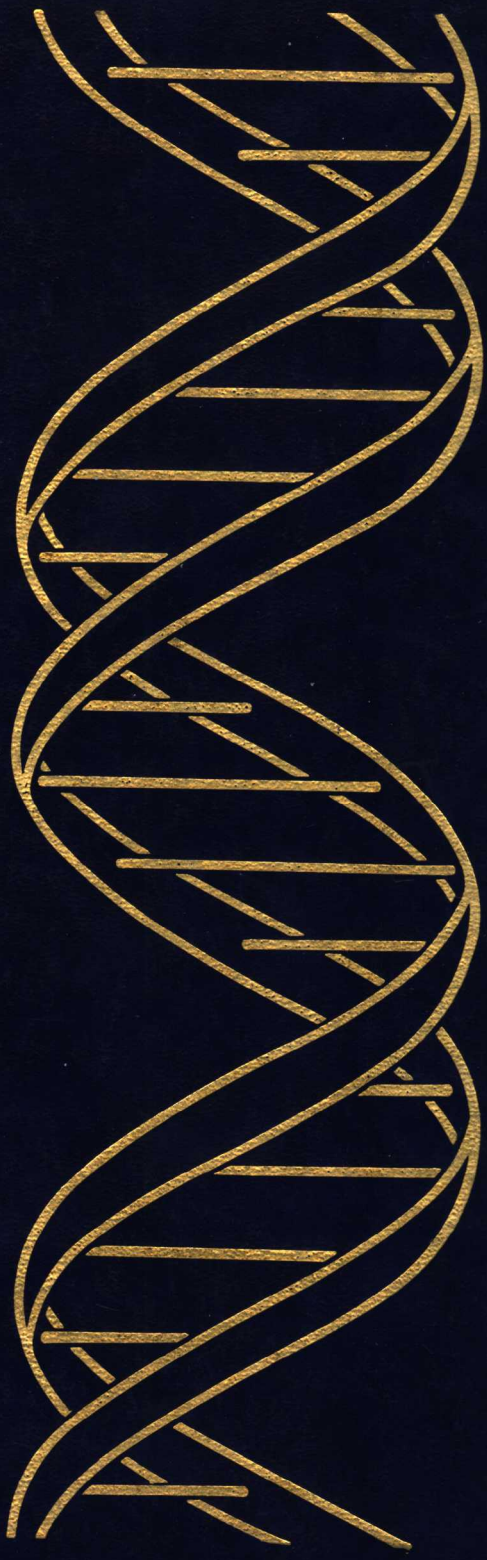
Topics in Nucleic Acid Structure

PART 3

EDITED BY

Stephen Neidle



TOPICS IN MOLECULAR AND STRUCTURAL BIOLOGY

9

TOPICS IN NUCLEIC ACID STRUCTURE Part 3

Edited by

STEPHEN NEIDLE

Cancer Research Campaign Biomolecular Structure Unit
Institute of Cancer Research
Sutton, Surrey, UK

First published 1987

Published by
THE MACMILLAN PRESS LTD
Houndmills, Basingstoke, Hampshire RG21 2XS
and London
Companies and representatives
throughout the world

Typeset by TecSet Ltd, Wallington, Surrey

Printed in Great Britain by
Camelot Press Ltd, Southampton

ISBN 0-333-33376-4
ISSN 0265-4377

The Contributors

G. I. Birnbaum
Division of Biological Sciences
National Research Council
OTTAWA K1A OR6
Canada

W. Fuller
Department of Physics
University of Keele
KEELE
Staffs ST5 5BG
UK

R. J. Greenall
Department of Physics
University of Keele
KEELE
Staffs ST5 5BG
UK

and
SERC Daresbury Laboratory
Daresbury
WARRINGTON WA4 4AD
UK

B. Lesyng
Department of Biophysics
Institute of Experimental Physics
University of Warsaw
Zwirki i Wigury 93
02-089 WARSAW
Poland

A. Mahendrasingam
Department of Physics
University of Keele
KEELE
Staffs ST5 5BG
UK

A. D. B. Malcolm
Department of Biochemistry
Charing Cross and Westminster Medical School
Fulham Palace Road
LONDON W6 8RE
UK

D. Shugar
Department of Biophysics
Institute of Experimental Physics
University of Warsaw
Zwirky i Wigury 93
02-089 WARSAW
Poland

C. K. Singleton
Department of Molecular Biology
Vanderbilt University
NASHVILLE
Tennessee 37235
USA

G. Snounou
National Institute for Medical Research
Mill Hill
LONDON NW7 1AA
UK

Preface

It is now five years since Volume 2 of *Topics in Nucleic Acid Structure* was published, and much of consequence has been established and published in this intervening period. This volume continues the pattern of the earlier ones of highlighting a few selected areas in the field that have seen especial progress, and that promise to be of continuing interest.

Variations in nucleic acid structure at a detailed level have been clearly shown by a number of single-crystal studies on short-length oligonucleotides. It is perhaps true that these major advances have tended to obscure the complementary information that continues to be provided by classic fibre diffraction techniques on polynucleotides. This area is reviewed by Fuller and Mahendrasingam who show that the power and scope of the technique is being greatly enhanced by the advent of high-flux X-rays from synchrotron sources. The controversy several years ago about the correctness of the Watson–Crick model for DNA compared to alternative 'side-by-side' ones is discussed by Greenall; he shows that these alternatives can be decisively rejected on the basis of X-ray diffraction considerations. The radically distinct left-handed Z-DNA structural polymorph has excited much attention, not least on account of its potential biological roles. These are reviewed by Singleton in relation to studies of the solution behaviour of Z-DNA.

The conformational flexibility of nucleotides is manifest in nucleic acid structures at both the oligomer and polymer levels. However, it remains true that as yet we still have an imperfect understanding of one of the major contributors to this flexibility – deoxyribose ring puckering changes. This topic is reviewed by Lesyng, who examines in detail the various factors contributing to particular pucker states. Nucleotide flexibility is also important when considering the functional features of modified nucleosides and nucleotides, a number of which have clinical importance as anticancer or antiviral agents by virtue of their abilities to inhibit critical enzymatic steps in nucleotide metabolism. Birnbaum and Shugar review conformational aspects of this large area, especially in relation to structure–activity relationships of enzyme interactions.

The advent of restriction enzymes has provided molecular biology with the tools to dissect, examine and manipulate biological DNA sequences at will, yet remarkably little is known about the mechanisms of specific DNA recognition by them. Malcolm and Snounou review this topic, which now that the first crystal structure of a restriction enzyme-oligonucleotide has been determined, may well be able to provide insights in more general aspects of protein-DNA interaction and recognition.

As ever, I am very grateful to many colleagues for advice and help, to the contributors for their efforts and patience, and to Harry Holt and Roger Osborne of The Macmillan Press for their continuing skills as publishers.

Sutton, Surrey, April 1987 S.N.

Contents

1
Biologically Active Nucleosides and Nucleotides: Conformational Features and Interactions with Enzymes

George I. Birnbaum and David Shugar

INTRODUCTION

The current widespread interest in analogues of nucleosides and nucleotides, and to some extent the aglycon moieties of these, as antimetabolic agents stems logically from the inference that they may inhibit one or more steps in the metabolic pathway(s) leading to the biosynthesis of nucleic acids, and that such inhibition may exhibit a certain degree of selectivity towards tumour cells or viruses (Herrmann, 1977; Cohen, 1979; Fields and Greene, 1982). A number of such heterocyclic bases and nucleoside analogues are, in fact, now widely employed as experimental antitumour and antiviral agents, with some of them approved for clinical use (Bloch, 1975; Herrmann, 1977; Collier and Oxford, 1980; Smith and Kirkpatrick, 1980; Smith *et al.*, 1980; Galasso, 1981; Shugar, 1981; Hirsch and Schooley, 1983; Stuart-Harris and Oxford, 1983; De Clercq and Walker, 1984; Shugar, 1984, 1985). At least in the case of antiviral agents, which will be emphasised in this chapter, a good deal of information has now been accumulated regarding their mechanism(s) of action, in large part because of the significant advances in our knowledge of the molecular biology of viral replication (Baltimore, 1976; Herrmann, 1977; Cohen, 1979; Fields and Greene, 1982). In particular, a relatively frequent requirement in the case of many nucleosides is the necessity for their prior intracellular phosphorylation (or 'activation') to the metabolically active forms.

Once a therapeutically active analogue has been found, subsequent research is directed not only to establishment of its mode of action, but also to its modification with a view to enhancement of its activity, or therapeutic index. By this method one seeks to determine structure-activity relationships (SAR), or quantitative SAR (QSAR) (Hansch, 1981, and references therein). Since the

therapeutically active form is generally that resulting from one or more modifications by intracellular enzymes, and since it manifests its activity by inhibition of some key enzymatic reaction (Cohen, 1979; De Clercq and Walker, 1984), it is clearly of interest to determine the role of the structure and conformation of a given analogue relative to those of the natural substrate. This is the object of the present review, such information being of obvious value in the clarification of the mode of action of a given analogue and in the design of more effective agents.

Considerable data have now been accumulated on the structures of both natural and synthetic analogues of purines and pyrimidines, and their nucleosides, both in the solid state and in solution, with the aid of X-ray diffraction, CD and NMR spectroscopy, and *ab initio* and semi-empirical theoretical calculations. The theoretical procedures are, of course, strictly applicable only to molecules *in vacuo* – a fact which must be borne in mind in assessing their applicability to a given problem. Each of these methods provides information about various aspects of structure and conformation, and usually more meaningful results are forthcoming from simultaneous application of more than one method. It should, however, be underlined that, while CD spectroscopy is a relatively simple procedure for following *changes* in conformation, its utility in the case of a nucleoside is somewhat limited, particularly when the aglycon is a purine or a purine analogue.

In contrast to the crystalline state, where conformational freedom is rather limited (but see Gavezzotti and Simonetta, 1982), a variety of conformations may, and does, exist in solution for molecules such as nucleosides and nucleotides, which possess a high degree of flexibility due to permissible rotations about single bonds and puckering of the five-membered furanose ring. NMR methods consequently provide information which describes the conformational equilibria in solution between two (or more) states and the predominant conformational form. The equilibrium for a given compound may be shifted by appropriate chemical modification.

The biological activity of a given analogue is frequently determined by, among other factors, its conformation, or rather by the changes of its conformation accompanying interaction with some cellular constituent, often an enzyme. Hence, the ideal approach for establishing these changes would be one where a conformational parameter may be 'fixed' by some chemical modification or, better still, by cocrystallisation of a substrate or inhibitor with the enzyme. Alternatively, in solution studies, one may follow, with the aid of CD and/or NMR spectroscopy, the changes in conformation resulting from such interactions.

Differentiation between conformational and other factors

A structural analogue may differ from its parent congener not only in conformation, but also in steric, electronic and other properties (including changes in

tautomeric equilibrium), and it is consequently necessary to differentiate between these. For example, 6-methylcytidine, which is constrained by the 6-methyl substituent exclusively to the *syn* conformation about the glycosidic bond (see p. 19), is not a substrate for the cytosine nucleoside deaminase of *Salmonella typhimurium* and of human lymphocytes (Krajewska and Shugar, unpublished). It is conceivable, in this case, that it is not the change in conformation, but the steric or other effects of the methyl group, which are responsible for the loss of substrate properties. In such instances it becomes of importance to establish whether the analogue is bound by the enzyme, in which case it will be an inhibitor of the reaction. If it is reasonably strongly bound, then it may be assumed that the change in conformation is not the major factor responsible for the loss of substrate properties, but that it is due to other effects resulting from the presence of the methyl group. If it is weakly bound, or not at all, then a straightforward answer is not forthcoming. The relative binding of an analogue and substrate may be determined from the values of the association constants K_i (for the analogue) and K_m (for the substrate). Various physicochemical techniques other than kinetics may be employed for measurements of binding constants. One may also use substituents with similar van der Waals' radii, but with different electronegative (or other) properties – e.g. CH_3 and Br.

Solid state and solution conformations

The extensive data now available on solid state structures of nucleosides and nucleotides have frequently been profited from in order to delineate the role of structure and conformation in the chemical reactivity, and substrate and/or inhibitory properties in some enzyme system, of a given analogue in solution (Sundaralingam, 1975; Saenger *et al.*, 1979). Such an approach must be treated with reserve, especially when based on one crystal structure. While diffraction methods provide the only accurate data on the spatial location of individual atoms in a molecule, it should be underlined that, for molecules with some degree of flexibility, the conformation so determined is usually *one* of a multitude of possible conformations in solution. There are instances where a given molecule may be crystallised in different polymorphic forms and/or with more than one independent molecule in the asymmetric unit. For example, 2′-fluoro-2′-deoxyguanosine crystallises with eight molecules in the asymmetric unit, each in a different conformation (Ikehara, personal communication), while 5-bromo-2′-deoxycytidine crystallises in a monoclinic and a triclinic form, with three and four molecules in the asymmetric unit, respectively (Low *et al.*, 1981a,b). The potent antiviral acycloguanosine (acyclovir) crystallises with three independent molecules in the asymmetric unit (Birnbaum *et al.*, 1981b, 1984a). While such examples provide useful information for analysis of conformational equilibria in solution, they do not, by themselves, provide any evidence for the conformation associated with the biologically active form.

The difficulty associated with the use of the foregoing information for deriving conclusions about the solution conformation may be partially circumvented

by making use of the known geometries of the large number of molecules, and defined molecular fragments, for which data have been classified. There are now about 50 000 published crystal structures of various organic and organometallic compounds (including over 500 nucleosides and nucleotides) which can be retrieved from the Cambridge Structural Database (Allen *et al.*, 1979), and studies on simple defined fragments have led to some general overall conclusions (see, for example, Murray-Rust, 1982). The probability of finding a given conformation in the crystal is dependent on its energy, and the most frequently found conformations correspond to global or local potential energy minima. Hence, if we have available a substantial number of crystal structures of related molecules, some generalisations may be drawn regarding the low-energy regions of the potential energy surface.

We therefore describe the conformations of nucleosides and nucleotides both in the solid state and in solution. The general features of these conformations furnish an essential tool in the determination of the mode of interaction of a given molecule with some protein or enzyme, either in solution or when cocrystallised in the solid state, which will be taken up in detail for specific cases.

In order to assess the role of conformation in the biological activity of a nucleoside, it is necessary to define the limits of various torsion angles within which the conformation may be regarded as 'normal'. In the following discussion, which is an extension of the brief summary presented in Chapter 1 of Part 1 of this series (Berman, 1981), we consider how the conformation is affected by various parameters, and to what extent some of the torsion angles are interdependent. The backbone torsion angles are designated according to recent recommendations of IUB–IUPAC:

$$P\text{-}O5'\overset{\beta}{}\text{-}C5'\overset{\gamma}{}\text{-}C4'\overset{\delta}{}\text{-}C3'\overset{\epsilon}{}\text{-}O3'\overset{\zeta}{}\text{-}P$$

As has become standard practice, we describe the puckering of the furanose ring in terms of P, the phase angle of pseudorotation (Altona and Sundaralingam, 1972); for puckers such as C3′-*endo*, C2′-*exo* and C1′-*exo*/C2′-*endo*, we occasionally employ the abbreviations ${}^3E$, ${}_2E$ and ${}_1T^2$, respectively. The glycosidic torsion angle χ is as given by Berman (1981).

CONFORMATION OF NUCLEOSIDES

Conformation of furanose rings

It is well known (Sundaralingam, 1975; de Leeuw *et al.*, 1980) that the most common conformations of the ribose and deoxyribose rings in nucleosides, in the solid state and in solution, are C2′-*endo* (${}^2E$) and C3′-*endo* (${}^3E$). With the crystallographic results available several years ago, it appeared that ribonucleosides exhibit a preference for C3′-*endo* puckering (Sundaralingam, 1975). Following acquisition of additional data, an analysis by de Leeuw *et al.* (1980)

revealed an almost equal distribution (55:47) of type *N* (^3E) and type *S* (^2E) conformers. By contrast, 2′-deoxyribonucleosides show a marked preference for type *S* puckering (18:4). On the basis of such data, one may infer that the ^2E and ^3E conformations represent true energy minima, and that the conformation in solution may be described by a C2′-*endo* ⇌ C3′-*endo* equilibrium (de Leeuw *et al.*, 1980). Indeed, most NMR and theoretical analyses of nucleoside conformations have taken only these two puckers into consideration (Davies, 1978a).

Several proposals have been advanced for three-state conformational equilibria for the furanose ring of some nucleosides in solution (Birnbaum *et al.*, 1979; Ekiel *et al.*, 1979b; Olson, 1981). The relevant data on which these are based were subsequently subjected to further analysis, with the use of a new empirical generalisation of the Karplus relation which takes into account the electronegativities and relative orientations of ring substituents (Haasnoot *et al.*, 1980), and a judicious use of X-ray crystallographic data. This led to an apparently satisfactory interpretation (de Leeuw and Altona, 1982) of the furanose ring pseudorotation on the basis of a two-state *N* ⇌ *S* equilibrium for β-D-ribo-, deoxyribo-, arabino-, xylo- and lyxonucleosides (Figure 1.1). However, the occurrence in crystal structures of ring conformations other than ^2E and ^3E – e.g. C1′-*exo*, C4′-*exo* and even O1′-*endo* – suggests that these may, in some instances, contribute to the conformational equilibrium in solution.

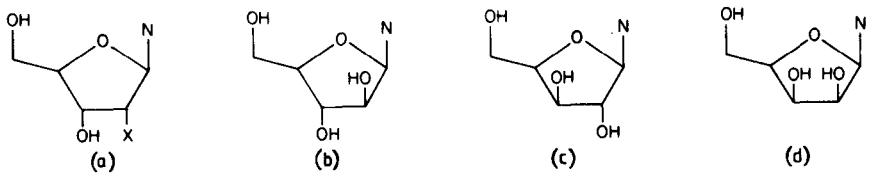


Figure 1.1 Structures of (a) ribo- (X = OH) and deoxyribo- (X = H), (b) arabino-, (c) xylo-, (d) lyxofuranose sugar moieties

Steric effects

The stabilities of the ^2E and ^3E conformers in ribonucleosides can be attributed, in part, to steric effects. When C2′ and/or C3′ are puckered out of the plane of the other three sugar ring atoms, the C2′–O2′ and C3′–O3′ bonds are staggered, so that the non-bonded interactions between the vicinal oxygens are reduced. With other ring conformations, these bonds are more eclipsed and are thus energetically less favourable. Attention was directed to this steric effect by Bishop and Cooper (1963) long before conformational analysis of sugars and nucleosides attained its present state of sophistication. However, if the steric effect were dominant, 2′-deoxynucleosides would be expected to adopt preferentially the C3′-*endo* conformation, in which the 3′-OH group is oriented

pseudoequatorially. In fact, as mentioned above, this turns out not to be the case. Steric interactions were recently examined by Klimke *et al.* (1980a,b), who found no quantitative correlation between the pentose ring conformations of various substituted araA and xyloA analogues and the steric requirements of the substituents. Their results show some correlation between the sugar ring conformations and the van der Waals' radii of the ring substituents. However, this correlation is neither linear nor very pronounced.

Electronegativity effects

Alternative interpretations of the influence of substituents on sugar ring conformation have been advanced by Uesugi *et al.* (1979) and by Guschlbauer and Jankowski (1980). The former examined a series of 2′-substituted 2′-deoxyadenosines and observed a linear relationship between the electronegativity of the substituent and the chemical shift of H1′. While the line fitted the points for H, I, Br, Cl, N_3, OMe and OH, there were very substantial deviations from linearity for NH_2 and F. A linear relationship was also found between the electronegativities of I, Br, Cl, N_3 and F and the mole fraction of the *N* conformer, the largest electronegativity (for F) corresponding to the highest proportion (0.67) of the N conformer; but, in this instance, the points for H, NH_2, OH and OMe did not fit the line. Nevertheless, it was concluded that it is feasible to predict the sugar ring conformation of adenine nucleoside analogues on the basis of the electronegativity of the 2′-substituent. It was further suggested that this relationship should hold for any pentofuranose moiety.

Guschlbauer and Jankowski (1980) carried out NMR analyses of 2′-substituted uridines and found that the electronegativities of the substituents are linearly related to the chemical shifts of C2′ and the coupling constants $^1J_{C2',H2'}$. Having also found that the mole fraction of the *N* conformer increases approximately linearly with increasing electronegativity of the 2′-substituent, they concluded that the polarity of the C2′–X bond has a dominant influence on the conformation of the pentose ring, because the most electronegative substituent pulls the ring carbon atom to which it is attached towards itself. They further proposed that 'the difference in electronegativity of H and OH is the dominating force in the differences between DNA and RNA'. The authors considered it desirable to investigate the theoretical basis for their findings, but did not themselves offer any explanation.

The gauche *effect*

As pointed out by Haasnoot *et al.* (1981) and by Olson (1982), what was observed by Uesugi *et al.* (1979) and Guschlbauer and Jankowski (1980) is actually the well-known '*gauche*-effect', earlier described by Wolfe (1972) as a 'tendency to adopt that structure which has the maximum number of *gauche* interactions between the adjacent electron pairs and/or polar bonds'. In ribonucleosides there are three such pairs of polar bonds on the sugar ring, if we exclude the exocyclic side chain, viz. O1′–C4′–C3′–O3′, O1′–C1′–C2′–O2′

and O2′–C2′–C3′–O3′. In both the ^2E and ^3E conformations, two of these pairs are *gauche* oriented, so that there is no preference for type *N* or type *S*. By contrast, in deoxyribonucleosides, where only the first of these pairs is present, the conformation is *gauche* for the ^2E forms and *trans* for the ^3E conformers. Hence the preference for type *S* puckering. This analysis can be extended to other pentose rings of nucleosides, with results listed in Table 1.1. From this table it may be readily inferred that deoxyribo- and arabinonucleosides prefer the type *S* sugar pucker, and the xylonucleosides type *N*, whereas in ribo- and lyxonucleosides the two conformations should be equally populated. These conclusions are similar to those reached on the basis of electronegativity, as discussed above.

The validity of the foregoing predictions for ribo- and deoxyribonucleosides has already been referred to. We can now examine the conformational preferences of nucleosides with other pentose moieties, both in solution and in the solid state. The results of several recent conformational analyses are, however, somewhat conflicting. A detailed ^{1}H NMR investigation of β-arabinonucleosides (Ekiel *et al.*, 1979b) indicated that araC and araU exhibit a slight preference for the type *S* conformation, while the opposite was taken to hold for araA. Klimke *et al.* (1980a) found that, at 40°C in ND_3, the *N* and *S* states are approximately equally populated, and that lowering the temperature to −60°C shifted the equilibrium in the direction of the type *N* state (from 0.53 to 0.65). From semi-empirical energy calculations, Yathindra and Sundaralingam (1979) concluded that the type *N* state is the most stable conformation, but no values were given.

On the other hand, similar studies of xyloA and a number of its derivatives (Ekiel and Shugar, 1979; Klimke *et al.*, 1980b) demonstrate a definite preference for the type *N* pucker, in accordance with the data in Table 1.1. Finally, lyxonucleosides also appear to prefer the type *N* conformation (Ekiel *et al.*, 1979a) – a result which cannot be attributed to the *gauche* effect.

In the antibiotic cordycepin (3′-deoxyadenosine) there is only one C–O/C–O interaction and, as expected, the preferred conformation of the pentose ring is C3′-*endo*, both in the crystal (Radwan and Wilson, 1980) and in solution

Table 1.1 *Gauche* and *trans* interactions in furanose rings

		O1′–C4′–C3′–O3′	O1′–C1′–C2′–O2′	O2′–C2′–C3′–O3′
Ribose type *N*,	^3E	*trans*	*gauche*	*gauche*
type *S*,	^2E	*gauche*	*trans*	*gauche*
Deoxyribose	^3E	*trans*		
	^2E	*gauche*		
Arabinose	^3E	*trans*	*trans*	*gauche*
	^2E	*gauche*	*gauche*	*trans*
Xylose	^3E	*gauche*	*gauche*	*trans*
	^2E	*trans*	*trans*	*gauche*
Lyxose	^3E	*gauche*	*trans*	*gauche*
	^2E	*trans*	*gauche*	*gauche*

(Westhof *et al.*, 1977), where the mole fraction of the type *N* state was found to be 0.92. On the other hand, a molecular orbital study, using the perturbative configuration interaction over localised orbitals (PCILO) method, did not indicate whether the ^3E or ^2E pucker is preferred (Saran and Patnaik, 1981).

The influence of the *gauche* effect may be further evaluated by examination of nucleoside analogues in which a C–O bond is replaced by the more polar C–F bond or the much less polar C–H bond. Some of these analogues exhibit antimetabolic (and chemotherapeutic) activities. As already mentioned above, NMR studies have shown that the population of the type *N* state is lower in 2′-deoxyuridine than in uridine, while it is higher in the 2′-deoxy-2′-fluoro analogue (Guschlbauer and Jankowski, 1980), in which there are two C–F/C–O *gauche* interactions. Similar observations have been made on the solid state structures of ribonucleosides and their 2′-deoxy and 2′-deoxy-2′-fluoro analogues (Hakoshima *et al.*, 1981). Furthermore, in the solid state structure of 2′-deoxy-2′-fluorocytidine, the sugar ring exhibits the expected C3′-*endo* pucker, with the C2′–F bond *gauche* to both the C1′–O1′ and C3′–O3′ bonds; whereas in 2′-deoxy-2′-fluorouridine the furanose ring was found to be in the unusual C4′-*exo*/O1′-*endo* conformation, ascribed to a short contact between the fluorine atom and the 3′-OH (Marck *et al.*, 1982).

A remarkable manifestation of the *gauche* effect is provided by the recent X-ray analysis of 2′-deoxy-2′-fluoroguanosine (Ikehara, personal communication) in which eight independent molecules were found in the asymmetric unit. The pseudorotation angles (*P*) in the furanose rings are as follows: 4, 12, 14, 27, 29, 44, 61 and 74°. Thus, while the puckers range from ^3_2T to ^0_4T, they are nevertheless all in the type *N* region.

When the –F substituent is *trans* to a vicinal –OH group (i.e. in derivatives of arabino- and xylonucleosides), the *gauche* effect will play only a minor role. An example is provided by 3′-deoxy-3′-fluoro-araA. In ND_3 solution at 40°C, this nucleoside occurs preferentially (70 per cent) in a type *S* state (Klimke *et al.*, 1980a). While this conformation is favoured because of the enhanced *gauche* interaction between the C3′–F and C4′–O1′ bonds, it also leads to C3′–F being *trans* to C2′–OH – an orientation which is not favoured (Wolfe, 1972). Surprisingly, Klimke *et al.* (1980a) ascribed the 'strong preference for the *S* state' to a *favourable trans* orientation of the two dipoles (C2′–OH and C3′–F), for which they calculated an interaction energy hundreds of J/mol lower than for the *N* state, in which these two bonds are in a *gauche* orientation. It is more likely, however, that the observed conformational equilibrium is due to the *gauche* interaction C1′–O1′/C2′–O2′. It should be pointed out that the electronegativity of an *O*-alkyl group (O1′–C4′) is only slightly different (lower) from that of an O–H. This at least partially accounts for the observation that replacement of a sugar O–H by O–CH_3 affects the ring conformation to only a minor extent (Remin and Shugar, 1973) – a finding of some significance in view of the fact that 2′-*O*-methyl nucleosides are widely encountered in tRNAs and, to a minor extent, in ribosomal RNA (Hall, 1971).

Relevant to the foregoing are conformational analyses of 2′-deoxy-2′-fluoro-5-iodo-arabinosylcytosine (FIAC; Figure 1.2), one of the more potent reported antiherpes agents (Watanabe *et al.*, 1979). On the assumption that the strongly electronegative fluorine atom would pull C2′ 'up', Haertlé *et al.* (1979) predicted the extreme *S* form for fluorinated arabinosyl analogues. However, an analysis of *gauche* interactions reveals that in the ^{2}E conformation the C2′–F bond is *gauche* to C1′–O1′ and *trans* to C3′–O3′, while in the ^{3}E pucker the situation is reversed (Figure 1.2). Thus, there is no strong preference for either conformation. In fact, a conformational analysis, based on 270 MHz ^{1}H NMR spectra, revealed a 50:50 mixture of the C2′-*endo* and C3′-*endo* conformers (Lipnick and Fissekis, 1980). An X-ray analysis of FIAC showed that in the solid state the sugar ring adopts the conformation C3′-*endo*/C2′-*exo* ($^{2}_{3}T$) (Birnbaum *et al.*, 1982b), possibly as a result of intermolecular hydrogen bonding (see below).

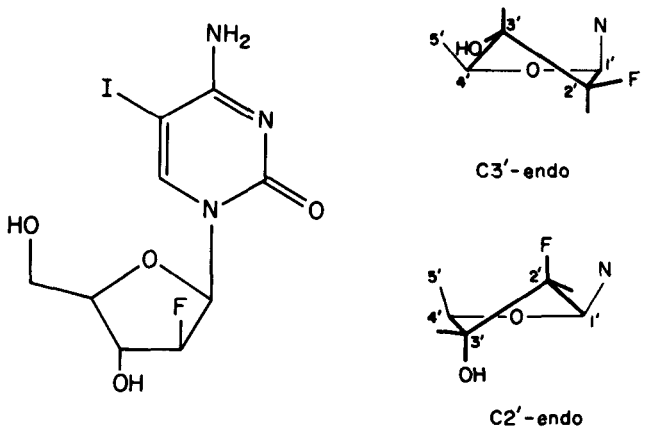


Figure 1.2 Structure of FIAC (*left*), and (*right*) the C3′-*endo* and C2′-*endo* conformations of its sugar ring. Reprinted with permission from Birnbaum *et al.* (1982b). Copyright 1982 American Chemical Society

Summing up, the conformation of the furanose ring in a nucleoside is influenced by the *gauche* effect, to an extent determined by the electronegativities of the ring substituents. Its approximate magnitude may be assessed on the basis of a study of the *gauche* effect in 1,2-disubstituted cyclohexanes (Zefirov *et al.*, 1978), which showed that the stabilisation due to this effect is approximately 2.73 kJ/mol in the case of vicinal C–O/C–O bonds and about 5.5 kJ/mol for C–O/C–F bonds.

Hydrogen bonding

Crystallographic studies have shown that the conformation of the pentofuranose ring in a nucleoside may be affected by hydrogen bonding, either inter- or intramolecular. In the former case, a change in ring pucker (pseudorotation) may facilitate formation of a hydrogen bond between an –OH substituent and a neighbouring molecule, thus stabilising the packing of the molecules in the crystal lattice. The extent of such packing forces is, by and large, unpredictable and would be expected to vary from structure to structure. There is, in fact, no evidence that this factor is of systematic significance in determining the conformations of nucleosides.

By contrast, the influence of *intra*molecular hydrogen bonding on the sugar ring pucker is well documented. Thus, an intramolecular O5′–H···N3 hydrogen bond (Figure 1.3), often observed in purine nucleosides with a *syn* conformation about the glycosidic bond, is usually accompanied by a C2′-*endo* pucker of the ribofuranosyl ring (Birnbaum and Shugar, 1978). Such a conformation, with associated hydrogen bonding, has also been observed in solution (Plochocka *et al.*, 1977; Koole *et al.*, 1986). It has been suggested (Silverton *et al.*, 1982) that this correlation of C2′-*endo* and *syn* conformations is valid only in the presence of the O5′–H···N3 hydrogen bond. Nevertheless, this conformation has also been found in several purine nucleosides in which this hydrogen bond is not present (Nagashima and Wakabayashi, 1974; Al-Mukhtar and Wilson, 1978; Ginell and Parthasarathy, 1978; Stolarski *et al.*, 1980b).

With xylonucleosides a different intramolecular hydrogen bond, O3′–H···N3 (Figure 1.3), is sterically feasible and has, indeed, been observed in an 8-bromo-

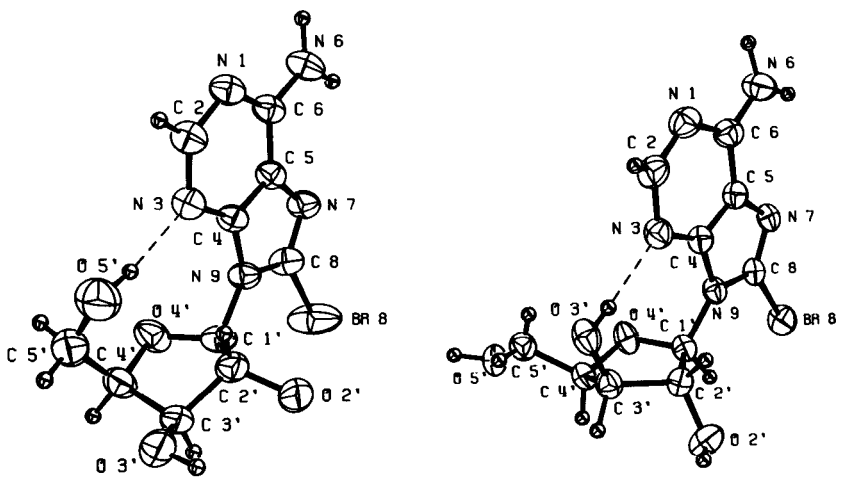


Figure 1.3 Examples of intramolecular hydrogen bonding in 8-bromo-9-β-D-xylofuranosyladenine: (*left*) O5′–H···N3; (*right*) O3′–H···N3. Reprinted with permission from Birnbaum *et al* (1982a). Copyright 1982 American Chemical Society