. We propose to assign it to δ(N–C–H). Its presence in the Raman spectrum (see Fig. 2) and the absence of a vibration at 1030 $cm^{-1}$ (corresponding to δ(C–N–H)) is explicable by the dipole character of NH. Another exocyclic, CH deformation, δ(C-5–C–H), is observed at 935 $cm^{-1}$ (i.r.) and 944 $cm^{-1}$ (Raman). This angle has only one atom belonging to the ring; it needs less energy of vibration than a deformation such as δ(N–C–H), which supports our assignment (see Table I). No frequencies were observed[4,5,16] at 935 $cm^{-1}$ for thymine and uracil. Vibrations at 983 $cm^{-1}$ (see Fig. 1) and 986 $cm^{-1}$ (see Fig. 2) were found[4] in the i.r. (993 $cm^{-1}$) and Raman (988 $cm^{-1}$) spectra of uracil, and assigned to ν and δ(ring).

Investigation[4] of i.r. and Raman spectra of thymine permitted observation of frequencies at 984 and 985 $cm^{-1}$ which were attributed to ($CH_3$). Our observed results (see Figs. 1 and 2) show that the i.r. band at 983 $cm^{-1}$ and the Raman line at 986 $cm^{-1}$ have approximately the same intensity. We propose assignment of this vibration to an endocyclic, δ(C–N–C) deformation. A similar vibration was observed[1] at 1025 $cm^{-1}$, and assigned to δ(C–N–C). The relatively symmetrical aspect of δ(C-2–N-3–C-4) supports observation of a relatively intense, Raman line at 986 $cm^{-1}$. However, this vibration is wide, and it could be partly due to δ(N-1–C-2–N-3), which is symmetrical compared to other ring-vibrations.

*d. Region of frequencies below 900 $cm^{-1}$.* As a general rule, ring vibrations and out-of-plane deformations of exocyclic CH and NH are found[20] between 900 and 700 $cm^{-1}$. The frequencies observed below 700 $cm^{-1}$ are inclusively attributed to skeletal modes. Two relatively strong absorptions are seen in Fig. 1 at 845 and 814 $cm^{-1}$. Only one vibration at 810 $cm^{-1}$ is observed in the Raman spectrum (see Fig. 2). The vibration at 845 $cm^{-1}$ is broad, and probably originates from an out-of-plane NH deformation. Its absence from the Raman spectrum is explicable according to the same arguments as those for the 1050–1030-$cm^{-1}$ region. A broad and diffuse i.r. band was observed by Kazakova[16] between 870 and 820 $cm^{-1}$ that disappears from the spectrum of the di-*N*-deutero derivative of uracil; this band has been regarded[16] as an out-of-plane NH deformation, and the same result was derived[4] from an isotopic-substitution study of uracil. Consequently the 845-$cm^{-1}$ vibration seems to be highly characteristic of the NH deformation. An individualized, i.r. absorption at 814 $cm^{-1}$, corresponding to a Raman line at 810 $cm^{-1}$, was assigned to endocyclic ν(C–C) (see Table I); this assignment is in agreement with previous work[1,3,21] on carbohydrates. Moreover, this vibration is present in the Raman spectrum of sodium hydroxide solutions of thymine, which indicates that no NH group is involved in this vibration. The thymine molecule (**2**) contains an exocyclic C–C bond that would be expected to stretch at a lower frequency than the endocyclic, C-4–C-5 bond, maintained rigid in the ring. The i.r. band observed at 760 $cm^{-1}$ is, accordingly, assigned to ν(C–$CH_3$).

Likewise, a Raman vibration observed at 746 $cm^{-1}$ could have as its origin an important contribution of ν(C–$CH_3$). This Raman band has a comparatively important width and comprises a shoulder, which comes from a combination of a stretching and a deformation. It was reported[17,18] in Raman investigations of DNA that a line for the thymine residue appears at 725 cm$^{-1}$; this vibration, which seems to be characteristic of thymine, is absent for uracil derivatives, supporting our assignment to a C–C stretching involving the CH$_3$ group. The observed i.r. absorption at 740 cm$^{-1}$ probably originates from an out-of-ring deformation of CH, as is generally considered[22] for rings having one hydrogen atom adjacent to it. A strong Raman line is observed at 622 cm$^{-1}$ (see Fig. 2). Strobel and Scovell[19] assigned the line at 623 cm$^{-1}$ to $\delta$(N-7–C-5–C-6) in 9-methyladenine.

From our previous work[1] on adenine, it may be seen that a Raman band at 628 cm$^{-1}$ was attributed to $\delta$(N–C–C); this vibration was also found at 620 cm$^{-1}$ in the Raman spectrum of the sodium hydroxide solution of thymine (see Fig. 2), which means that no influence of alkaline pH was detected. That is why we are inclined to assign the line at 622 cm$^{-1}$ to $\delta$(N-3–C-4–C-5). Three Raman bands having the same intensity are seen (see Fig. 2) at 560, 480, and 432 cm$^{-1}$. We propose assignment of the line at 560 cm$^{-1}$ to $\delta$(C–C=C), that at 480 cm$^{-1}$ to $\delta$(N–C=C), and that at 432 cm$^{-1}$ to out-of-phase C=O deformations. The C=O bending modes have been discussed in regard to the use of normal-coordinate calculations[6] and the Raman[17,19] effect. The calculation results permitted assignment of the line at 421 cm$^{-1}$ to $\delta$(ring) and $\delta$(C=O) in uracil; the Raman investigation of 1-methylthymine[19] led to the conclusion that the line at 429 cm$^{-1}$ comes from carbonyl deformation. A frequency at 435 cm$^{-1}$ observed in the Raman

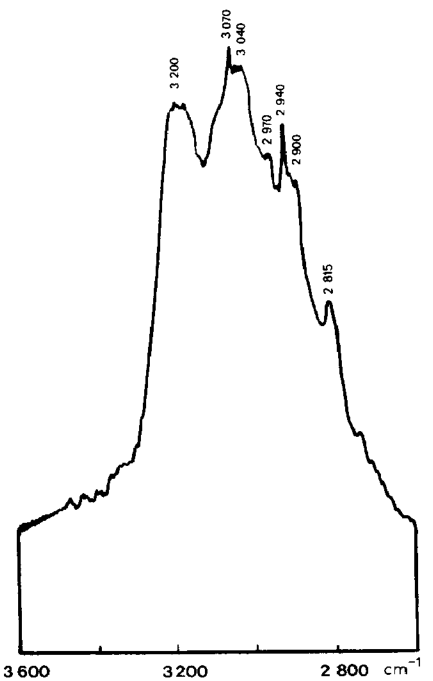


Fig. 3. F.t.-i.r. spectrum of solid thymine (2) in the region of frequencies of 3600–2700 cm$^{-1}$.

TABLE II

BANDS OBSERVED[a] IN THE F T -I R SPECTRUM OF THYMINE (THE 3600–2000-$cm^{-1}$ REGION)

| $\nu$ ($cm^{-1}$) | *I* | *Assignments* |
|---|---|---|
| 2815 | 57.4 | $\nu$(C-5–H) |
| 2900 | 77.5 | |
| 2940 | 86.8 | $\nu_s(CH_3)$ |
| 2970 | 81.4 | |
| 3040 | 96.6 | $\nu_a(CH_3)$ |
| 3070 | 100[b] | |
| 3200 | 90.7 | $\nu$(NH) |

[a]Key: I = relative intensity; $\delta$ = bending mode; $\nu$ = stretching mode; r = rocking mode. [b]Taken as reference.

spectrum of DNA was assigned to $\delta$(ring) and $\delta$(C=O). The assignments proposed in Table I for this region of frequencies are in agreement with those made by Sanyal *et al.*[23]. Two very weak bands are observed (see Fig. 2) at 348 and 290 $cm^{-1}$. Assignments in this region of frequencies are scarce. As proposed earlier[1,2], hydrogen bonding is expected to give rise to Raman vibrations at 348 and 290 $cm^{-1}$.

*e. Frequencies in the region between 3600 and 2700 $cm^{-1}$.* The F.t.-i.r. spectrum of solid thymine (**2**) is shown in Fig. 3. The frequencies observed and their assignments are listed in Table II. The proposed attribution of the band at 3200 $cm^{-1}$ to $\nu$(N–H) is in agreement with a deuteration study[16] of uracil, and with our previous results[1]. The symmetrical and antisymmetrical C–H stretching in $CH_3$ are

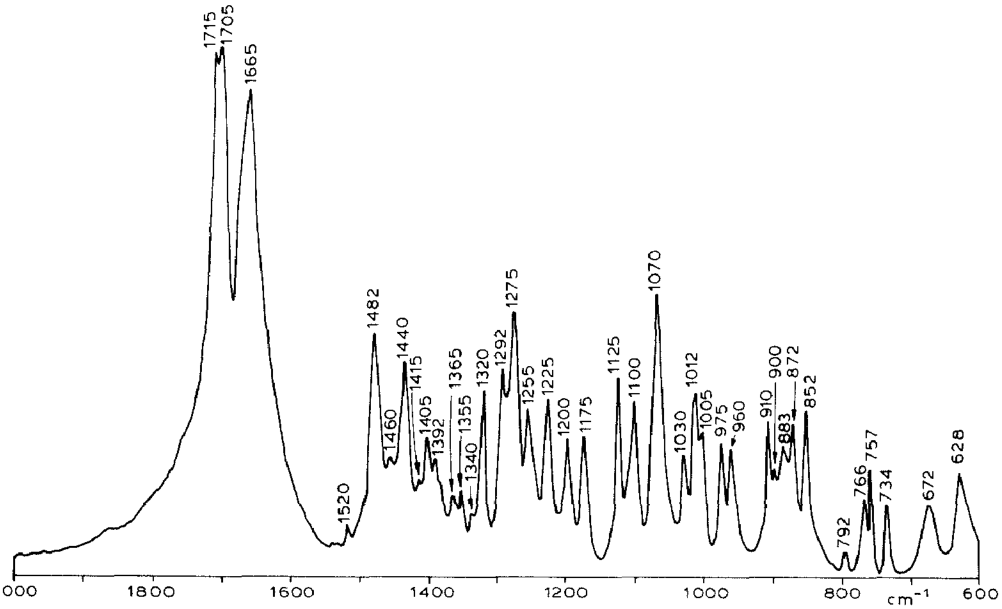


Fig. 4. F.t.-i.r. spectrum of solid thymidine (**3**).

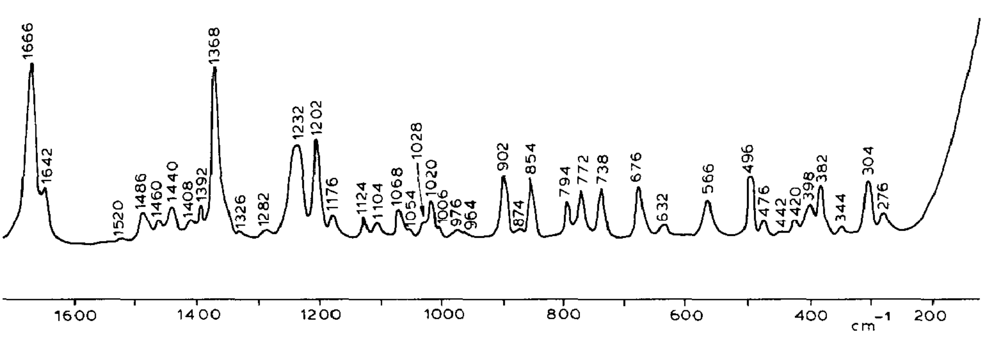


Fig. 5. Laser-Raman spectrum of solid thymidine (**3**).

assigned according to classical results[15]. The environment of the methyl group in thymine is probably the origin of a shift of $\nu_s$ $(CH_3)$ and $\nu_a$ $(CH_3)$ towards higher frequencies. The vibration at 2815 $cm^{-1}$, having a lower intensity than $\nu$(C–H) in $CH_3$, is assigned to $\nu$(C-5–H). Comparison of our results to those in the literature[4,5,16] showed that the F.t.-i.r. technique gives rise to numerous, well defined bands not observed in classical, i.r. spectroscopy.

*B. Thymidine* (**3**). — F.t.-i.r. and Raman spectra of solid thymidine (**3**) are shown in Figs. 4 and 5, respectively. The general aspect of the vibrations observed for the nucleoside is completely different from that of the pyrimidine base. From comparison of the spectra of the ribonucleoside adenosine and D-ribose, it was found[1] that addition of the D-ribosyl group to the adenine molecule influences the vibrations, especially in the region of 1200–800 $cm^{-1}$. However, it emerges from Figs. 1 and 4 that, apart from the carbonyl region, the whole i.r. spectrum is perturbed. This is probably due to the structure of the deoxy sugar, which is capable of inducing frequency shifts, and variations of relative intensities, when the environment is changed[3].

*a. Observed bands identified from the spectra of thymine and 2-deoxy-*D-erythro*-pentose.* Two very strong bands are observed in Fig. 4, at 1715–1705 and 1665 $cm^{-1}$. They are comparable to the two peaks observed in Fig. 1 in the carbonyl region. However, shifts in frequencies, and modification of the relative intensities, are seen, especially for the 1735-$cm^{-1}$ frequency, assigned to a combination of $\nu$(C-2=O) and $\nu$(C-4=O) (see Table I), which is shifted towards lower frequencies and split into two vibrations, at 1715 and 1705 $cm^{-1}$. Analysis of the hydrogen bonding of the thymidine crystal[24] shows that N-3–H acts as a donor of an H bond and O-4 as an acceptor thereof, whereas no hydrogen bonding is noted for C-2=O-2. This may be the origin of the differentiation of the two carbonyls. The band at 1665 $cm^{-1}$ (see Fig. 4) and 1666 $cm^{-1}$ (see Fig. 5) corresponds to the i.r. absorption at 1680 $cm^{-1}$ (see Fig. 1) and the Raman line at 1674 $cm^{-1}$ (see Fig. 2); it is assigned to $\nu$(C-5=C-6). The frequencies between 1500 and 1300 $cm^{-1}$ concern the $CH_3$ and $CH_2$ deformations. Assignments of the observed frequencies to the sugar or the base residue were derived from previous results (see ref. 3 and Table

TABLE III

BANDS OBSERVED[a] IN THE F T -I R AND LASER-RAMAN SPECTRA OF THYMIDINE (3)

| *I.r.* | | *Raman* | | |
|---|---|---|---|---|
| *ν (cm$^{-1}$)* | *I* | *ν (cm$^{-1}$)* | *I* | *Assignments (modes)* |
| | | 276 | 10 2 | D |
| | | 304 | 28.0 | D |
| | | 344 | 5.1 | T |
| | | 382 | 28.0 | D |
| | | 398 | 24.2 | D |
| | | 420 | 8.9 | D |
| | | 442 | 3.2 | D |
| | | 476 | 8.9 | D |
| | | 496 | 31.8 | δ(C–C–C) |
| | | 566 | 20.4 | D + T |
| 628 | 20.1 | 632 | 6.4 | D |
| 672 | 14 0 | 676 | 28.0 | D |
| 734 | 13.2 | 738 | 28.0 | D + T |
| 757 | 20.3 | | | D + T |
| 766 | 15.2 | 772 | 25.5 | D |
| 792 | 4.7 | 794 | 20.4 | δ(N-1–C-1′–C-2′) |
| 852 | 31.9 | 854 | 33.1 | T |
| 872 | 29.4 | 874 | 3.8 | δ(C-2′–C-1′–O-4′) |
| 883 | 25.0 | | | D |
| 900 | 20.3 | 902 | 35.0 | D |
| 910 | 29.7 | | | δ(N-1–C-1′–O-4′) |
| 960 | 24.5 | 964 | 3.2 | δ(C-1′–N-1–C) |
| 975 | 25.7 | 976 | 3.8 | D + T |
| 1005 | 27.2 | 1006 | 6.4 | δ(O-4′–C-1′–H) |
| 1012 | 34.8 | | | D |
| | | 1020 | 21.7 | D |
| 1030 | 23.0 | 1028 | 8.9 | T |
| | | 1054 | 6.4 | D + T |
| 1070 | 53.4 | 1068 | 16.6 | δ(N-1–C-1′–H) |
| 1100 | 33.3 | 1104 | 8.9 | D |
| 1125 | 37 7 | 1124 | 12.7 | D |
| 1175 | 27.0 | 1176 | 12.7 | D |
| 1200 | 26.5 | 1202 | 53.5 | D + T |
| 1225 | 33.6 | 1232 | 51.0 | ν(C-1′–N-1) |
| 1255 | 31.9 | | | D + T |
| 1275 | 50.5 | 1282 | 5.1 | D |
| 1292 | 39.2 | | | |
| 1320 | 35.3 | 1326 | 5.1 | D |
| 1340 | 11.8 | | | D |
| 1355 | 16.2 | | | D |
| 1365 | 15.4 | 1368 | 93.0 | T |
| 1392 | 22.5 | 1392 | 17.8 | D |
| 1405 | 26.0 | 1408 | 10.2 | T |
| 1415 | 18.1 | | | D + T |
| 1440 | 40.4 | 1440 | 17.8 | D |
| 1460 | 22.5 | 1460 | 11.5 | T |
| 1482 | 46.1 | 1486 | 15.3 | T |
| 1520 | 9.3 | 1520 | 1.9 | |
| | | 1642 | 25.0 | |

TABLE III (*continued*)

| *I.r.* | | *Raman* | | |
|---|---|---|---|---|
| *ν (cm⁻¹)* | *I* | *ν (cm⁻¹)* | *I* | *Assignments (modes)* |
| 1665 | 91.7 | 1666 | 100[b] | T |
| 1705 | 100[b] | | | T |
| 1715 | 98.5 | | | T |

[a]Key: I = relative intensity; $\delta$ = bending mode; $\nu$ = stretching mode. T = thymine (detailed assignments are given in Table I); D = 2-deoxy-D-*erythro*-pentose (detailed assignments are given in ref. 2). [b]Taken as reference.

I), and are listed in Table III. It must be pointed out that, except for a few peaks, such as the strong one at 1368 cm$^{-1}$ assigned to symmetrical $\nu$(CH$_3$) (see Fig. 5), most of the vibrations for thymine are masked or shifted, whereas those of 2-deoxy-D-*erythro*-pentose retain the same relative intensities and exhibit only very slight frequency shifts.

*b. Observed bands differentiating thymidine* (**3**) *from thymine* (**2**) *and 2-deoxy-*D-erythro-*pentose.* The vibrations not referred to as "D" (for the deoxyglycosyl group) or "T" (for the thymine residue) in Table III are specific to the nucleoside and will be discussed hereafter. As expected, these vibrations come from the stretching and deformations around the glycosylic bond, C-1′–N-1. A new vibration is shown in the i.r. spectrum (see Fig. 4) at 1225 cm$^{-1}$, and in the Raman spectrum (see Fig. 5) at 1232 cm$^{-1}$. It may be assigned to $\nu$(C-1′–N-1), as is generally indicated[20,22] in correlation charts. The crystallographic data[24] give the C–N bond-lengths in thymidine as N-1–C-1′, 148.0, N-1–C-2, 138.5, N-3–C-2, 138.1, N-3–C-4, 137.8, and N-1–C-6, 137.4 pm.

From these results, it is to be expected that C-1′–N-1, which is the longest bond, should stretch at the lowest frequency. The fact that $\nu$(C-1′–N-1) is localized at a frequency relatively higher than those of the other C–N stretchings is certainly due to the nature of the vibration, which acts more like a linkage between the two residues than as a bond between the two atoms. The i.r. bands at 1070 and 1005 cm$^{-1}$ (see Fig. 4) are stronger than the corresponding Raman vibrations at 1068 and 1006 cm$^{-1}$ (see Fig. 5). This difference in intensities is attributable to the non-symmetrical character of such vibrations. As may be seen in Table III, these frequencies are respectively assigned to $\delta$(N-1–C-1′–H) and $\delta$(O-4′–C-1′–H).

The i.r. band observed at 983 cm$^{-1}$ for thymine (see Fig. 1) was assigned to $\delta$(C–N-C). Two i.r. absorptions are shown, at 975 and 960 cm$^{-1}$, by thymidine (see Fig. 4). The vibration at 975 cm$^{-1}$ corresponds to the thymine vibration at 983 cm$^{-1}$, and that at 960 cm$^{-1}$ could originate from the deformation of the C-1′–N-1–C) angle. The modification of environment due to the glycosylic linkage affects the deformations around C-1′, even for the sugar angles. The observed new frequencies concern the endo- and exo-cyclic deformations around C-1′. Assignments are proposed for 910 cm$^{-1}$ (i.r.) to $\delta$(N-1–C-1′–O-4′), 872 cm$^{-1}$ (i.r.) and 874

TABLE IV

BANDS OBSERVED[a] IN THE F T -I R SPECTRUM OF THYMIDINE (THE 3600–2000-cm$^{-1}$ REGION)

| *$\nu$ (cm$^{-1}$)* | *I* | *Assignments* |
|---|---|---|
| 2845 | 45.3 | $\nu$(C–H) |
| 2900 | 29.4 | Fermi resonance |
| 2940 | 31.9 | $\nu$(CH$_3$) |
| 2980 | 52.0 | $\nu$(C-5′–H$_2$) |
| 3000 | 50.5 | $\nu$(C-2′–H$_2$) |
| 3025 | 53.9 | $\nu_a$(CH$_3$) |
| 3095 | 36.5 | |
| 3110 | 35.8 | |
| 3160 | 43.1 | (N–H) |
| 3320 | 100[b] | (O–H) |

[a]Key: I = relative intensity; $\nu$ = stretching mode ($\nu_a$, antisymmetrical stretching and $\nu_s$, symmetrical stretching). [b]Taken as reference

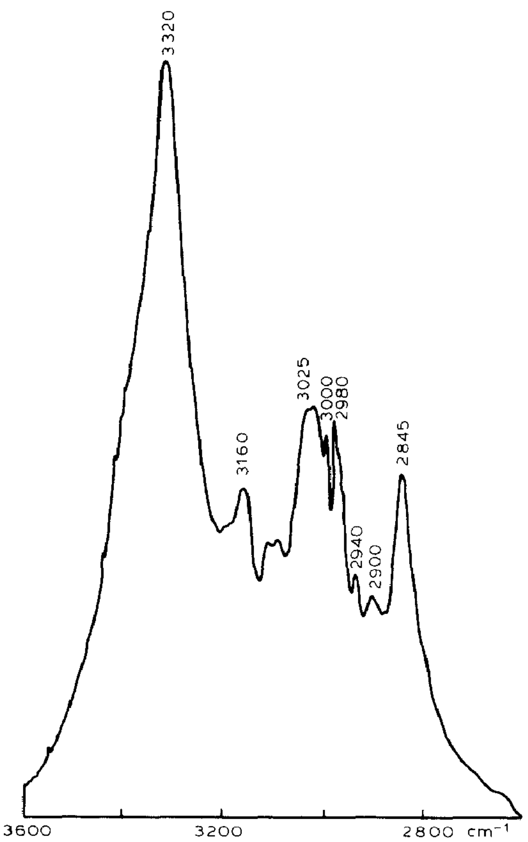


Fig. 6. F.t.-i.r. spectrum of solid thymidine (3) in the region of frequencies of 3600–2700 cm$^{-1}$.

$cm^{-1}$ (Raman) to $\delta$(C-2′–C-1′–O-4′), and 792 $cm^{-1}$ (i.r.) and 794 $cm^{-1}$ (Raman) to $\delta$(N-1–C-1′–C-2′). It may be noted that these frequencies have values higher than those generally accepted for the skeletal-deformation modes.

It may be concluded from these assignments that the sugar moiety is more perturbed than the thymine residue, which is in agreement with the structure derived from X-ray diffraction[24]. Indeed, it was found that the pyrimidine ring in thymidine is planar, and that the C-1′ atom is coplanar with that ring. The rest of the furanoid ring is puckered, with C-3′ *exo,* which is an exception to the commonest conformations, having C-2′ or C-3′ *endo*. Such a puckering results in a modification of the C–C–C angles. The Raman line at 496 $cm^{-1}$ (see Fig. 5) could correspond to displacement of the vibration at 508 $cm^{-1}$ observed in the Raman spectrum[2] of 2-deoxy-D-*erythro*-pentose, assigned to $\delta$(C–C–C) in the furanose ring. This shift towards lower frequencies may be explained by deformation of the C–C–C angles.

*c. The region of frequencies between 3600 and 2700 $cm^{-1}$*. Only the F.t.-i.r. spectrum of crystalline thymidine was recorded (see Fig. 6 and Table IV). Comparison of this spectrum with that of thymine (see Fig. 3) shows an important absorption, at 3320 $cm^{-1}$, that originates from O–H stretchings; the hydrogen bonding and the packing arrangement in the crystal explain the intensity and the position of this frequency. The vibration at 3160 $cm^{-1}$ is assigned to $\nu$(N–H). The aspect of this band is different from the wide absorption observed at 3200 $cm^{-1}$ in Fig. 3 for thymine, because of the absence of N-1–H in the nucleoside. The remaining peaks in Fig. 6 come from C–H vibrations. The group of bands at ~3000 $cm^{-1}$ has as its origin $CH_3$ or $CH_2$ stretchings, whereas the lower-frequency band at 2845 $cm^{-1}$ is assigned to other C–H vibrations. As in thymine, asymmetrical $CH_3$ stretching may be localized at 3025 $cm^{-1}$, and the $\nu_s(CH_3)$, at 2940 $cm^{-1}$. The endocyclic $CH_2$ in the sugar, which is relatively more restricted, should stretch at 3000 $cm^{-1}$, and C-5′–$H_2$ of the 4-(hydroxymethyl) group, at 2980 $cm^{-1}$. The vibration at 2900 $cm^{-1}$ could be due to a Fermi resonance between $\nu(CH_2)$ and the first overtone of $\delta(CH_2)$.

## CONCLUSION

Analysis of the vibrational spectra of thymine and thymidine shows a resemblance between those of the nucleoside and the deoxy sugar in the whole range of frequencies. The influence of the 2-deoxy-D-*erythro*-pentofuranosyl group on the vibrations of thymidine are probably due to its particular conformation with C-3′ *exo*. The region of frequencies that is common to thymine and thymidine, but less perturbed by the sugar vibration, is the carbonyl region. Besides the bands arising from the sugar and the pyrimidine base, the spectrum of the nucleoside shows specific vibrations due to the glycosylic linkage.

The aim in proposing assignment of frequencies is the characterization of the constitutive elements of nucleic acids. Such an investigation should be helpful in precisely determining the sites of fixation of metals. Indeed, when cellular division takes place, the DNA molecule is opened at the level of bases fixing metallic ions[3].

## REFERENCES

1 M. Mathlouthi, A.-M. Seuvre, and J. L. Koenig, *Carbohydr. Res* , 131 (1984) 1–15.

2 E. Guille, J. Grisvard, and I. Sissoeff, in D. R. Williams (Ed.), *Biocompatibility,* Vol. 5, C.R.C. Press, Boca Raton, Florida, 1980, pp. 39–45.

3 M. Mathlouthi, A.-M. Seuvre, and J. L. Koenig, *Carbohydr. Res.*, 122 (1983) 31–47.

4 H. Susi and J. S. Ard, *Spectrochim. Acta,* 27A (1971) 1549–1562.

5 H. Susi and J. S. Ard, *Spectrochim. Acta,* 30A (1974) 1843–1853.

6 Y. Nishimura, H. Haruyama, K. Nomura, A. Y. Hirakawa, and M. Tsuboi, *Bull. Chem. Soc. Jpn.*, 52 (1979) 1340–1345.

7 M. Tsuboi, S. Takahashi, and I. Harada, *Phys. Chem. Prop. Nucleic Acids,* 2 (1973) 91–145.

8 T. Theophanides, *NATO Adv. Study Inst. Ser., Ser. C,* 43 (1978) 205–223.

9 R. M. Izatt, J. J. Christensen, and J. H. Rytting, *Chem. Rev.*, 71 (1971) 439–481.

10 R. Taylor and O. Kennard, *J. Mol. Struct.*, 78 (1982) 1–28.

11 R. U. Lemieux, *Can. J. Chem.*, 39 (1961) 116–129.

12 R. U. Lemieux and M. Hoffer, *Can. J. Chem.*, 39 (1961) 110–115.

13 P. A. Levene and R. S. Tipson, *Z. Physiol. Chem.*, 234 (1935) v; *Science,* 81 (1935) 98; *J. Biol. Chem.*, 109 (1935) 623–630.

13a H. T. Miles, *Proc. Natl. Acad. Sci. U.S.A.*, 51 (1964) 1104–1109.

14 R. C. Lord and G. J. Thomas, Jr., *Spectrochim. Acta,* 23A (1967) 2551–2591.

15 E. F. H. Brittain, W. O. George, and C. H. J. Wells, *Introduction to Molecular Spectroscopy: Theory and Experiment,* Academic Press, New York, 1970, pp. 114–213.

16 G. V. Kazakova, *Zh. Strukt. Khim.*, 10 (1969) 618–623.

17 K A. Hartman, R. C. Lord, and G J. Thomas, Jr., *Phys. Chem. Prop Nucleic Acids,* 2 (1973) 1–89.

18 W. L. Peticolas, in G. L. Cantoni and D. Davis (Eds.), *Procedures in Nucleic Acid Research,* Vol. 2, Harper and Row, New York, 1971, pp. 94–136.

19 J. L. Strobel and W. M. Scovell, *Biochim. Biophys. Acta,* 608 (1980) 201–214.

20 L. J. Bellamy, *Advances in Infrared Group Frequencies,* Methuen, London, 1968.

21 V. V. Sivchik and R. G. Zhbankov, *Zh. Prikl. Spektrosk.*, 28 (1978) 1038–1045.

22 N. B. Colthup, L. H. Daly, and S. E. Wilberly, *Introduction to Infrared and Raman Spectroscopy,* Academic Press, New York, 1974.

23 N. K. Sanyal, S. L. Srivastava, and R. K. Goel, *Indian J. Phys.*, 52 (1977) 108–115.

24 D. W. Young, P. Tollin, and H. R. Wilson, *Acta Crystallogr.*, 25 (1969) 1423–1432.