

## Planar valence force constants and assignments for pyrimidine derivatives

H. SUSI and J. S. ARD  
Eastern Regional Research Center\*  
Philadelphia, Pennsylvania 19118 U.S.A.

(Received 27 November 1973)

**Abstract**—Overlay calculations have been carried out to obtain a set of transferable valence force constants for the planar modes of pyrimidine derivatives. A 42 parameter force field was adjusted to reproduce 233 experimental frequencies of 10 molecules. Computations were based on previous work on isotopic analogs of uracil and cytosine, and new data on thymine, 1-methylthymine, 1-methyluracil, and N-deuterated analogs. The average frequency error was below one percent. Assignments and characteristic frequencies are discussed on the basis of the potential energy distribution.

### INTRODUCTION

THE VIBRATIONAL spectra of biologically important pyrimidine derivatives have been examined by several investigators [1–3], but the high complexity and low symmetry of these molecules has made detailed interpretation very difficult. Over the past few years normal coordinate calculations have been carried out for the planar modes of uracil and cytosine, and assignments have been reported for a total of six isotopic analogs [4, 5]. The present communication describes the spectra of thymine, 1-methylthymine, and 1-methyluracil, and reports the results of Overlay calculations involving a total of ten isotopic species of five pyrimidine derivatives. The 1-methyl derivatives are of importance because their spectra resemble the spectra of related biopolymers [1–3].

The following molecules were involved in Overlay calculations: uracil; C,C-dideutero uracil; perdeutero uracil; cytosine; cytosine- $d_3$ ; thymine; thymine- $d_2$ ; 1-methyluracil; 1-methyluracil- $d$ ; 1-methylthymine. The first five are previously investigated species [4, 5]. The frequencies of the most complex molecule, 1-methylthymine, were given zero weight in the refinement in order to provide an indication of the transferability of the calculated force constants.

Simultaneous refinement of force constants becomes progressively more difficult as the complexity of the involved molecules increases [6–11]. Best results have been

\* Agricultural Research Service, U.S. Department of Agriculture.

- [1] C. L. ANGELL, *J. Chem. Soc. (London)*, 504 (1961).
- [2] T. SHIMANOCHI, M. TSUBOI, and Y. KYOGOKU, *Advances in Chemical Physics* **7**, p. 435. Interscience, N.Y. (1964).
- [3] R. C. LORD and G. J. THOMAS, JR., *Spectrochim. Acta* **23A**, 2551 (1967).
- [4] H. SUSI and J. S. ARD, *Spectrochim. Acta* **27A**, 1549 (1971).
- [5] H. SUSI, J. S. ARD, and J. M. PURCELL, *Spectrochim. Acta* **29A**, 725 (1973).
- [6] J. H. SCHACHTSCHNEIDER and R. G. SNYDER, *Spectrochim. Acta* **19**, 117 (1963).
- [7] J. R. SCHERER, *Spectrochim. Acta* **19**, 601 (1963); *ibid.* **21**, 321 (1965); *ibid.* **24A**, 747 (1968).
- [8] N. NETO, M. SCROCCO and S. CALIFANO, *Spectrochim. Acta* **22**, 1981 (1966).
- [9] C. LA LAU and R. G. SNYDER, *Spectrochim. Acta* **27A**, 2073 (1971).
- [10] R. G. SNYDER and G. ZERBI, *Spectrochim. Acta* **23A**, 391 (1967).
- [11] P. COSSEE and J. H. SCHACHTSCHNEIDER, *J. Chem. Phys.* **44**, 97 (1966).

obtained with saturated hydrocarbons [6] (only  $\sigma$ -type CC and CH bonds); good results have been obtained with aromatic hydrocarbons [7, 8], chlorinated benzenes [7] and alkyl benzenes [9]. With ethers [10], and aldehydes and ketones [11], somewhat higher frequency errors must be accepted. Pyrimidine derivatives involve CC, CN, and CO bonds of ill-defined 'bond order' [12, 13] in highly conjugated configurations. The present study should throw some light on the feasibility of applying the Overlay approach to molecules of such complexity.

#### STRUCTURAL CONSIDERATIONS AND $G$ -MATRICES

All calculations were carried out on the basis of  $C_s$  molecular symmetry. The  $G$ -matrix [14] for uracil was based on X-ray data by STEWART and JENSEN [15], for cytosine on data by BARKER and MARSH [13], for thymine on data by GERDIL [16] and for 1-methylthymine on data by HOOGSTEEN [17]. The parameters for 1-methyluracil were taken to be the same as for uracil, with the 1-methyl group identical with the one of 1-methylthymine. The structural parameters reflect the fact that in these highly conjugated molecules classical bond orders have a limited meaning. Thus, in cytosine, the C—NH<sub>2</sub> bond and the ring CN 'double bond' have approximately the same length [13], the corresponding force constants are expected to have similar values, and no C=N group frequency is expected in the double bond region [5]. The crystals of uracil, thymine, and 1-methylthymine are monoclinic, space group  $C_{2h}^5 - P2_1/a$ , with four molecules per unit cell. Cytosine is orthorhombic, space group  $V^4 - P2_12_12_1$  with four molecules per unit cell, and 1-methyluracil is orthorhombic, space group  $V_h^{26}$ -Ibam, eight molecules per unit cell [18]. The detailed packing of the molecules is different for each investigated pyrimidine derivative. It results in different intermolecular hydrogen bonding [13, 15–18], and in variations of CO and NH frequencies which are not determined by the intramolecular force field alone. Because of factor group splitting [19], infrared and Raman frequencies do not coincide, except for the  $B_1$ ,  $B_2$ ,  $B_3$  branches of cytosine [5]. The splitting for most fundamentals is small [3–5]. Raman values were used for calculations except as noted in Tables 3–6. For molecules containing methyl groups, symmetrized  $G$ -matrices were constructed with generally accepted sets of symmetry coordinates [20] for the methyl groups. Tetrahedral angles were assumed because of the uncertainty of the available structural data. All other atoms are on the molecular plane of symmetry. Internal in-plane coordinates were defined as bond stretching and angle bending displacements in the usual manner [14].

Complications caused by crystal structure could be avoided by examining solution spectra, but unfortunately no suitable solvents for obtaining all necessary data could be found.

[12] R. E. MARSH, R. BIERSTEDT, and E. L. EICHHORN, *Acta Cryst.* **15**, 310 (1962).

[13] D. L. BARKER and R. E. MARSH, *Acta Cryst.* **17**, 1581 (1964).

[14] E. B. WILSON, JR., *J. Chem. Phys.* **7**, 1047 (1939).

[15] R. E. STEWART and L. H. JENSEN, *Acta Cryst.* **23**, 1102 (1967).

[16] R. GERDIL, *Acta Cryst.* **14**, 333 (1961).

[17] K. HOOGSTEEN, *Acta Cryst.* **16**, 28 (1963).

[18] S. FURBERG, A. HØRDVIK, and K. TAVGBOL, *Acta Chem. Scand.* **10**, 135 (1956).

[19] H. WINSTON and R. S. HALFORD, *J. Chem. Phys.* **17**, 607 (1949).

[20] W. T. KING and B. CRAWFORD, JR., *J. Molecular Spectr.* **5**, 421 (1960).

For molecules of high complexity a carefully chosen valence force field frequently yields better results than the often employed Urey-Bradley approximation [6, 7, 11, 21, 22]. It also has the advantage that bond length-force constant correlations can be used in a straightforward manner [4, 5, 21].

For the methyl groups, symmetry force constants based on standard symmetry coordinates were used, as indicated in Table 1. The choice is analogous to the one used by THOMPSON and FLETCHER for methyl azide [22] and permits a comparison with a variety of other molecules with methyl groups [21-24]. Available data did not permit the evaluation of additional interaction terms for the methyl groups.

Table 1. Force constants associated with methyl groups\*

	CCH <sub>3</sub>	NCH <sub>3</sub>
Sym. str.	4.956 (0.034) [4.78†]	4.626 (0.032) [4.793‡]
Antis. str.	4.746 (0.032) [4.76†]	4.761 (0.032) [4.694‡]
Sym. bend.	0.569 (0.009) [0.57†]	0.603 (0.010) [0.626‡]
Antis. bend.	0.538 (0.009) [0.54†]	0.530 (0.015) [0.498‡]
Rock	0.663 (0.021) [0.65†]	0.852 (0.038) [0.897‡]
X-CH <sub>3</sub> str.	4.889 (0.178)	4.889 (0.178)
X-CH <sub>3</sub> bend.	1.014 (0.078)	0.671 (0.055)
Sym. b, X-CH <sub>3</sub> st.	-0.465†	-0.320‡
X-CH <sub>3</sub> st., ring st.	0.471 (0.115)	0.471 (0.115)

\* Based on commonly used symmetry coordinates [20]. Standard error in parentheses. Units: stretching, mdyn/Å; bending, mdyn Å/rad<sup>2</sup>; stretch-bend interactions, mdyn/rad.

† Values for saturated hydrocarbons [6, 23].

‡ Values for N<sub>3</sub>CH<sub>3</sub>.

For the modified pyrimidine rings (including the attached H and O atoms, the 'outer' CC and CN linkages, and the NH<sub>2</sub> group), a scheme strictly analogous to previous studies on pyrimidine derivatives [4, 5] was adhered to. To summarize: (a) there is a force constant for each chemical bond and each bond angle; (b) bending force constants on either side of outer atoms and on either side of the -NH<sub>2</sub> and -CH<sub>3</sub> groups are assumed to be equal; (c) in general, only stretch-stretch interactions with one common atom and stretch-bend interactions with two common atoms are included; (d) *ortho* and *para* stretch-stretch interactions are included for the pyrimidine rings; and (e) interactions involving X-H stretching modes are ignored (because they could not be evaluated).

No exceptions were made to these rules, in order to provide as much orderliness as possible, although the inclusion of some additional interaction constants results in a slight improvement of some calculated frequencies [4, 5]. The aim of this study

[21] H. SUSI and J. R. SCHERER, *Spectrochim. Acta* **25A**, 1243 (1969).

[22] W. T. THOMPSON and W. H. FLETCHER, *Spectrochim. Acta* **22**, 1907 (1966).

[23] J. L. DUNCAN, *Spectrochim. Acta* **20**, 1197 (1964).

[24] A. B. HARVEY and M. K. WILSON, *J. Chem. Phys.* **45**, 678 (1966).

was to obtain a set of transferable constants for as wide a range of molecules as possible, without the use of special terms and exceptions.

Ring stretching force constants were calculated from bond length data by previously described procedures [4, 5, 21] and were not refined. This appears to be one way to obtain a sensible force field for rings with no symmetry except the molecular plane.

The described force field, even though it is severely constrained, contains a fair number of force constants. In order to make numerical calculations possible, and to provide transferable values, many similar terms were assumed to have identical values, in analogy with previous calculations on uracil and cytosine [4, 5]. The final

Table 2. Valence force constants for pyrimidine derivatives\*

Stretching†		Bending‡	
C-C	6.202	N-H } C-H <i>u</i> }	0.415 (0.003)
C=C	8.702	C-H <i>c</i>	0.329 (0.005)
C-N	6.380	C=O	1.170 (0.030)
C=N	7.340	C-NH <sub>2</sub>	0.492 (0.112)
C=O (2) <i>u</i>	11.000		
C=O (4) <i>u</i> } C=O <i>c</i> }	10.500		1.421 (0.130)
C-NH <sub>2</sub>	8.527		
C-H	5.160 (0.014)		0.644 (0.056)
N-H <i>u</i>	5.382 (0.019)		
N-H <i>c</i>	5.556 (0.040)		1.062 (0.089)
HNH <i>c</i>	6.030 (0.035)		1.353 (0.225)
			0.368 (0.012)
Stretch-stretch interactions		Stretch-bend interactions	
Ortho Ring, C-NH <sub>2</sub> }	1.013 (0.040)		0.420 (0.029)
Meta	-0.185 (0.045)		
Para	0.456 (0.065)		
Ring, C=O	1.560 (0.091)		0.223 (0.013)

\* The constants associated with methyl groups are listed in Table 1. Standard error is given in parentheses. Units: as in Table 1; *u*, *c*-uracil or cytosine derivatives only.

† Skeletal stretching constants calculated from bond length. Cf. Reference [4, 5].

‡ See text and Reference [4], Fig. 4, for further explanation. E.g., the N-H bending constant applies to the CNH bond angle on either side of the N-H linkage.

force constant matrix thus contained a relatively large number of off-diagonal elements, while the number of variable force constant values is within manageable limits, as indicated in Table 2.

#### EXPERIMENTAL DATA AND PROCEDURE

Laser Raman spectra and i.r. spectra of polycrystalline uracil, C,C-dideutero uracil, perdeutero uracil, cytosine and cytosine- $d_3$  have been reported [4, 5]. Infrared data of the remaining molecules were obtained as previously described. Raman data

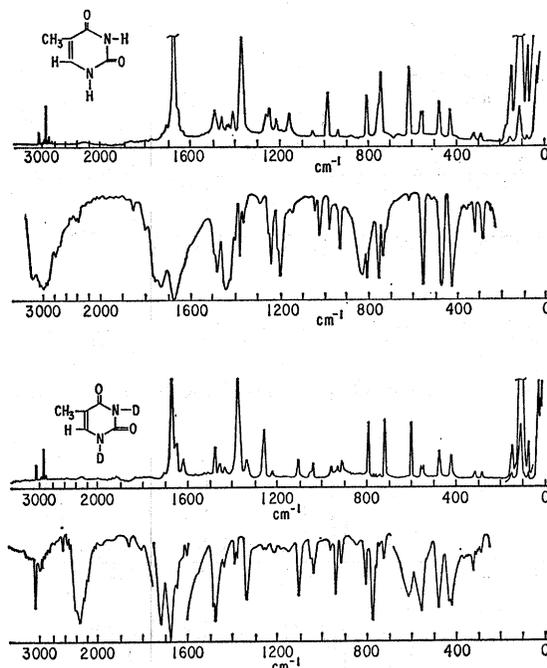


Fig. 1. Laser-Raman and infrared spectra of thymine and thymine- $d_2$ .

were obtained with a Spex Ramalog system\* equipped with an Argon ion laser. The 514.5 nm line was used for excitation, the power at the sample was  $\sim 200$  mW, the spectral slit width  $3\text{ cm}^{-1}$ . Samples were obtained from the Cyclo Chemical Corporation (grade I) and used without further purification. N-deuterated samples were prepared by repeated recrystallization from  $D_2O$ , until no NH bands were observed in the i.r. spectra of the products. The spectra of thymine and N-deuterated thymine are given in Fig. 1, of 1-methyluracil and 1-methyluracil- $d$  in Fig. 2 and of 1-methylthymine in Fig. 3. Numerical computations, based on the GF matrix formulation [14], were carried out by previous procedures [4, 5, 21] and computer programs.

\* Mention of commercial items is for your convenience and does not constitute an endorsement of this product over others of a similar nature by the U.S. Department of Agriculture.

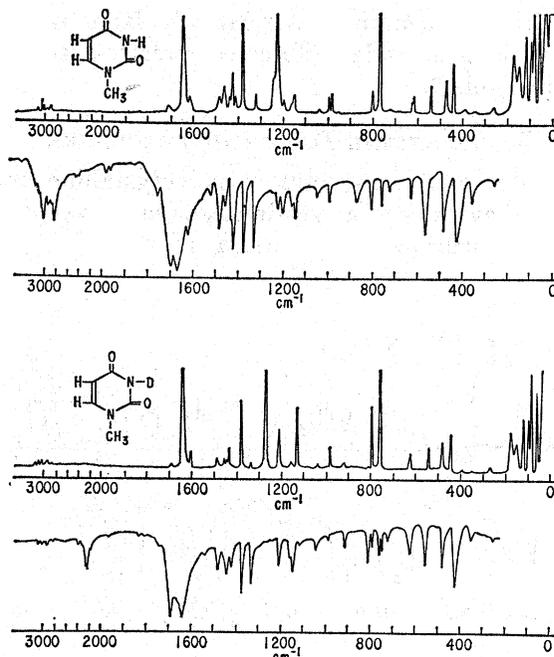


Fig. 2. Laser-Raman and infrared spectra of 1-methyluracil and 1-methyluracil-*d*.

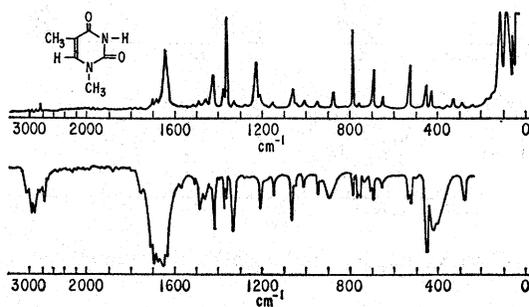


Fig. 3. Laser-Raman and infrared spectra of 1-methylthymine.

#### COMPUTATIONS AND ASSIGNMENTS

To obtain an initial assignment for the molecules which had not been previously investigated (see Introduction), zero order calculations were carried out with transferred force constants. Values transferred from the work of THOMPSON and FLETCHER [22] were used for the  $\text{NCH}_3$  groups. Symmetry force constants for the  $\text{CCH}_3$  groups were taken from saturated hydrocarbons, as reported by DUNCAN [23] on the basis of original work by SCHACHTSCHNEIDER and SNYDER [6]. The remaining values were taken from our previous work on uracil [4] and cytosine [5]. The zero order calculations were surprisingly successful, considering the fact that the force constants were taken from data of three different groups of investigators and derived from

different types of molecules. The average frequency error ranged from 12–18  $\text{cm}^{-1}$  for the different types of pyrimidine derivatives, the potential energy distribution appeared reasonable. It might be mentioned that an empirical assignment for methylated pyrimidine derivatives is next to impossible because of pronounced coupling between ring stretching modes, C–CH<sub>3</sub> and N–CH<sub>3</sub> stretching modes, and CH<sub>3</sub> rocking modes. In deuterated derivatives the situation is even worse because CD and ND deformations interact with skeletal vibrations [4, 5]. The in-plane assignments for thymine, 1-methylthymine, and 1-methyluracil are given in Tables 3–5. (The potential energy distribution values are from final refinements described below.) The planar fundamental frequencies of thymine-*d*<sub>2</sub> and 1-methyluracil-*d* are listed in Table 6 without a more detailed description in order to conserve space.

Subsequent refinements were carried out simultaneously for all molecules. The force constants mentioned in the previous paragraph were used as initial values. The frequencies of 1-methylthymine (the most complex one of the studied species) were given zero weight to obtain a check on the validity of the computed force field. All other frequencies were given a weight of  $1/\lambda$ . Frequency errors were minimized by iteration [6, 24]. All pyrimidine force constants were refined, except the skeletal stretching terms listed at the beginning of Table 2, which were calculated from bond length data, as described [4, 5]. The results of the first refinement were used to consolidate the force constants. If the computational uncertainty of similar constants (*e.g.*, ring and ring–ligand stretch–bend interactions, see Table 2) was higher than the difference between the computed values, these constants were set

Table 3. In-plane fundamentals of thymine ( $\text{cm}^{-1}$ )

	Infrared*	Raman	Calc.	Assignment†	
1	—	—	3129	$\nu$ NH (99)	
2	—	—	3126	$\nu$ NH (99)	
3	3063	3066	3083	$\nu$ CH (99)	
4	2993	2991	2990	$\nu^a$ CH <sub>3</sub> (99)	
5	2930	2934	2934	$\nu^s$ CH <sub>3</sub> (99)	
6	(1735)	1706	1701	$\nu$ C=O <sub>(2)</sub> (56)	U(I)
7	1677	1674	1674	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (54)	U(II)
8	—	1600	1606	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (77)	U(III)
9	1495	1492	1502	$\delta$ NH (64)	
10	1483	—	1484	$\nu$ -ring (52)	U(IV)
11	(1447)	1461	1461	$\delta^a$ CH <sub>3</sub> (81)	
12	1406	1409	1430	$\delta$ NH (78)	
13	1383	1379	1378	$\delta^s$ CH <sub>3</sub> (99)	
14	1366	1370	1357	$\delta$ CH (43) $\nu$ -ring (40)	
15	1245	1248	1250	$\nu$ -ring (60) $\delta$ CH (34)	U(V)
16	1203	1216	1214	$\nu$ -ring (65), C–Me (35)	
17	1152	1157	1153	$\nu$ -ring (72)	
18	1028	—	1018	$\nu, \delta$ -ring (63)	
19	984	985	984	r CH <sub>3</sub> (70)	
20	815	—	815	$\nu, \delta$ -ring (49), C–Me (24)	
21	—	806	803	$\nu, \delta$ -ring (52)	U(IX)
22	617	618	631	$\delta$ C=O (52)	U(XI)
23	560	562	569	$\delta$ -ring (59)	U(X)
24	475	475	477	$\delta$ -ring (43)	U(XII)
25	—	—	392	$\delta$ C=O (63)	U(XIII)
26	321	321	319	$\delta$ C–Me (66)	

\* Tentative assignments in parentheses.

† PED in parentheses; U(I) to U(XIII) indicate corresponding uracil modes, Ref. [4];  $\nu$ -stretching,  $\delta$ -bending, r-rocking.

Table 4. In-plane fundamentals of 1-methyl thymine

	Infrared*	Raman	Calc.	Assignment†	
1	3146	—	3127	$\nu$ NH (99)	
2	3064	3070	3083	$\nu$ CH (99)	
3	—	3009	2995	$\nu^a$ NCH <sub>3</sub> (99)	
4	2980	2989	2990	$\nu^a$ CCH <sub>3</sub> (99)	
5	2930	2927	2935	$\nu^s$ CCH <sub>3</sub> (99)	
6	2835	2838	2835	$\nu^s$ NCH <sub>3</sub> (99)	
7	1700	1704	1701	$\nu$ C=O <sub>(2)</sub> (61)	U(I)
8	1660	1645	1677	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (62)	U(II)
9	(1575)	—	1599	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (78)	U(III)
10	1516	1510	1504	$\nu$ -ring (58)	
11	1488	1487	1470	$\nu^a$ NCH <sub>3</sub> (47) $\nu^a$ CCH <sub>3</sub> (25)	
12	1465	1457	1460	$\nu^s$ CCH <sub>3</sub> (63) $\nu^a$ NCH <sub>3</sub> (14)	
13	(~1430)	1431	1434	$\delta$ NH (63)	
14	(1423)	—	1411	$\delta^s$ NCH <sub>3</sub> (42), $\nu$ -ring (23)	
15	1379	1382	1378	$\delta^s$ CCH <sub>3</sub> (96)	
16	1368	1368	1367	$\delta^s$ NCH <sub>3</sub> (51), $\nu$ -ring (25)	
17	(1337)	(1335)	1296	$\delta$ CH (44), $\nu$ -ring (22)	
18	—	1232	1244	$\nu$ -ring (90)	U(V)
19	1210	1210	1214	$\nu$ -skel (62)	
20	1152	1151	1131	$\nu$ -skel (82)	
21	1068	1063	1065	rNCH <sub>3</sub> (46)	
22	1012	1007	996	rCCH <sub>3</sub> (67)	
23	935	933	930	$\nu$ -skel (54)	
24	788	787	810	$\delta$ , $\nu$ -ring (54)	U(IX)
25	(756)	(760)	734	$\nu$ -skel (38), $\delta$ -ring (35)	
26	650	646	660	$\delta$ C=O (51)	U(XI)
27	520	522	509	$\delta$ -ring (43)	U(X)
28	450	448	463	$\delta$ -ring (48)	U(XII)
29	~390	390	384	$\delta$ C=O (62)	U(XIII)
30	—	323	321	$\delta$ C-CH <sub>3</sub> (54)	
31	275	—	265	$\delta$ N-CH <sub>3</sub> (71)	

\* Tentative assignments in parentheses.

† Skel refers to the PED sum of CC and CN stretching modes. PED in parentheses; U(I) to U(XIII) indicate corresponding uracil modes, reference [4];  $\nu$ -stretching,  $\delta$ -bending, r-rocking.Table 5. In-plane fundamentals of 1-methyl uracil (cm<sup>-1</sup>)

	Infrared*	Raman	Calc.	Assignment†	
1	(3145)	(3131)	3128	$\nu$ NH (99)	
2	3085	3079	3087	$\nu$ CH (98)	
3	3085	3079	3079	$\nu$ CH (99)	
4	—	2981	2994	$\nu^a$ CH <sub>3</sub> (99)	
5	2835	2841	2835	$\nu^s$ CH <sub>3</sub> (99)	
6	1695	(1720)	1700	$\nu$ C=O <sub>(2)</sub> (73)	U(I)
7	1662	1654	1648	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (70)	U(II)
8	(1621)	(1624)	1588	$\nu$ C=O <sub>(4)</sub> , $\nu$ C=C (75)	U(III)
9	1484	1491	1495	$\delta^a$ CH <sub>3</sub> (35), $\nu$ -ring (43)	
10	1459	1470	1458	$\delta^a$ CH <sub>3</sub> (44), $\nu$ -ring (40)	
11	1434	1444	1442	$\delta$ CH (57)	
12	1422	1420	1419	$\delta$ NH (65)	
13	1378	1386	1385	$\delta^s$ CH <sub>3</sub> (89)	
14	1331	1329	1302	$\nu$ -ring (43), $\delta$ CH (27)	
15	1225	1231	1256	$\nu$ -ring (92)	U(V)
16	1202	1202	1209	$\nu$ -ring (42), $\delta$ CH (54)	
17	1161	1156	1154	$\nu$ N-Me (45), $\nu$ -ring (29)	
18	1049	1047	1050	rCH <sub>3</sub> (42), $\nu$ -ring (44)	
19	994	998	1018	$\delta$ , $\nu$ -ring (68)	
20	804	806	830	$\delta$ , $\nu$ -ring (52), $\nu$ N-Me (20)	} U(IX)
21	—	775	785	$\delta$ -ring (44), $\nu$ N-Me (12)	
22	627	628	610	$\delta$ C=O (61)	U(XI)
23	560	549	556	$\delta$ -ring (60)	U(X)
24	480	482	482	$\delta$ -ring (50)	U(XII)
25	—	398	393	$\delta$ C=O (65)	U(XIII)
26	260	268	277	$\delta$ N-Me (89)	

\*, † As in Table 3.

Table 6. Observed\* and calculated frequencies

Thymine- $d_2$		1-methyl uracil- $d$		Cytosine		Cytosine- $d_3$	
Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
3064	3083	3087	3087	3354	3364	3088§	3086
2989	2990	3087	3079	3230	3248	—	3079
2934	2935	2980†	2995	3176	3177	2511	2506
2290†	2311	2835†	2835	—	3086	2376	2355
2290†	2305	(2260)†	2309	3055§	3079	2338	2342
1693†	1696	1693	1698	1694	1668	1648	1657
1661†	1669	1645	1644	1653	1650	1598	1607
1608†	1595	1609	1586	1612	1612	1514	1529
1482	1483	1489	1491	1533	1538	1491	1511
1460	1460	1457	1454	1498	1501	1383	1380
1378	1379	1434	1435	1462	1445	1316	1312
1375	1367	1383	1384	1361	1361	1282	1250
1338	1318	1336	1310	1276	1276	1190	1197
1259	1229	1272	1266	1247	1218	1130	1176
1223	1194	1213	1213	1148§	1142	1038	1053
1107	1122	1157	1173	1108§	1077	965§	983
1050	1049	1131	1134	990	991	949§	947
963	962	1043	1036	971	982	783	790
914	885	914†	914	792	775	777	755
802†	807	812†	816	597	583	550§	556
795	787	796	779	546	553	524	538
602	611	622	601	533	520	504	505
553	558	543	552	400	397	369	372
477	476	478	482				
—	390	395	391				
319	318	266	277				

Uracil		C,C-dideutero uracil		Perdeutero uracil	
Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
3130	3129	3130	3129	2326	2314
3130	3126	3130	3126	2316†	2307
3100	3087	2318	2309	2302	2304
3085	3079	2298	2281	2277†	2281
1715†§	1698	1713†§	1697	1699†§	1692
1664†	1646	1637†	1637	1628†	1629
1611	1597	1586	1570	1577	1559
1507	1498	1499	1496	1448	1450
1462	1467	1456	1459	1331	1329
1422	1430	1416	1422	1261	1238
1398	1411	1279	1288	1224	1211
1236	1251	1203	1208	1082	1097
1217†	1222	1089	1090	952	959
1104	1124	982	976	938†	906
1010	1027	912	900	848	866
988	977	851	861	838†§	839
792	798	791	783	784	782
579	584	567	566	561	556
558	569	539	561	552§†	551
540	543	525	533	534§	531
398	395	390	394	387	392

\* Raman frequencies used, expt. if otherwise indicated.

† Average between i.r. and Raman frequency. Cf. Ref. [4].

‡ Infrared only.

§ Assignment differs from Reference [4, 5].

equal. (Attempts to calculate more force constants than the available data justify are meaningless, both from the standpoint of physical reality and pragmatic transferability). The final set of force constants is given in Tables 1 and 2. The final overall average frequency error was below  $10\text{ cm}^{-1}$  (ranging from  $6\text{ cm}^{-1}$  for thymine to  $14$  for cytosine- $d_3$ , where ND deformation frequencies are strongly influenced by H-bonding). There might be some uncertainty about CH and NH stretching constants, because a great number of bands, evidently in Fermi resonance, are observed around  $3000\text{ cm}^{-1}$ , and it is not easy to pinpoint the fundamentals. Because these modes are not coupled with other fundamentals, this uncertainty does not influence the rest of the force constants.

## DISCUSSION

### *The force field*

The described approach constitutes a compromise between efforts to obtain a physically meaningful force field and a set of pragmatic transferable force constants. The number of off-diagonal  $F$ -matrix elements is reasonably high, and a scheme has been devised for systematic selection. The number of distinct values is low. Objections could be raised, for instance, to the 'averaged' *meta* and *para* ring interactions. It was not possible to calculate more specific values with any degree of reliability. The set of reported constants does reproduce the frequencies of ten complex molecules with fair accuracy. The nature of the force field is quite similar to valence force fields employed for chlorinated benzenes [7], alkyl benzenes [9], and diverse compounds with methyl groups [21–24]. This facilitates intercomparison and transfer of force constants. The final values for methyl groups, in particular, were reasonably close to initial, transferred values, as indicated in Table 1. The constants for the pyrimidine rings were reasonably close to values obtained previously for uracil [4] and cytosine [5].

We hope that the systematic scheme for selection of off-diagonal elements, and the relatively low number of constants with distinct values, facilitates the construction of valence force fields for a still broader range of complex molecules and the calculation of their vibrational frequencies.

### *Potential energy distribution and calculated frequencies*

Predominant terms in the Potential Energy Distribution (PED) for thymine, 1-methylthymine, and 1-methyluracil are given in parentheses in the fifth column of Tables 3–5. The nature of the complete PED matrix is too complex to permit much insight into the detailed nature of most vibrations. The frequencies of deuterated species and of previously studied molecules are summarized in Table 6. The PED and the calculated frequencies for the latter are very close to previously reported results [4, 5]. Some minor changes in assignment of observed frequencies are indicated in the table.

Most fundamentals are heavily mixed, as expected [1–3]. In particular,  $\text{CH}_3$  rocking and  $\text{X}-\text{CH}_3$  stretching modes are strongly coupled with skeletal vibrations of the ring. The skeletal modes which do resemble corresponding vibrations of the parent compounds are indicated by Roman numerals and briefly discussed in the next section.

The frequency agreement appears acceptable for highly complex conjugated structures observed in the crystalline state. In addition to the usual uncertainties caused by the harmonic approximation and the simplified force field [6], complications arise because of hydrogen bonding and crystal field effects. Thus the poorest frequency agreement was observed with cytosine and cytosine- $d_3$ , which have a very complex hydrogen bonding network [13] and exhibit apparently irregular NH (ND) deformation frequencies. In contrast, the frequency agreement for 1-methylthymine, the most complex of the studied molecules, was as good as the average although the frequencies were given zero weight in the computations. Some caution should be exercised in drawing conclusions from the exact carbonyl frequencies because these are strongly influenced by the particular hydrogen bonding scheme in each crystal. In crowded spectral regions where some calculated frequencies are only 10–20  $\text{cm}^{-1}$  apart assignment errors are possible. We attempted to minimize such errors by proceeding systematically from simpler to more complex molecules. The small change of transferred  $\text{CH}_3$  force constants is also an encouraging sign.

#### *Group vibrations*

A few words are in order regarding regularities observed in the spectra of this group of molecules, i.e., about group frequencies, in the light of the described computations.

The skeletal modes are indicated in the last column of Tables 3–5 by a previously described numbering system [4, 5] (U I to U XIII are the planar modes of the heavy-atom skeleton of uracil). The double bond stretching modes (I, II, III), the *V* mode (which resembles the  $B_{2u}$  stretch of benzene [4, 7]) and the low frequency modes IX to XIII retain some similarity throughout the series. The IV, VI, VII, and VIII vibrations are strongly coupled with  $X\text{-CH}_3$  modes. Skeletal modes are generally more intense and more easily identified in the Raman effect. Medium to strong characteristic lines are observed around 1230–1250  $\text{cm}^{-1}$  (*V*, ring stretch) and close to 800  $\text{cm}^{-1}$  (IX, complex bending). Despite its complex nature, the 800  $\text{cm}^{-1}$  mode appears to be stable and characteristic. (The Cartesian displacements for these vibrations are indicated in Fig. 5, References 4 and 5.) In deuterated derivatives the frequency variations of the skeletal modes are larger and their identification is more difficult because of coupling with ND deformations.

Of the CH and NH deformation modes, the bending vibration of the 3-position NH linkage of uracil derivatives is observed between 1400 and 1430  $\text{cm}^{-1}$  throughout the series (PED 60–80%). This mode is useful in conformational studies of biopolymer models [25]. The  $\text{CH}_3$  ‘umbrella modes’ retain high purity and a relatively constant frequency in most studied molecules. [In 1-methylthymine some coupling with skeletal motion is observed (Table 4).] The asymmetric bending modes are less stable. The rocking modes are strongly coupled to other vibrations in a complex manner.

---

[25] H. T. MILES, T. P. LEWIS, E. D. BECKER, and J. FRAZIER, *J. Biol. Chem.* **248**, 1115 (1973).