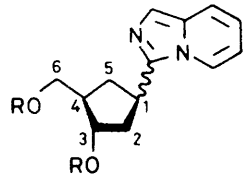
Carbocyclic 2'-Deoxyribonucleoside Analogues: Assignment of Epimeric Configuration and Conformation by ^{1}H Nuclear Magnetic Resonance Spectroscopy

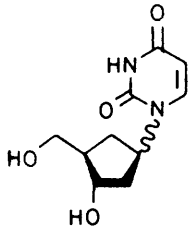
Richard C. Cookson and Philip J. Dudfield
Department of Chemistry, The University, Southampton SO9 5NH
Graham Klinkert*
Analytical Research Department, Glaxo Group Research Ltd., Ware, Herts., SG12 ODJ

The determination of epimeric configuration of some novel carbocyclic 2'-deoxyribo-*C*-and -*N*-nucleosides by ^{1}H n.m.r. spectroscopy is described. Assignments were derived initially from $^3J_{HH}$ vicinal coupling constants and then confirmed by n.O.e. difference spectroscopy. In both the *C*-nucleosides and *N*-nucleosides the β-epimers appear to adopt the same predominant conformation in solution, whereas in the α-epimers the equilibrium conformation is an average, with contribution from a number of different conformations.

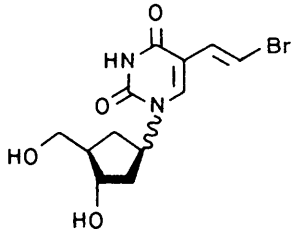
Our synthetic pathway to the carbocyclic 2'-deoxyribonucleosides produced a mixture of α- and β-epimers† which have been separated in three cases; the imidazo[1,5-*a*]pyridine-*C*-nucleosides (**1**) and the uridine-(**2**) and bromovinyluridine (**3**) -*N*-nucleosides.[1] The spectroscopic problem which faced us was to distinguish between the epimers which is not easy in view of the conformational lability of the cyclopentane ring. Both the nature and orientation of electronegative substituents affect vicinal coupling constants in the ring and this is an additional complicating factor. In order to substantiate deductions made from coupling constants, nuclear Overhauser effect (n.O.e.) difference experiments were used wherever possible since these usually provide unambiguous information.



(**1**) a; R = H, α-epimer
b; R = H, β-epimer
c; R = CF_3CO, α-epimer
d; R = CF_3CO, β-epimer



(**2**) a; α-epimer
b; β-epimer

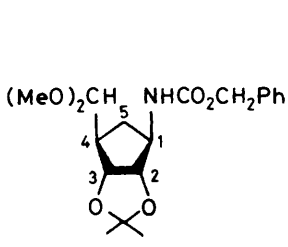


(**3**) a; α-epimer
b; β-epimer

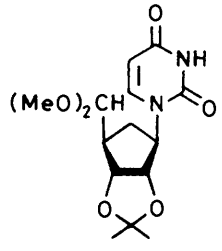
† For consistency, α- and β-epimers in the carbocyclic series correspond, respectively, to the α- and β-*anomer* terminology of their ribofuranosyl counterparts.

Results and Discussion

1. *Isopropylidene Derivatives.*—Before we turn our attention to the nucleosides themselves, it is instructive to consider the n.m.r. spectra of the two isopropylidene intermediates (**4**) and (**5**), prepared *en route* to the ribose carbocyclic nucleosides. Both of these have the β-configuration from their synthetic



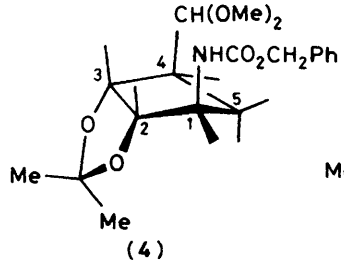
(**4**)



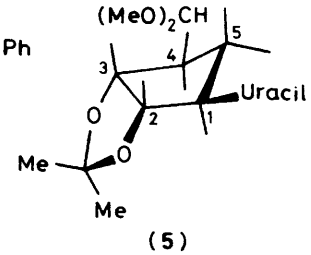
(**5**)

origin but they adopt quite different conformations in dueteriochloroform solution. Thus in the urethane (**4**), $J_{1,2}$ and $J_{3,4}$ are very small (see Table 1), indicating dihedral angles of *ca.* 90°. This means that the conformation must be as shown in Figure 1, with the C-1 and C-4 substituents axial. Further evidence is provided by the relatively small values of $J_{1,5\beta}$ and $J_{4,5\beta}$. In contrast, in the uracil (**5**) one of the C-5 protons shows two vicinal couplings of 11 Hz. Vicinal couplings of this magnitude can only be associated with two axial–axial relationships, hence the conformation must be as shown, with the C-1 and C-4 substituents equatorial.

Although at first sight it would appear that the *endo*-envelope conformation of compound (**4**) is energetically less favourable than the *exo*-envelope conformation of compound (**5**) it has



(**4**)



(**5**)

Figure 1. Conformations of the isopropylidene intermediates (**4**) and (**5**)

Table 1. ^{1}H N.m.r. chemical shifts (δ) [$CDCl_3$] and coupling constants (Hz) of the isopropylidene intermediates

	1-H	2-H	3-H	4-H	5-H_α	5-H_β	$J_{1,2}$	$J_{1,5\alpha}$	$J_{1,5\beta}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5\alpha}$	$J_{4,5\beta}$
(4)	4.15	4.38	4.58	2.3—2.5		1.60	1.5	—	1—2	6	1	—	1—2
(5)	4.6[a]	4.75	4.6[a]	2.52	2.30	2.05	5.5	7.5	11	7	4.5	7.5	11

[a] Approximate value since multiplets overlap.

Table 2. ^{1}H N.m.r. chemical shifts (δ) [$CDCl_3$] and coupling constants (Hz) of imidazo[1,5-*a*]pyridine carbocyclic 2′-deoxyribo-*C*-nucleosides

	1-H	2-H_α	2-H_β	3-H	4-H	5-H_α	5-H_β	6-H	6-H′	$J_{1,2\alpha}$	$J_{1,2\beta}$	$J_{1,5\alpha}$	$J_{1,5\beta}$	$J_{2\alpha,3}$	$J_{2\beta,3}$	$J_{3,4}$	$J_{4,5\alpha}$	$J_{4,5\beta}$
(1a)	3.7[a]	2.2—2.5		4.22	2.3[a]	2.09	1.82	3.6—3.8		7	—	2	8	6.5	4	4	8	9.5
(1b)	3.75[a]	2.27, 2.17		4.40	2.28	2.43	1.92	3.7—3.8		8	8	8	8	5.5	5.5	5.5	8	8
(1c)	3.69	2.84, 2.47		5.28	2.99	2.61	2.08	4.4—4.6		7.5	7.5	6.5	8.2	7.5	7.5	7	9	8.2
(1d)	3.82	2.41	2.65	5.45	2.78	2.60	2.04	4.5		7	10	7.5	10	3	7	4	8	9[b]

[a] Approximate value since multiplets overlap. [b] Also a long-range (*W*) coupling $^4J_{2.5\alpha}$ 1.5 Hz.

Table 3. ^{1}H N.m.r. chemical shifts (δ) and coupling constants (Hz) of uridine carbocyclic 2′-deoxyribonucleosides

	1-H	2-H_α	2-H_β	3-H	4-H	5-H_α	5-H_β	6-H	6-H′	$J_{1,2\alpha}$	$J_{1,2\beta}$	$J_{1,5\alpha}$	$J_{1,5\beta}$	$J_{2\alpha,3}$	$J_{2,\beta3}$	$J_{3,4}$	$J_{4,5\alpha}$	$J_{4,5\beta}$
(2a)[a]	4.90	1.88$_\alpha$	2.50	4.11	2.30	2.1, 2.0		3.7—3.6		8	8	8	8	8	6	7	—	—
(2b)[a]	5.02	2.0—2.2		4.24	2.1[c]	2.36	1.59	3.8—3.6		—	—	7	10	5.5	5.5	5.5	7	10
(3a)[b]	5.00	1.84	2.43	4.10	2.28	2.06		3.6—3.5		8	8	8	8	5.5	6	5.5	8	8
(3b)[a,d]	4.80	1.89	1.98	4.03	1.92	2.15	1.40	3.6—3.4		8.7	9	7.7	10.3	4.5	7.5	5.7	7.5	10.3[e]

[a] Run in D_2O. [b] Run in CD_3OD. [c] Approximate value since multiplets overlap. [d] 500 MHz Spectrum. [e] Also a long-range (*W*) coupling $^4J_{2\alpha.5\alpha}$ 1.2 Hz.

been shown that the former is generally preferred for bicyclic systems in which a cyclopentane ring is fused to either a three[2] or a four-membered[3] ring. According to work by Paulsen and Brauer, this preference appears to hold for the 2,4-dioxabicyclo[3.3.0]octane system.[4] Presumably, in compound (**5**) non-bonded interactions with the bulky uracil group are sufficient to tip the balance in favour of the *exo*-envelope conformation.

This information proved to be extremely helpful in our assignment of the carbocyclic 2′-deoxyribose epimers (*vide infra*).

2. *Imidazopyridine Carbocyclic 2′-Deoxyribo-C-nucleosides.*—The ^{1}H n.m.r. spectra of the α- and β-epimer (**1a**) and (**1b**) in deuteriochloroform were very similar (see data in Table 2); furthermore there was extensive overlap of the aliphatic signals which prevented a full analysis of the spectra. Those vicinal couplings which could be measured were neither very large nor very small, suggesting no major single conformation for either epimer, thus making assignment of configuration on this basis impossible. However, the C-5 methylene protons could be differentiated from the C-2 protons by decoupling of the 3-H proton. An n.O.e. difference experiment, saturating 3-H, did identify 5-H_β in each epimer but did not allow assignment of configuration because of overlapping signals and the complication of selective population transfer (SPT) effects masking the n.O.e.s. [A method of eliminating SPT effects (described below) was used successfully in the *N*-nucleoside series but was not developed until after the work on the *C*-nucleosides had been completed].

In an attempt to simplify the spectra, the 3,6-bis(trifluoroacetyl) derivatives were prepared *in situ*. This succeeded in resolving all the individual aliphatic proton signals, enabling measurement of all the coupling constants (see Table 2). Once again, irradiation of 3-H allowed differentiation of the C-2 and C-5 methylene groups, but assignment of the individual protons was not easy. However, one of the epimers, (**1d**), had a C-5 proton with vicinal couplings of 10 and 9 Hz and, by analogy with the isopropylidene derivatives, this can only be accommodated by two axial–axial couplings, thus establishing a *cis* relationship between the C-1 and C-4 substituents, *i.e.*, the β-epimer with a probable conformation for the cyclopentane ring as shown in Figure 2. The other vicinal couplings are consistent with this conformation.

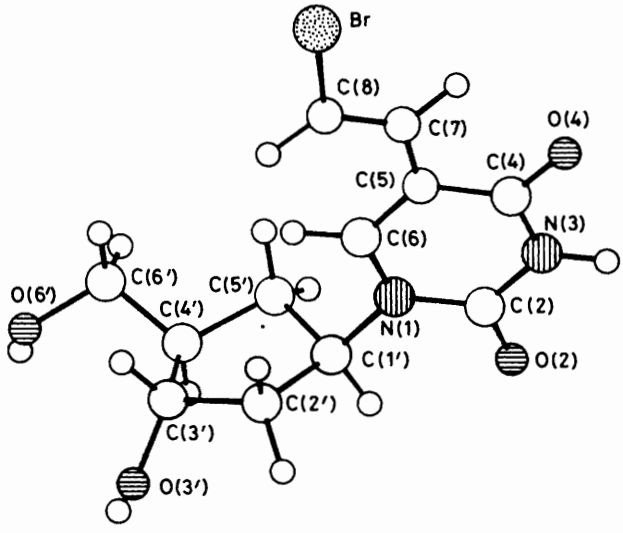


Figure 2. Conformation of β-carbocyclic (*E*)-5-(2-bromovinyl)-2′-deoxyuridine (**3b**)

In the other epimer, (**1c**), now assumed to be the α-epimer, it was not possible to assign the individual C-2 and C-5 protons because all the vicinal couplings were in the 6—9 Hz range.

Attempted n.O.e. difference experiments to confirm these assignments failed due to decomposition of the trifluoroacetyl derivatives.

3. *Uridine- and Bromovinyluridine-2′-deoxy Carbocyclic N-Nucleosides.*—These compounds were insoluble in deuteriochloroform so their n.m.r. spectra were recorded in either D_2O, CD_3OD, or $(CD_3)_2SO$. The conformation of the cyclopentane ring, inferred from the vicinal coupling constants, did not vary

significantly in these three solvents. As expected the aliphatic region was essentially the same for the corresponding uridine and bromovinyluridine epimers so their spectra will not be discussed separately. The chemical shifts and coupling constants are presented in Table 3.

Comparison of the spectra of the epimeric *N*-nucleoside analogues again showed one epimer having a C-5 proton with two vicinal couplings of 10 Hz which, as discussed above, can only arise in the β-epimer. In the case of the α-epimers, as in the imidazopyridine series, all the couplings were in the 6—9 Hz range.

The cyclopentane ring is known to be very flexible and can take up a large number of non-planar conformations which are interconverted by a pseudorotation of puckering.[2] The pseudorotation may be described by a mathematical equation involving a phase angle of puckering which can vary from 0—360°. Algorithms have been written to describe this pseudorotation and display it using molecular graphics.[5] At each point in the pseudorotation circuit, the program calculates a potential-energy function, involving terms for Van der Waals and torsional non-bonded interactions. It also calculates vicinal coupling constants using a Karplus-type equation and matches these against the observed values. The quality of fit is given by computation of the root-mean-square (r.m.s.) deviation [equation (1)] of each conformation.

$$\text{rms deviation} = \sqrt{\frac{\sum_{i=1}^{N} (J_{i(\text{calc})} - J_{i(\text{obs})})^2}{N}} \qquad (1)$$

In cyclopentane itself there is little difference in energy between conformations, but in trisubstituted cyclopentanes interactions involving the substituents will result in larger energy variations over the pseudorotation circuit. In favourable situations there will be a single potential-energy minimum, corresponding to a single conformation, and in such cases one would expect good agreement between calculated and observed coupling constants. This was in fact the case when the measured *J* values for the β-epimer (**3b**) were input to the pseudorotation program and an excellent r.m.s. deviation of ±0.95 Hz was obtained with the pseudorotational phase angle = 233°. The conformation at this point is shown in Figure 2.

In contrast, when the program was run for the α-epimer, there was no single potential-energy minimum, indicating contributions from several conformations in solution, and no r.m.s. deviation lower than ±2.3 Hz was obtained.

In order to re-check the assignment of configuration to the epimers (**3a**) and (**3b**), the program was re-run to associate the experimental coupling constants found for (**3a**) with the β-epimer and *vice versa*. Gratifyingly, in both cases the r.m.s. deviation was worse than before, thus substantiating our original assignments. In other instances this has proved to be a useful technique for confirming assignments of configuration of 5-membered-ring stereoisomers made on the basis of vicinal coupling constants.[6]

N.O.E. Difference Experiments.—Advances in n.m.r. instrumentation in recent years have made it possible to detect n.O.e.s of less than 1% with relative ease using difference spectroscopy,

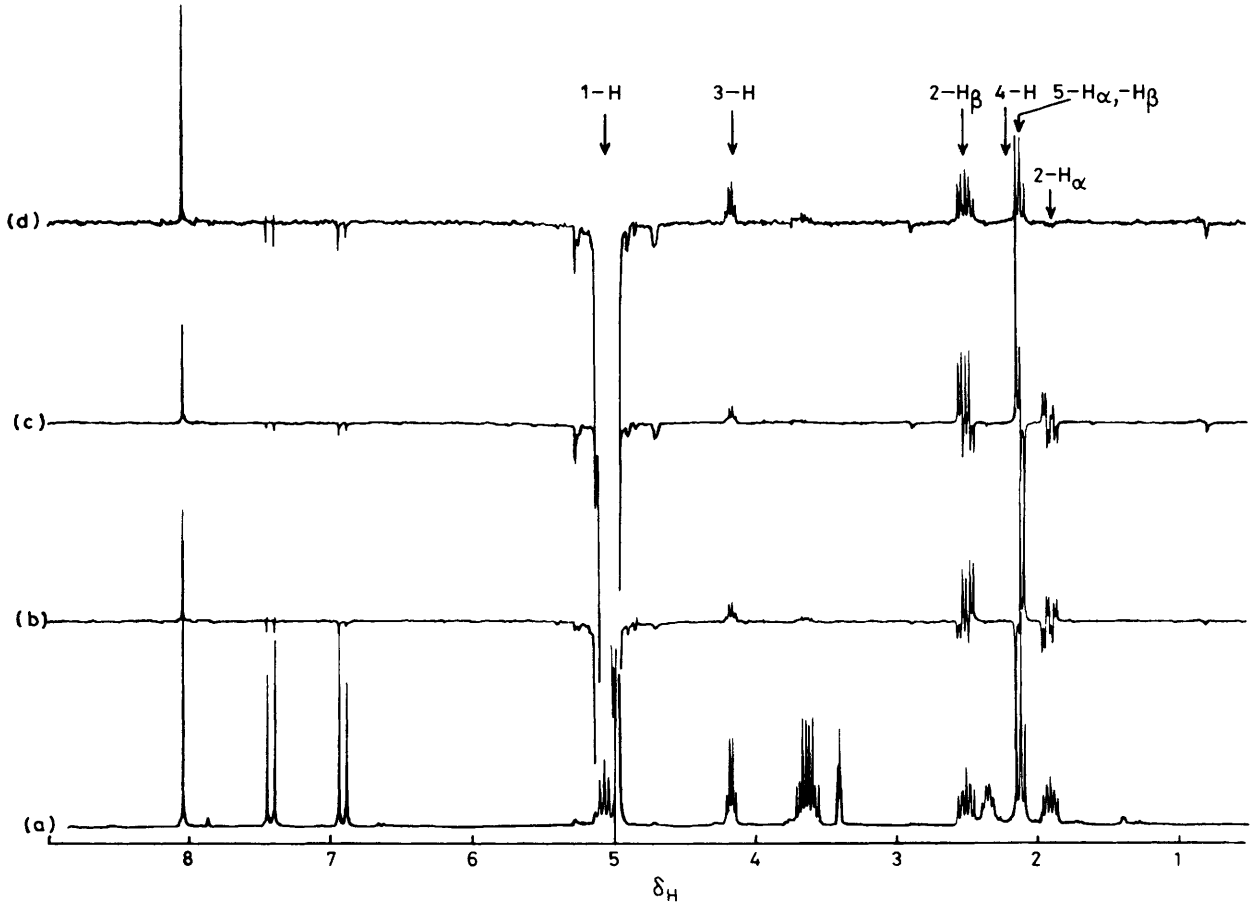


Figure 3. N.O.e. difference spectra of α-carbocyclic (*E*)-5-(2-bromovinyl)-2′-deoxyuridine (**3a**). (a) Normal spectrum (irradiation off-resonance); (b) n.O.e. difference spectrum, irradiation of line 2 of the 1-H multiplet; (c) n.O.e. difference spectrum, irradiation of line 4 of the 1-H multiplet; (d) spectrum (b) + spectrum (c)

and this has become an important method for providing unambiguous stereochemical information about molecules in solution.

Although n.O.e.s have been used in the nucleoside field to investigate the *syn-anti* conformation of the base relative to the sugar ring,[7,8] with one early exception[9] they do not appear to have been applied to determination of anomeric configuration. We have recently reported such an application to some 2-deoxyribo-*C*-nucleosides[10] and, in view of the simplicity of the method, have now extended it to carbocyclic analogues.

Our approach has been to saturate a signal well separated from the others. This is usually either the 1-H proton, which resonates at lowest field in the *N*-nucleoside series, or the 3-H proton in the *C*-nucleoside series. An n.O.e. at 4-H upon saturation of 1-H means that the two protons are on the same face of the ring, *i.e.*, a β-epimer, whilst an n.O.e. at 3-H on saturation of 1-H is indicative of an α-configuration. The nuclear Overhauser effect is strongly distance-dependent, therefore the C-2 and C-5 protons also show enhancements. Unfortunately, because 2-H_α, 2-H_β, 5-H_α, and 5-H_β are all scalar (*J*)-coupled to 1-H, SPT also occurs, which can obscure the n.O.e.s. SPT occurs whenever the lines of a multiplet are being saturated to different extents (which is normally the case at the low decoupler power levels used for n.O.e. experiments). The result of SPT is a distortion of intensities of the lines of any multiplet which is scalar-coupled to the signal being irradiated—some are increased while others are decreased. (SPT is the phenomenon observed in INDOR experiments.) Thus it can be seen that for signals which are both dipolar (n.O.e.). and scalar (*J*)-coupled to the signal being irradiated, the resulting difference spectrum will show an unpredictable combination of n.O.e. and SPT effects, with some lines of the multiplet positive, some negative, and some even cancelled altogether.[11]

To remove unwanted SPT effects we have utilised the method of Neuhaus.[12] To ensure that the individual lines of the multiplet are saturated to the same extent, he includes the frequency of each line in the frequency list for automatic acquisition of the free induction decays (FIDs). By adding the resulting FIDs together and subtracting an appropriately weighted control spectrum, the SPT effects cancel, leaving pure n.O.e. effects.

In our case, for compounds (**3a**) and (**3b**) we found it unnecessary to irradiate every line in the 1-H multiplet, merely two 'complementary' lines (lines 2 and 4). This does gain an advantage in experiment time. The results for (**3a**) are shown in Figure 3. As expected for the α-epimer, there is an n.O.e. observed for 3-H but none for 4-H. On saturation of line 2 or line 4, a typical SPT pattern is observed for both 2-H_α and 2-H_β and the C-5 protons (which happen to be isochronous in this epimer), but when these two difference spectra are added the

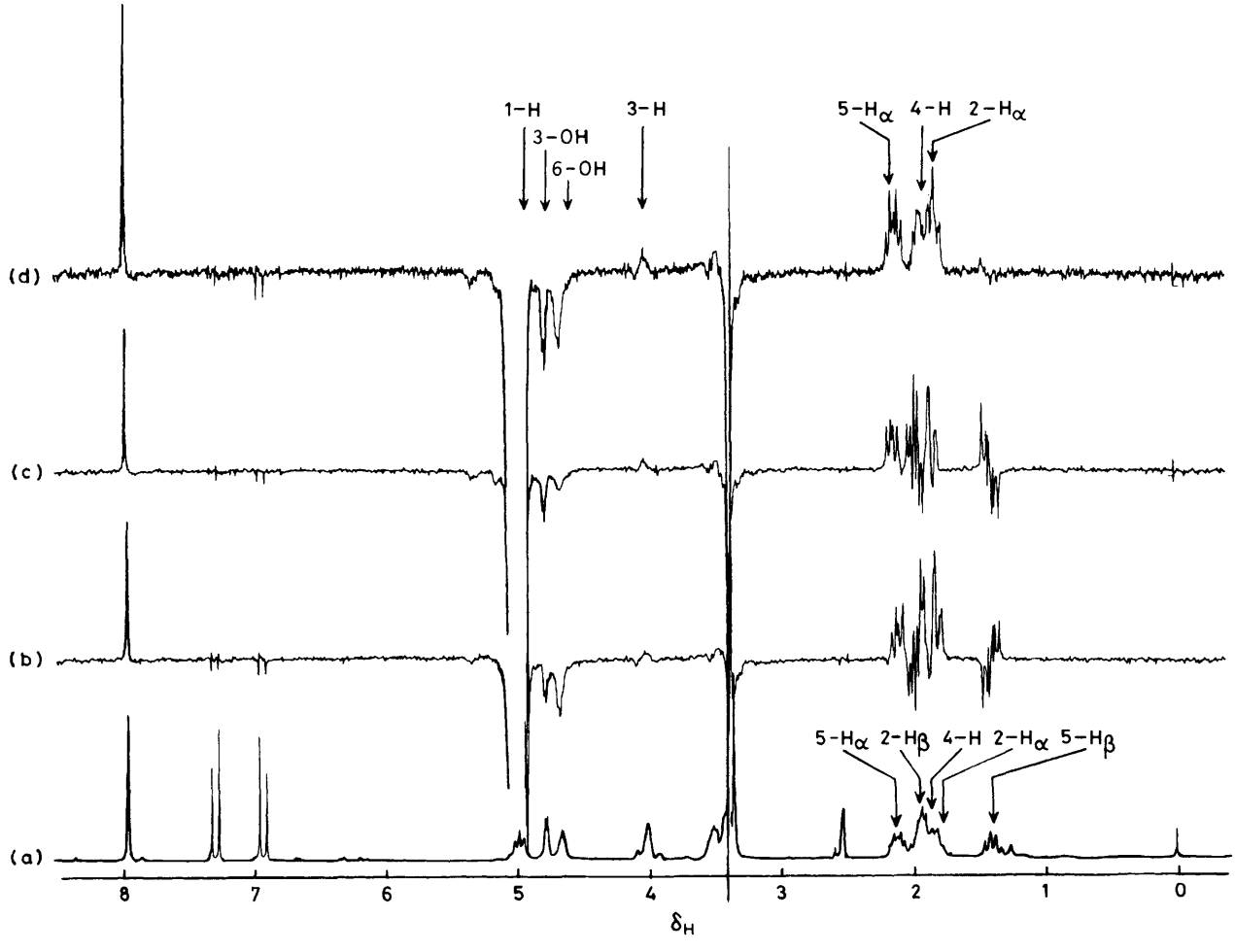


Figure 4. N.O.e. difference spectra of β-carbocyclic (*E*)-5-(2-bromovinyl)-2′-deoxyuridine (**3b**). (a) Normal spectrum (irradiation off-resonance); (b) n.O.e. difference spectrum, irradiation of line 2 of the 1-H multiplet; (c) n.O.e. difference spectrum, irradiation of line 4 of the 1-H multiplet; (d) spectrum (b) + spectrum (c)

SPT effects are cancelled and 2-H_α (with no n.O.e.) disappears from the spectrum, thus allowing the unequivocal assignment of 2-H_α and 2-H_β.

For compound (**3b**) (see Figure 4) the same experiment provides even more dramatic results; here, the signals due to 2-H_β and 5-H_β disappear from the spectrum, leaving n.O.e.s for 5-H_α, 4-H, and 2-H_α only, which confirm the β-configuration. Although (at 250 MHz) (signals for 2-H_α, 2-H_β, and 4-H partially overlap, examination of a 500 MHz spectrum of compound (**3b**) leaves no doubt which signals have disappeared. [The small n.O.e. for 3-H, which is somewhat misleading, is from the unintended partial saturation of the two OH resonances during the experiment.]

Conclusions

Over the years much effort has been expended on devising empirical rules for the determination of epimeric configuration in nucleosides by n.m.r. spectroscopy. A recent review[13] summarises this work, all of which relies on characteristic ^1H or ^{13}C chemical-shift differences or differences in the appearance of particular multiplets in the ^1H spectra. These methods all suffer the disadvantage of being restricted to a closely related series of structures. New rules have to be devised for each new series. For the carbocyclic nucleosides described in this paper, it is possible to assign epimeric configuration on the basis of vicinal coupling constants but only in circumstances where a single conformer predominates in solution. Even in such cases it is often necessary to obtain a high-field (500 MHz) spectrum or to prepare a derivative in order to extract the coupling constants. Nuclear Overhauser effect difference experiments, particularly when modified to eliminate SPT effects, provide a rapid, unambiguous, and generally applicable means of establishing the epimeric configuration, whether or not there is a single conformation present in solution.

Experimental

The n.m.r. experiments described in this paper were all carried out at 250 MHz on a Bruker WM250 spectrometer using 16K data points and a spectral width of 4.5 KHz. For the n.O.e. difference experiments, 32 transients were collected at each irradiating frequency in turn, including the control (off-resonance) frequency, and this cycle was repeated overnight, resulting in a total of *ca.* 3 000 scans for each frequency. The decoupler power employed was 30L (γH_2 3.5 Hz). The addition and subtraction of FIDs and spectra were accomplished using standard Bruker software. No line-broadening was used.

The nuclear Overhauser enhancements have not been quantified. Although it is our normal practice to integrate the positive n.O.e. signals against the negative signal due to the saturated multiplet in the difference spectra, the experimental variables[14] and in this case the manipulation of the difference spectra to eliminate SPT effects removes any quantitative significance from the results.

All the coupling constants referred to in the text have been measured directly from the spectra. None have been calculated.

The trifluoroacetyl derivatives (**1c**) and (**1d**) were prepared by the addition of a few drops (excess) of trifluoroacetic anhydride to a *ca.* 0.2M solution of (**1a**) or (**1b**) in $CDCl_3$ (0.5 ml) in a 5 mm n.m.r. tube. Excess of acid was neutralised by shaking with saturated sodium hydrogen carbonate. The organic phase was dried (Na_2SO_4) and filtered into a second n.m.r. tube to record the spectrum.

Acknowledgements

We thank Dr. P. Murray-Rust and Dr. A. P. Tonge of the Chemical Research Dept., Glaxo Group Research Ltd., Greenford for the use of the pseudorotation program and molecular graphics facilities, and Dr. A. Derome of Dyson Perrins Laboratory, Oxford University for the 500 MHz spectrum of compound (**3b**).

References

1 R. C. Cookson, P. J. Dudfield, and D. I. C. Scopes, preceding 2 papers.
2 B. Fuchs, 'Topics in Stereochemistry,' eds E. L. Eliel and N. L. Allinger, Interscience, New York, 1977, vol. 10, p. 1.
3 A. Brown, R. Glen, P. Murray-Rust, J. Murray-Rust, and R. F. Newton, *J. Chem. Soc., Chem. Commun.*, 1979, 1178.
4 H. Paulsen and O. Brauer, *Chem. Ber.*, 1977, **110**, 342.
5 P. Murray-Rust and A. P. Tonge, *J. Mol. Graphics*, 1984, **2**, 58.
6 D. P. Reynolds, personal communication.
7 P. A. Hart and J. P. Davis, *Biochem. Biophys. Res. Commun.*, 1969, **34**, 733.
8 M. Legraverend, J. M. Lhoste, and E. Bisagni, *Tetrahedron*, 1984, **40**, 709.
9 R. Cushley, B. L. Blitzer, and S. R. Lipsky, *Biochem. Biophys. Res. Commun.*, 1972, **48**, 1482.
10 L. J. S. Knutsen, B. D. Judkins, R. F. Newton, D. I. C. Scopes, and G. Klinkert, *J. Chem. Soc., Perkin Trans. 1*, 1985, 621.
11 J. K. M. Sanders and J. D. Mersh, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1982, **15**, 353.
12 D. Neuhaus, *J. Magn. Reson.*, 1983, **53**, 109.
13 C. K. Chu, F. M. El-Kabbani, and B. B. Thompson, *Nucleosides Nucleotides*, 1984, **3**, 1.
14 L. D. Hall and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1980, **102**, 5703.

Received 10*th May* 1985; *Paper* 5/783